interactive Journal of Medical Research

Impact Factor (2022): 2.0
Volume 1 (2012), Issue 2 ISSN 1929-073X

Contents

Corrigenda and Addenda

Original Papers

Review

Original Paper

Corrigenda and Addenda

Correction: Improving Interoperability in ePrescribing

Sten-Erik Öhlund[1,2,3], BSC, PhLic; Bengt Åstrand[2], MPharm, PhD; Göran Petersson[2], MD, PhD

[1]Logica Sverige AB, Kalmar, Sweden

[2]eHealth Institute, Linnaeus University, Kalmar, Sweden

[3]Information Systems, Department of Management and Engineering, Linköping University, Linköping, Sweden

Corresponding Author:
Sten-Erik Öhlund, BSC, PhLic
Logica Sverige AB
Skeppsbrogatan 49
Kalmar, 391 23
Sweden
Phone: 46 0733982117
Fax: 46 480 8557
Email: sten.erik.ohlund@logica.com

Related Article:

Correction of: http://www.i-jmr.org/2012/2/e17/

(Interact J Med Res 2012;1(2):e19) doi:10.2196/ijmr.2457

In the recent paper by Öhlund SE et al. [Improving Interoperability in ePrescribing, Interact J Med Res 2012;1(2):e17], the percentage of prescriptions in the post-NEF (national ePrescription format) sample with errors should be 0.9% (13,735/1,479,588), rather than 0.009%. The change is minor and does not affect the overall results and conclusions of the paper. The percentage was corrected in 3 places in the paper: 1) Abstract/Results: "In the post-NEF sample, only 0.9% of the prescriptions had errors"; 2) Results/Errors per Prescription and Prescription Set: "The percentage of post-NEF prescription sets with at least one error was 0.9% (13,735/1,479,588)"; 3) Table 7 forth row: "Prescription sets with error, % 98.6 0.9". The online version on i-JMR was corrected on November 29, 2011, one week after publication, before submission to PubMed Central, Medline, or other databases.

Edited by G Eysenbach; submitted 27.11.12; this is a non–peer-reviewed article;accepted 27.11.12; published 29.11.12.

Please cite as:
Öhlund SE, Åstrand B, Petersson G
Correction: Improving Interoperability in ePrescribing
Interact J Med Res 2012;1(2):e19
URL: http://www.i-jmr.org/2012/2/e19/
doi:10.2196/ijmr.2457
PMID:23612367

Original Paper

Assessing and Comparing Information Security in Swiss Hospitals

Sarah Landolt[1], MSc; Jürg Hirschel[2]; Thomas Schlienger[3], PhD; Walter Businger[4], PhD; Alex M Zbinden[5], MD, PhD

[1]AAC Infotray, Winterthur, Switzerland

[2]Ypsomed AG, Burgdorf, Switzerland

[3]TreeSolution Consulting GmbH, Hünenberg, Switzerland

[4]Engineering and Information Technology, Bern University of Applied Sciences, Bern, Switzerland

[5]Medical Technology Center, Bern University of Applied Sciences, Bern, Switzerland

Corresponding Author:
Sarah Landolt, MSc
AAC Infotray
Zur Kesselschmiede 31
Winterthur, CH-8400
Switzerland
Phone: 41 52 260 3162
Fax: 41 52 203 6929
Email: sarah.landolt@alumni.ethz.ch

Abstract

Background: Availability of information in hospitals is an important prerequisite for good service. Significant resources have been invested to improve the availability of information, but it is also vital that the security of this information can be guaranteed.

Objective: The goal of this study was to assess information security in hospitals through a questionnaire based on the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) standard ISO/IEC 27002, evaluating *Information technology – Security techniques – Code of practice for information-security management*, with a special focus on the effect of the hospitals' size and type.

Methods: The survey, set up as a cross-sectional study, was conducted in January 2011. The chief information officers (CIOs) of 112 hospitals in German-speaking Switzerland were invited to participate. The online questionnaire was designed to be fast and easy to complete to maximize participation. To group the analyzed controls of the ISO/IEC standard 27002 in a meaningful way, a factor analysis was performed. A linear score from 0 (not implemented) to 3 (fully implemented) was introduced. The scores of the hospitals were then analyzed for significant differences in any of the factors with respect to size and type of hospital. The participating hospitals were offered a benchmark report about their status.

Results: The 51 participating hospitals had an average score of 51.1% (range 30.6% - 81.9%) out of a possible 100% where all items in the questionnaire were fully implemented. Room for improvement could be identified, especially for the factors covering "process and quality management" (average score 1.3 ± 0.8 out of a maximum of 3) and "organization and risk management" (average score 1.3 ± 0.7 out of a maximum of 3). Private hospitals scored significantly higher than university hospitals in the implementation of "security zones" and "backup" (P = .008).

Conclusions: Half (50.00%, 8588/17,177) of all assessed hospital beds in German-speaking Switzerland are in hospitals that have a score of 49% or less of the maximum possible score in information security. Patient data need to be better protected because of the data protection laws and because sensitive, personal data should be guaranteed confidentiality, integrity, and availability.

(Interact J Med Res 2012;1(2):e11) doi:10.2196/ijmr.2137

KEYWORDS

information security; information protection; computer security standards; electronic health records organization & administration; hospital information systems; Switzerland.

Introduction

Information management, especially in emergency medicine, enhances the instantaneous and ubiquitous availability of digital patient records and can significantly improve clinical practice [1]. On the other hand, poor patient data security represents a major problem that must be addressed with more sophisticated hospital information technology (HIT) [2], but the protection

of information represents a growing challenge [3]. For example, it is increasingly difficult to safeguard the integrity of digital radiology images and protect them from unauthorized manipulation [4]. Furthermore, the growing integration of complex hospital information systems [5], the widespread use of mobile devices [6], and the increasing amount of communication between health care providers require special attention regarding information security.

To implement an adequate information-security management system, it is first necessary to evaluate information security and assess its risks, and subsequently to find suitable measures to control risks and improve security measures [7]. TheInternational Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) have defined the international standards for information and data security (ISO/IEC 2700x, *Information technology – Security techniques*) that are widely accepted and can be used to evaluate levels of security [8]. The standards identify three main components of information security: confidentiality, availability, and integrity. They also describe requirements for an information-security management system (ISO/IEC 27001), a code of practice (ISO/IEC 27002) [9], implementation guidelines (ISO/IEC 27003), parameters to be measured (ISO/IEC 27004), and risk management (ISO/IEC 27005).

Education is an important component of successful management of information security [10]. To determine appropriate actions and education efforts, chief information officers (CIOs) need to know the status quo in their organization and have both a measuring tool and benchmark values at their disposal. However, no study has compared hospitals with respect to information security. This might be because information about the security level of an institution is delicate and might influence the hospital's perceived trustworthiness or that assessing it might itself be a security threat. The lack of an effective benchmark tool for the assessment of the status quo of information security may be another explanation for the absence of such comparisons: The comprehensive and time-intensive character of commercially available tools, such as Verinice [11], rules out their use for a widescale comparison of hospitals.

Switzerland has a national implementation strategy for efficient and safe eHealth systems in which, for reasons of legal rights and acceptability, information security plays a central role [12]. The goal of this study is to evaluate the current status of information security in Swiss hospitals. As a first step, an ISO/IEC 27002-compliant tool that allows for both a rapid nationwide assessment of hospital security and the provision of benchmark data for CIOs was developed. By using this tool, the present investigation aims to evaluate information security focusing on differences between hospitals of different sizes and types (ie, private vs public hospitals and academic vs non-academic hospitals).

Methods

Questionnaire

The goal was to develop an online questionnaire that covered most chapters of the ISO/IEC standard 27002, *Information technology – Security techniques – Code of practice for information-security management*, and required less than 20 minutes to fill out. The online "EFS Survey" tool [13] was used to design and host the questionnaire. The questionnaire incorporated 24 parameters defined in the ISO/IEC 27002 standard, *Information technology – Security techniques – Code of practice for information-security management*, with some parameters combined into one question (see Appendix 1 for the complete set of questions with the corresponding chapters of the standard). All questions were identically structured and had four possible answers: (1) "unknown/not implemented," (2) "partially implemented," (3) "completely implemented," and (4) "completely implemented and continuously monitored and improved." The same layout and order of answers was used for all questions to reduce visual complexity. The questionnaire consisted of 19 screens, with 2 questions displayed on each screen. Questions regarding general parameters of the hospital, such as type of hospital, number of beds, number of full-time equivalents (FTEs) of job positions, total number of employees, number of FTEs in information technology (IT), total number of employees in IT, and number of computer workstations at the hospital, were assessed before the actual questionnaire. The questionnaire was reviewed by several national experts from the fields of medical informatics and information security not directly involved in the survey. To ensure technical functionality, the questionnaire was comprehensively tested by three test participants prior to its distribution.

The CIOs of all 112 hospitals in the German-speaking portion of Switzerland were informed via email about the planned study 6 months before its inception. They were also informed that the study would be conducted by students of a Master of Advanced Studies in Medical Informatics at the Berne University of Applied Science. The survey was announced a second time in a personal letter [14] and, 2 weeks later, a third time via a personal email that contained a hyperlink with a personal key. Participants were informed in the correspondence that the survey should take a maximum of 15 to 20 minutes to complete and that confidential treatment of data was guaranteed.

To encourage timely responses, a genographic test from National Geographic [15] was offered as draw prize to one of the first 50 participants. Furthermore, the participants were ensured a detailed benchmark analysis. Participants were asked to respond within 2 months. To improve the response rate, two reminder emails were sent out 10 days after this 2-month period and again 1 week later.

The participants gave informed consent by affirming the opening question of the questionnaire: "Do you agree to participate in our survey? And do you give your consent that the data may be published in an anonymized form?"

There was a unique key included in the individual hyperlinks sent out to participants to access the survey tool. The key was logged by the tool and exported with the data; therefore, duplicate entries from the same user were precluded. The tool was configured to continue only after a question was completed. A back button allowed for corrections.

Data with personal information were stored in encrypted form when exported from the survey tool. Persons involved in the

statistical evaluation were blinded and worked with anonymized data. A monitoring group was in charge of protecting the data and the interests of the participants. All students involved in the investigation were required to sign a confidentiality agreement.

Statistical Analysis

Scores were introduced to perform a statistical analysis of the data collected. A linear score from 0 (answer 1) to 3 (answer 4) was used, as shown in Table 1. The higher a hospital's overall score, the more sophisticated its data security management.

Table 1. The four possible answers to questionnaire items and the assigned score points.

Answer	Score points
1. Unknown, not implemented	0
2. Known, partially implemented	1
3. Completely implemented	2
4. Completely implemented, under continuous improvement	3

Hospitals were classified into (1) academic (university) hospitals with a research mandate from the state, (2) non-academic public hospitals with an emergency ward, (3) rehabilitation clinics, and (4) private hospitals. Furthermore, the hospitals were split into two groups based on hospital size (ie, hospitals with > 150 beds and hospitals with ≤ 150 beds.

For data reduction, a factor analysis with varimax rotation and Kaiser normalization using SPSS version 15 (SPSS Inc, Chicago, IL, USA) was applied to group-related questions into independent factors.The Kaiser normalization eliminated all components with eigenvalues under 1.0, thus extracting 7 reasonable factors. No further cutoff criteria for determining the optimal number of factors were explored.

Since normal distribution could not be shown using a Shapiro-Wilk test, a two-way non-parametric analysis of variance (Friedman test) was performed for both the type of risk factors determined by the factor analysis and the group of hospitals. The difference between hospital types was then determined using pairwise testing with Bonferroni's corrections. Subsequently, the influence of each of the 7 risk factors on the differentiation between hospital types was calculated using a non-parametric one-way analysis of variance (Kruskal-Wallis test). The effect of hospital size (number of beds) was determined using a Friedman test, also taking into account the 7 risk factors.

Results

Questionnaire

Of the 112 CIOs invited to participate in the survey, 69 (61.6%) responded. Of these, 7 did not give informed consent, 9 aborted the questionnaire while answering the general questions about the hospital, and 2 aborted the questionnaire while answering the questions about information security. In Table 2, "responded," "participated," and "completed" indicate that the survey page was visited, that the informed consent page was filled out, and that the questionnaire was fully completed, respectively. Thus, there was a 90% (62/69) participation rate and a 74% (51/69) completion rate [16]. Only the 51 completed datasets were used for further analysis.

Table 2. Analysis of the number and percentage of returned questionnaires with respect to hospital type and hospital size.

Group	Invited		Responded		Participated		Completed	
	n	%	n	%	n	%	n	%
Hospital type (total)	112	100%	69	62%	62	90%	51	74%
University hospital	11	100%	9	82%	6	67%	4	44%
Public hospital	54	100%	39	72%	36	92%	29	74%
Rehabilitation clinic	13	100%	7	54%	7	100%	6	86%
Private hospital	34	100%	14	41%	13	93%	12	86%
Hospital size (total)	112	100%	69	62%	62	90%	51	74%
≤ 150 beds	45	100%	20	44%	18	90%	16	80%
> 150 beds	67	100%	49	73%	44	90%	35	71%

Of the 51 hospitals in which a CIO had completed the questionnaire, 4 were university hospitals, 29 were public hospitals, 6 were rehabilitation clinics, and 12 were private hospitals.The hospitals which completed the questionnaire had total scores ranging from 30.6% to 81.9% out of a maximum score of 100%. These scores are presented in Table 3 for the two hospital sizes and for the four hospital types.

To visualize the overall distribution of information security for hospitals in German-speaking Switzerland, each hospital's individual score was calculated as a percentage of the maximum score. These percentages are shown in Figure 1 as functions of the number of beds in a hospital. Additionally, a least squares regression curve was laid over the cumulated scores. The curve characterizes the distribution of information security per hospital

bed. The 50% line shows that, according to the regression curve, 50% of all hospital beds reached a score of 49.2% or less of the maximum score.

The factor analysis extracted 7 factors, explaining 70% of the total variance. The questions were assigned to the factor with the highest correlation (Table 4).

The grouping of the questions into factors gave interesting insights into their relationship and made it possible to assign a group term to each of the 7 groups of questions (Table 5).

Table 3. Scores for each hospital type and for the different hospital sizes.

Group	Average score		Minimum score		Maximum score	
	Mean	%	Mean	%	Mean	%
Hospital type (total)	36.8	51.1%	22	30.6%	59	81.9%
University hospital	32.8	45.5%	24	33.3%	40	55.6%
Public hospital	36.4	50.6%	22	30.6%	59	81.9%
Rehabilitation clinic	35.2	48.8%	23	31.9%	56	77.8%
Private hospital	39.9	55.4%	30	41.7%	53	73.6%
Hospital size (total)	36.8	51.1%	22	30.6%	59	81.9%
≤ 150 beds	36.8	51.1%	22	30.6%	58	80.6%
> 150 beds	36.8	51.2%	23	31.9%	59	81.9%

Table 4. Results of the factor analysis (rotated component matrix).

1. Process and quality management							
5. Classification of information	0.79[a]	–0.08	0.19	–0.07	–0.05	0.01	0.27
7. Awareness and end-user training	0.45[a]	0.22	0.25	0.25	0.21	0.27	–0.11
9. Documented business processes	0.57[a]	–0.06	–0.01	0.36	–0.11	0.11	0.13
21. Security incidents reporting	0.70[a]	0.30	0.19	0.01	0.11	0.09	–0.08
22. Learning from incidents	0.81[a]	0.23	0.05	0.26	0.15	0.13	–0.11
23. Ensuring hospital business continuity	0.69[a]	0.10	0.21	–0.23	0.12	–0.09	0.42
2. Access control and procurement							
13. Policies for handling mobile storage devices	0.02	0.54[a]	0.14	0.18	0.14	0.43	–0.01
15. User management and access rights	0.13	0.77[a]	0.06	0.10	–0.10	0.20	0.17
16. Remote access control	–0.02	0.82[a]	0.17	–0.16	0.30	0.09	–0.03
18. Secure procurement	0.28	0.69[a]	0.02	0.27	0.21	–0.01	0.19
3. Organization and risk management							
1. Security-risk analysis	0.38	0.07	0.58[a]	–0.24	–0.09	0.00	0.03
2. Information-security policies	0.09	0.08	0.72[a]	0.42	0.21	0.10	0.01
3. Management commitment	0.10	0.09	0.88[a]	0.08	0.01	0.21	0.00
4. IT inventory and data ownership	0.38	0.17	0.50[a]	0.47	–0.31	–0.06	–0.08
4. Control and monitoring							
6. Employment-contract rules	0.22	0.07	0.07	0.61[a]	0.02	0.49	–0.13
14. Monitoring	–0.10	0.13	0.14	0.74[a]	0.22	–0.05	0.26
17. System-login security	0.34	0.44	–0.13	0.46[a]	0.05	0.12	0.36
5. Attack protection							
11. Malware protection	0.35	0.36	–0.30	0.16	0.56[a]	0.07	–0.09
20. Patch management	–0.30	0.20	–0.09	0.15	0.64[a]	0.25	0.17
24. Security assessments	0.30	0.06	0.40	–0.04	0.65[a]	0.05	–0.09
6. Encryption and staging							
10. Staging (separation of development, test, and productive environment)	0.42	0.27	0.13	0.04	0.17	0.57[a]	–0.14
19. Encryption of mobile data	–0.02	0.15	0.11	–0.01	0.11	0.84[a]	0.18
7. Backup and security zones							
8. Security zones	0.34	0.22	0.10	0.20	0.47	0.29	0.50[a]
12. Backup	0.06	0.12	–0.04	0.12	–0.02	0.03	0.86[a]

[a]highest correlation value per question.

Table 5. Terms given to the seven factor groups of questions.

Factor	Term
Factor 1	Process and quality management
Factor 2	Access control and procurement
Factor 3	Organization and risk management
Factor 4	Control and monitoring
Factor 5	Attack protection
Factor 6	Encryption and staging
Factor 7	Backup and security zones

The null hypothesis that all hospital types reach the same scores could be rejected ($P < .05$) on the basis of the Friedman test for all 28 groups (7 risk-factor groups and 4 types of hospitals). With the Kruskal-Wallis test, a significant difference ($P < .05$) in factor 7 (backup and security zones) between hospital types was found, with university hospitals ranking lowest and private hospitals highest (Figure 2). Using (conservative) pairwise testing and Bonferroni's correction, however, no significant difference was found ($P = .02$ which is greater than .05/7 = .0071). No significant effect was observed with respect to hospital size (Figure 3).

Figure 1. Cumulated scores by hospital beds.

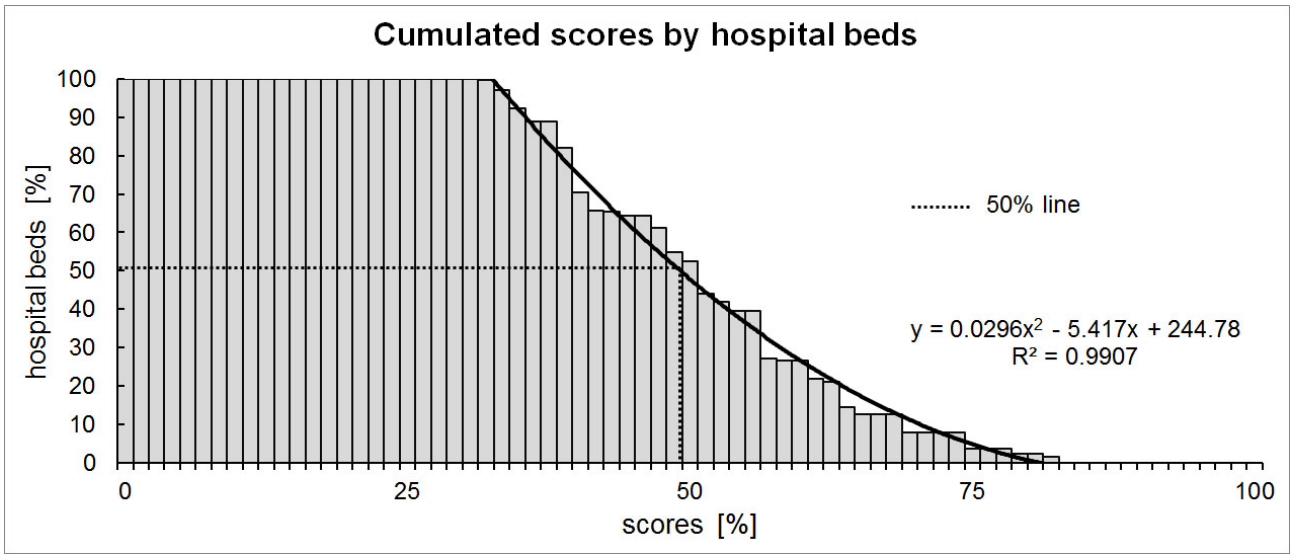


Figure 2. Scores by hospital groups.

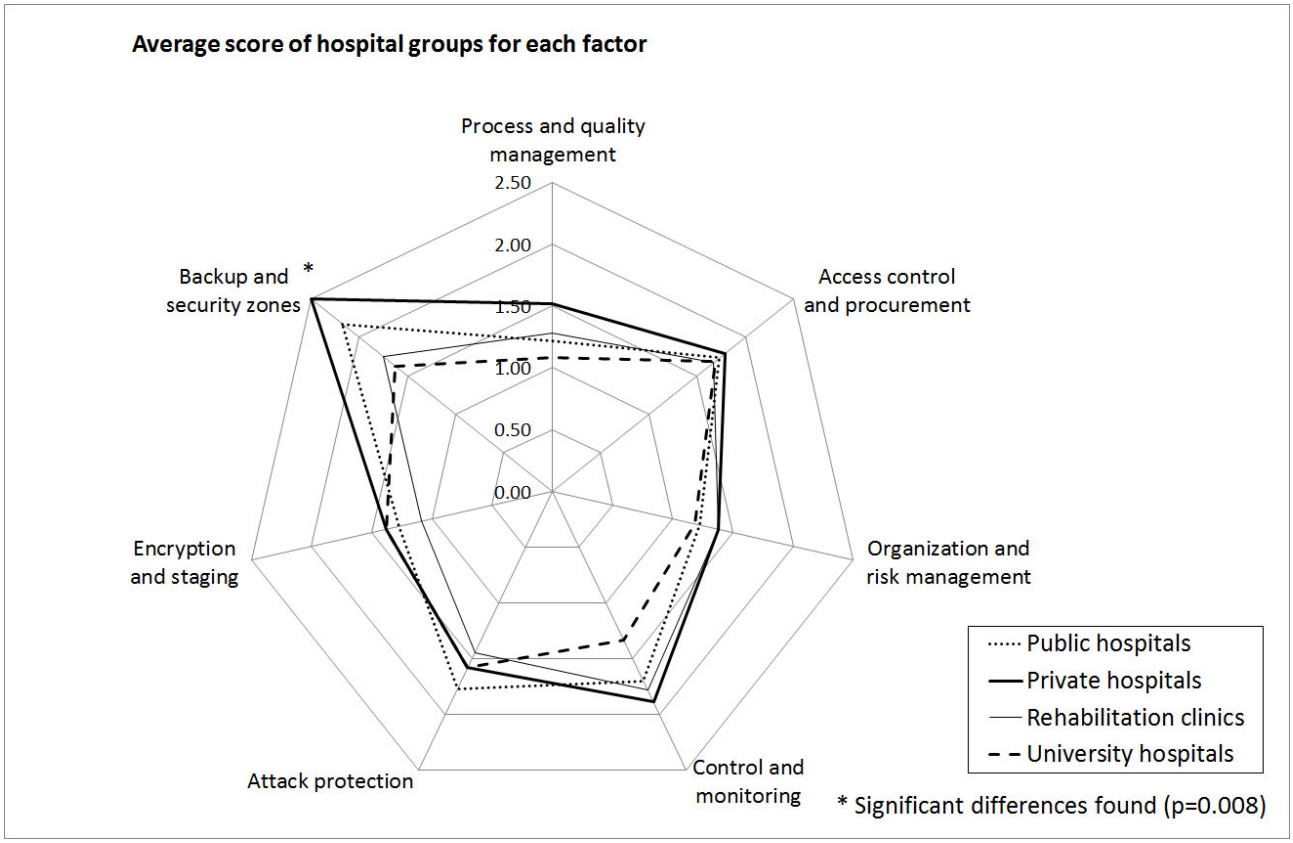
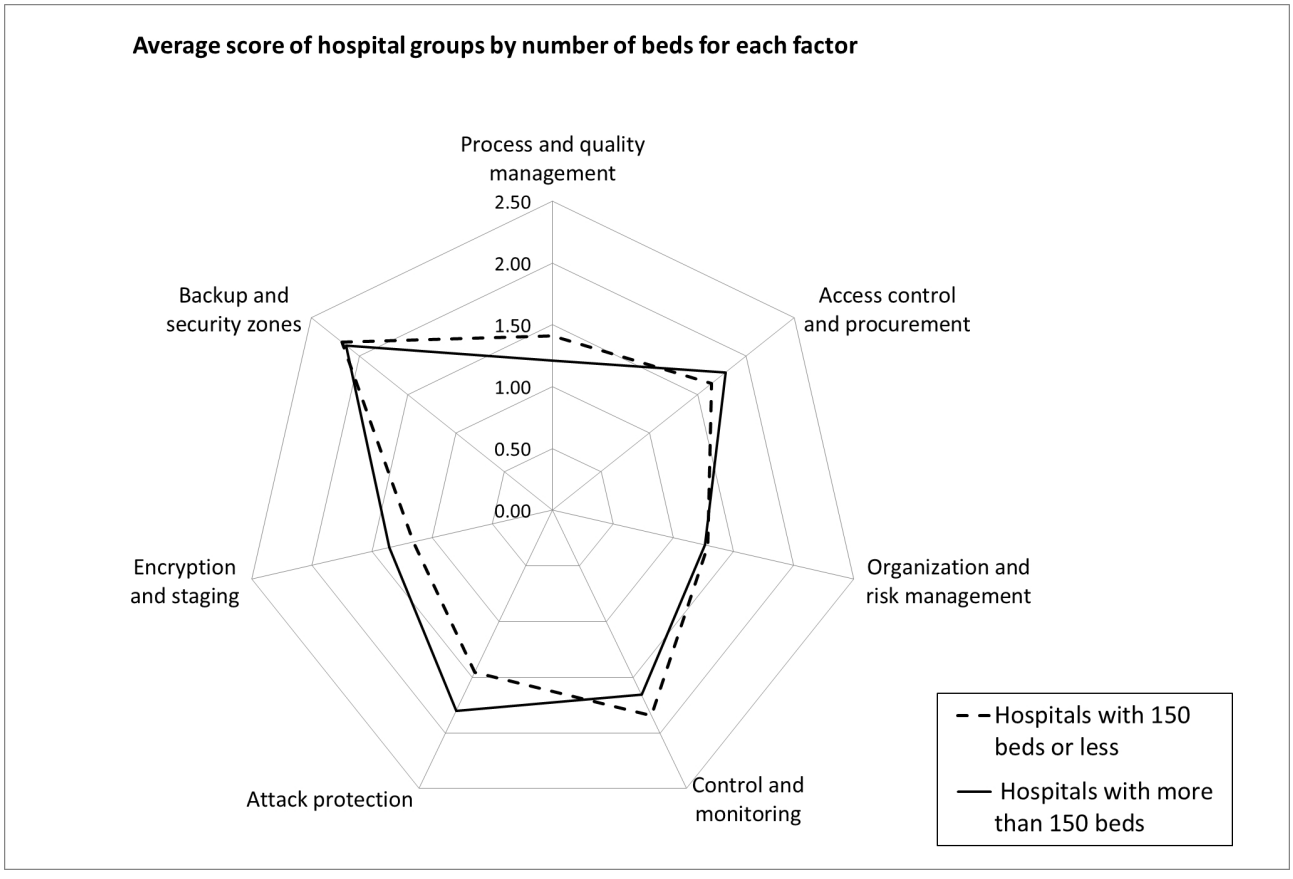


Figure 3. Scores by hospital size.



Discussion

Main Results

In this investigation, a comprehensive, but efficient and rapid, method to survey information security in institutions was introduced and successfully applied in 51 Swiss hospitals. Half (50.00%, 8588/17,177) of all hospital beds had a score of less than 49.2% of the maximum possible score (100%). In other words, a patient in one of these hospitals runs a 50% risk that he or she will lie in a hospital bed for which information security scores only reach 49%. Furthermore, university hospitals had lower scores for basic security features than private hospitals, although this difference does not reach statistical significance when conservative testing procedures are applied.

Methods

The lack of tools to quickly and inexpensively assess the information security of large numbers of hospitals led us to develop an effective and comprehensive survey tool. Only the 24 most-pertinent parameters of the total 133 parameters in the ISO/IEC 27002 standard were included to keep the resulting questionnaire manageable in a reasonable amount of time and to restrict the amount of data generated. This was possible by combining several subchapters into one parameter and selecting questions especially relevant to hospitals.

Tools, such as the one presented here, will become increasingly important as more and more countries need to address issues of information security in their health care systems.

Questionnaire

Only the data of hospitals that fully completed the questionnaire were analyzed. However, a selection bias may have influenced participation: Hospitals envisaging a potential for improvement in their security management may have been more willing to fill out the questionnaire to receive free advice through our benchmark report. On the other hand, hospitals apprehensive of a bad ranking might have refrained from answering. Unfortunately, this bias cannot be verified retrospectively.

Only 44% (20/45) of the smaller hospitals responded to the survey, compared to 73% (49/67) of the larger hospitals (Table 2). It is possible that smaller hospitals have fewer IT resources and, therefore, did not take the time to fill out our questionnaire. However, it should be noted that larger hospitals aborted the questionnaire more frequently. Fewer university hospitals completed the questionnaire than other hospitals (Table 2). The number of invited university hospitals may appear misleadingly high, as the German-speaking part of Switzerland only has three university hospitals. However, several university hospitals have subunits (eg, children's hospitals) with completely or partially independent IT structures. Fortunately, for each university hospital, exactly one eligible person representing the entire institution filled in the questionnaire. It is likely that these institutions appointed someone to respond, which also explains the frequency of aborted questionnaires.

Whether the responders filled out the questionnaire truthfully, whether they portrayed information security as more sophisticated than it actually is, or whether some respondents even understated their hospital's performance to be able to apply for more funds for their department remains unclear. These questions can only be explored with an on-site investigation.

Data Processing

Meaningful groups of security items were formed using factor analysis. The items of the first 5 factors dealt with similar topics. The final 2 factors, however, mixed different topics. This led to a decrease of eigenvalues and of the explanatory power of higher factor numbers as a consequence because of the very nature of factor analysis. Although all items in factor 6 (encryption and staging) showed an unambiguous high correlation with their factor, the mapping of question 8 about "security zones" to factor 7 (backup and security zones) was less straightforward because this question also showed high correlation with factor 5 (attack protection) to which it might also have been attributed based on its content. However, the authors decided to base factor attribution on the highest correlation and accepted the automated mapping suggested by the factor analysis.

Limitations

Switzerland is a small country with four different linguistic regions. Because only hospitals in the German-speaking part of Switzerland were included, the number of hospitals surveyed was relatively low. It would be interesting to perform the study in a larger, more uniform country to be able to work with larger numbers.

Risks for Patients

Of the 133 controls in the ISO/IEC 27002, *Information technology – Security techniques – Code of practice for information-security management*, the 24 that referred to issues of basic security were selected for our questionnaire.

Secure information processing in hospitals, such as preventing the loss or the (conscious or unintentional) manipulation of data, is crucial for patients' health. Moreover, patient health data is protected by law in Switzerland: All patient data must be stored, transmitted, and processed in a secure way that ensures confidentiality and integrity. Our results showed that only 50% of the hospital beds reach 50% of the maximum security score, implying a substantial need for improvement in many of the controls surveyed.

Recommendations for Hospitals

To address the most evident risks found in Swiss hospitals, we recommend considering the following points. We limit ourselves to the 8 questions with the lowest scores (see Appendix 1 for questions and for the reference to the ISO/IEC 27002 section [9]):

1. Risk assessments should be conducted regularly to identify, quantify, and prioritize risks (Question 1). The results should guide and determine the appropriate management action and help to prioritize controls to protect against these risks.
2. Information-security policies, standards, and guidelines should be created, approved by management, published, and communicated to all employees and relevant external parties (Question 2). These documents define how security

is managed within the hospital. They should be regularly reviewed to ensure their suitability, adequacy, and effectiveness.

3. Management should actively support security within the hospital (Question 3). Administrative management and medical management should support security through clear direction, demonstrated commitment, explicit assignment, and acknowledgment of information-security responsibilities. The support of the management is one of the most important pillars of a strong information-security culture.
4. Information classification should be implemented in the hospitals (Question 5). Information has different values and may be subject to different regulations. Knowing the value, sensitivity, and importance of hospital data allows for prioritizing the protection measures.
5. A policy on the use of cryptographic controls should be developed (Question 19). Today an increasing number of internal and external systems exchange data with each other over the Internet. These data are subject to data protection law and have to be protected. Cryptographic controls allow a secure data exchange and guarantees the integrity and authenticity of the hospital data.
6. Responsibilities and procedures should be established to handle information-security incidents effectively once they have been reported (Question 22). A process of continual improvement should be applied to learn from such events.
7. A business continuity plan should be developed and maintained for business continuity throughout the hospital (Question 23). Business continuity protects critical business processes from the effects of major failures of information systems or disasters. It is especially important for the critical infrastructure of hospitals.
8. The security of the hospital information systems should be reviewed regularly (Question 24). Different approaches exist to review the compliance of information processing with policies and standards, such as baseline audits, penetration tests, or vulnerability scans. Such reviews will reveal weaknesses and allow for prioritizing the protection measures.

Implementing these measures will close the most important information security gaps in Swiss hospitals. They also lay the foundation for further security optimizations.

Conclusions

In this paper, a comprehensive and efficient survey tool to obtain meaningful data concerning information security was introduced. Applied to assess information security in hospitals within the German-speaking section of Switzerland, it revealed surprisingly low security scores, especially for basic security issues. These results raise serious questions as to whether Swiss hospitals meet their patients' expectations and the country's legal requirements with regard to the level of information security they can guarantee. Our survey identified an urgent need for action to improve information security in hospitals, independent of their size and type.

In the future, the need for secure information handling in hospitals will increase greatly because of increased IT usage and digitalization in the health care sector. Information must also be protected from cyber threats that are increasing in number and sophistication. In the future, we will see more cyber threats that will directly attack industrial plants or a country's or region's critical infrastructures [17,18]. Hospitals are part of this critical infrastructure of a country; therefore, they must be protected from such information security breaches.

Acknowledgments

The authors would like to thank the students of the 2010/2011 "Research and Project Management" class at the Berne University of Applied Science who conducted this study as part of a class project: Amonn Christian, Bernegau Hans-Georg, Blüchel Tobias, Bolzern Silvio, Bozkurt Deniz, Brunner Heinz-René, Brunner Manuel, Fehlmann Daniel, Graf Schläppi Franziska, Herren Lukas, Hjortsberg Linn, Hügle Markus, Iseli Martin, Jeitziner Marie-Madlen, Kaltenrieder Jürg, Kaufmann Reto, Maurer Corinne, Mühlenbein Markus, Napoletano Daniel, Nilsen-Luginbühl Marianne, Nyfeler Franziska, Pulver Lilly, Schärer Benjamin, Schweizer Stefan, Siegenthaler Christoph, Sigg Stephan, Weber Jürg, and Zettler Sybille.

We would also like to thank Marc Matter who helped with the scientific writing.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questions assessed in the online survey.

[PDF File (Adobe PDF File), 34KB - ijmr_v1i2e11_app1.pdf]

References

1. Dinh M, Chu M. Evolution of health information management and information technology in emergency medicine. Emerg Med Australas 2006 Jun;18(3):289-294. [doi: 10.1111/j.1742-6723.2006.00855.x] [Medline: 16712540]
2. Einbinder JS, Bates DW. Leveraging information technology to improve quality and safety. Yearb Med Inform 2007:22-29. [Medline: 17700900]

3. Katsikas S, Lopez J, Pernul G. The challenge for security and privacy services in distributed health settings. Stud Health Technol Inform 2008;134:113-125. [Medline: 18376039]
4. Calberson FL, Hommez GM, De Moor RJ. Fraudulent use of digital radiography: methods to detect and protect digital radiographs. J Endod 2008 May;34(5):530-536. [doi: 10.1016/j.joen.2008.01.019] [Medline: 18436029]
5. Lovis C, Spahni S, Cassoni N, Geissbuhler A. Comprehensive management of the access to the electronic patient record: towards trans-institutional networks. Int J Med Inform 2007 Jun;76(5-6):466-470. [doi: 10.1016/j.ijmedinf.2006.09.014] [Medline: 17084663]
6. Joshi A, Finin T, Kagal L, Parker J, Patwardhan A. Security policies and trust in ubiquitous computing. Philos Transact A Math Phys Eng Sci 2008 Oct 28;366(1881):3769-3780 [FREE Full text] [doi: 10.1098/rsta.2008.0142] [Medline: 18672450]
7. International Organization for Standardization/International Electrotechnical Commission. ISO/IEC 27000:2009 Information technology -- Security techniques -- Code of practice for information security management. Geneva: International Organization for Standardization; 2009.
8. Gomes R, Lapão LV. The adoption of IT security standards in a healthcare environment. Stud Health Technol Inform 2008;136:765-770. [Medline: 18487824]
9. International Organization for Standardization/International Electrotechnical Commission. ISO/IEC 27002:2008 Information technology - Security techniques - Code of practice for information security management. Geneva: International Organization of Standardization; 2008.
10. Glaser J, Aske J. Healthcare IT trends raise bar for information security. Healthc Financ Manage 2010 Jul;64(7):40-44. [Medline: 20608415]
11. SerNet GmbH. verinice. 2011 URL: http://www.verinice.org/en/ [accessed 2011-10-17] [WebCite Cache ID 62VaOxG6a]
12. Bundesamt für Gesundheit (BAG). Strategie „eHealth" Schweiz. 2007. URL: http://www.e-health-suisse.ch/index.html?lang=de&download=NHzLpZeg7t,lnp6I0NTU042l2Z6ln1acy4Zn4Z2qZpnO2Yuq2Z6gpJCDdHt5f2ym162epYbg2c_JjKbNoKSn6A--- [WebCite Cache ID 62Vayvxib]
13. Globalpark AG. Online surveys software for Universities. 2011. Online surveys software for Universities URL: http://www.unipark.com/1-1-unipark-homepage.htm [accessed 2011-10-17] [WebCite Cache ID 62VbO7QlH]
14. Cook C, Heath F, Thompson RL. A meta-analysis of response rates in Web- or Internet-based surveys. Educational and Psychological Measurement 2000 Dec;60:821-836.
15. National Geographic Society. The Genographic Project. 2011 URL: https://genographic.nationalgeographic.com/genographic/index.html [accessed 2011-10-17] [WebCite Cache ID 62VbkBIrp]
16. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
17. Chen T, Abu-Nimeh S. Lessons from Stuxnet. Computer 2011;44(no. 4):91-93. [doi: 10.1109/MC.2011.115]
18. Michael B, Voas J, Laplante P. Cyberpandemics: History, Inevitability, Response. IEEE Security and Privacy, 2009;7(1):63-67. [doi: 10.1109/MSP.2009.4]

Abbreviations

CIO: chief information officer
FTE: full-time equivalents
IEC: International Electrotechnical Commission
ISO: International Organization for Standardization
ISO/IEC 27002 standard: Information technology – Security techniques – Code of practice for information security management

Edited by G Eysenbach; submitted 16.04.12; peer-reviewed by B von Solms; comments to author 14.06.12; revised version received 28.07.12; accepted 25.09.12; published 07.11.12.

Please cite as:
Landolt S, Hirschel J, Schlienger T, Businger W, Zbinden AM
Assessing and Comparing Information Security in Swiss Hospitals
Interact J Med Res 2012;1(2):e11
URL: http://www.i-jmr.org/2012/2/e11/
doi:10.2196/ijmr.2137
PMID:23611956

Original Paper

Diabetes Management Using Modern Information and Communication Technologies and New Care Models

Emmanouil G Spanakis[1*], Ph.D, MSc, BCompSci; Franco Chiarugi[1*], MSc; Angelina Kouroubali[1*], Ph.D; Stephan Spat[2*], MSc; Peter Beck[2*], Ph.D; Stefan Asanin[3*], BCompSci, Msc Candidate; Peter Rosengren[3*], Tech. Lic; Tamas Gergely[4*], Ph.D; Jesper Thestrup[5*], MSc

[1]Computational Medicine Laboratory (CML), Institute of Computer Science (ICS), Foundation for Research and Technology - Hellas (FORTH), Heraklion, Crete, Greece

[2]JOANNEUM RESEARCH, Forschungsgesellschaft mbH, Graz, Austria

[3]CNet Svenska AB, Danderyd, Sweden

[4]Applied Logic Laboratory, Budapest, Hungary

[5]In-JeT ApS, Birkeroed, Denmark

[*]all authors contributed equally

Corresponding Author:
Emmanouil G Spanakis, Ph.D, MSc, BCompSci
Computational Medicine Laboratory (CML)
Institute of Computer Science (ICS)
Foundation for Research and Technology - Hellas (FORTH)
Nikolaou Plastira 100, Vassilika Vouton
Heraklion, Crete, Greece, P.O Box 1385
Heraklion, Crete, GR-70013
Greece
Phone: 30 00302810391446
Fax: 30 00302810391428
Email: spanakis@ics.forth.gr

Abstract

Background: Diabetes, a metabolic disorder, has reached epidemic proportions in developed countries. The disease has two main forms: type 1 and type 2. Disease management entails administration of insulin in combination with careful blood glucose monitoring (type 1) or involves the adjustment of diet and exercise level, the use of oral anti-diabetic drugs, and insulin administration to control blood sugar (type 2).

Objective: State-of-the-art technologies have the potential to assist healthcare professionals, patients, and informal carers to better manage diabetes insulin therapy, help patients understand their disease, support self-management, and provide a safe environment by monitoring adverse and potentially life-threatening situations with appropriate crisis management.

Methods: New care models incorporating advanced information and communication technologies have the potential to provide service platforms able to improve health care, personalization, inclusion, and empowerment of the patient, and to support diverse user preferences and needs in different countries. The REACTION project proposes to create a service-oriented architectural platform based on numerous individual services and implementing novel care models that can be deployed in different settings to perform patient monitoring, distributed decision support, health care workflow management, and clinical feedback provision.

Results: This paper presents the work performed in the context of the REACTION project focusing on the development of a health care service platform able to support diabetes management in different healthcare regimes, through clinical applications, such as monitoring of vital signs, feedback provision to the point of care, integrative risk assessment, and event and alarm handling. While moving towards the full implementation of the platform, three major areas of research and development have been identified and consequently approached: the first one is related to the glucose sensor technology and wearability, the second is related to the platform architecture, and the third to the implementation of the end-user services. The Glucose Management System, already developed within the REACTION project, is able to monitor a range of parameters from various sources including glucose levels, nutritional intakes, administered drugs, and patient's insulin sensitivity, offering decision support for insulin dosing to professional

caregivers on a mobile tablet platform that fulfills the need of the users and supports medical workflow procedures in compliance with the Medical Device Directive requirements.

Conclusions: Good control of diabetes, as well as increased emphasis on control of lifestyle factors, may reduce the risk profile of most complications and contribute to health improvement. The REACTION project aims to respond to these challenges by providing integrated, professional, management, and therapy services to diabetic patients in different health care regimes across Europe in an interoperable communication platform.

(Interact J Med Res 2012;1(2):e8) doi:10.2196/ijmr.2193

KEYWORDS

Medical Information Systems; Medical Expert Systems; Biomedical Engineering; Biomedical Informatics; Biomedical Computing; Telemedicine; Diabetes

Introduction

Diabetes mellitus is a metabolic disorder characterized by hyperglycemia (high blood sugar) resulting from defects in the production or response to insulin [1]. The disease has two main forms: type 1 and type 2. Type 1 disease is characterized by diminished insulin production resulting from the loss of beta cells in the pancreatic islets of Langerhans, in most cases caused by immune-mediated cell destruction. Disease management entails administration of insulin in combination with careful blood glucose monitoring. Type 2 diabetes patients, mostly over 50 years old (although more and more young people develop type 2 diabetes) with additional health problems (eg, cardiovascular disease), in the early stages are often characterized by high plasmatic insulin concentration. In the fasting state, the basal insulin secretion rate increases as a function of the progressive insulin resistance [2]. In later stages of type 2 diabetes, beta cells are unable to produce enough insulin, and then type 2 becomes more similar to type 1 [3]. Management principally involves the adjustment of diet and exercise level and the use of oral anti-diabetic drugs and, eventually, insulin to control blood sugar. Type 2 diabetes is one of the faster growing chronic conditions in the developed world and is closely linked to the emerging epidemic of obesity and unhealthy lifestyles, which are among the main causes of preventable health problems [4].

Glucose Control in Diabetes Therapy for Insulin-Dependent Patients

Blood glucose is typically measured in a drop of capillary blood using a disposable dry chemical strip and reader device, an uncomfortable and slow process. Tight glucose control requires almost continuous measurements, and different sensors for continuous blood glucose measurement have been under development in the last two decades. Minimally invasive sensors able to measure glucose in interstitial fluid, more suitable for self-monitoring, have also been developed. To date, however, none of these has delivered a level of performance sufficient for use in routine glucose monitoring. Robust, clinically acceptable devices are expected to become widely available in the near future [5].

Although several guidelines for treatment regimen for management of diabetes have been defined [6-8], no adequate implementations of treatment regimen have been found for the establishment of glycemic control of hospitalized patients [9,10]. Studies have shown that frequency of hyperglycemia in surgical intensive care units can amount to 50-70% of all admitted patients [11].

Comprehensive management of diabetes has to be performed both in hospitals and outside health care premises. Based on the emerging clinical evidence from several clinical studies, there are worldwide increasing efforts to establish tight glycemic control in critically ill and hospitalized patients [12-14]. Achieving the goal of tight glycemic targets requires extensive nursing efforts, including frequent glucose monitoring, training of patients and carers to handle control algorithms, guidelines with intuitive decision making, and most importantly, additional responsibility to prevent hypoglycemic episodes. Traditional diabetes therapies for insulin-dependent patients try to achieve normal glycemia by administrating synthetic insulin to control patient's blood sugar level. Given that insulin cannot be taken orally, patients must turn to special type of devices to administer insulin.

Insulin Delivery Devices for Diabetes Therapy

The insulin needs of the body are covered with basal insulin, representing the background insulin (taking into account the daily activity level), which is normally secreted by the pancreas irrespectively of meals, and bolus insulin, representing the extra insulin necessary for balancing the glucose taken with food, which depends on the size of the meal.

During recent years, a number of insulin delivery systems have become available making insulin administration much easier. Many factors influence the choice of the appropriate device for each patient including patient conformance and self-care capacity. Most commonly, insulin is delivered using a needle injection or a syringe, an insulin pen, or an insulin pump.

A syringe is the simplest device used for the injection of insulin, where the patients typically use disposable units to prevent contamination and infection. An insulin pen is an injection device the size of a pen that includes a needle and holds a vial of insulin. There are two different types of insulin pens: a) durable with replaceable insulin cartridges; and b) disposable prefilled with insulin. The advantage of insulin pens over insulin syringes is that they are much more convenient and easier to transport and use than traditional vial and syringe. They can repeatedly administer more accurate dosages (especially for patients with visual or motor skills impairments). Insulin pens usually cause less pain to the patient. One of the disadvantages of insulin pens over insulin syringes is that, unlike traditional

syringes, two different insulin solutions cannot be mixed in an insulin pen. Also, using pens needles is usually more expensive than using the traditional vial and syringe method. A major disadvantage of insulin injection devices is that they are designed to administer insulin only in large boluses, which can cause peaks and valleys in the blood sugar levels in patients.

A solution to this problem of insulin shots is provided when using insulin pumps. These devices, which are worn outside of the patient's body, can be programmed to deliver a steady supply of insulin throughout the day at a basal rate and/or programmed to deliver larger boluses of insulin before or after meals. With a pump, continuous doses of background insulin are delivered to support the body's needs between meals, and with a button press it is possible to obtain an "on demand" dose of insulin (a bolus) to cover instant needs. Each basal rate can be variable since the pump can be programmed to deliver different basal rates throughout the day. The main advantage of insulin pumps is that the individuals do not need to take multiple injections every day, allowing them to continue with their daily actions without any problem. The main disadvantage of insulin pumps, apart from the obvious discomfort that a person might feel wearing it, is the cost of the pump and its maintenance, which can get very high. Furthermore, patient's activity could force the infusion set of the pump to slide off and stop the necessary delivery of insulin to the patient's body. Therefore, it is very important for patients to monitor their blood glucose levels much more frequently making sure the pump is working correctly and avoiding risks of ketoacidosis.

Methods

The Need for New Care Models

Health care practice supported by electronic processes and communication (eHealth) provides new possibilities for revising central parts of the established care models for chronic diseases. Planning new care models for the future involving eHealth is highly complex and involves a number of different factors that influence the opinions and attitudes of the participants in health care systems and their ability to carry out changes that are needed to implement new care models [15].

Several prediction studies have been published by health institutions as well as private observers. These studies differ in their approach and in the aspects they consider. Some studies emphasize technological aspect [16,17]; other focus on health care policies [18,19]. A study from the Australian Centre for Health Research focuses on the aspects of health services [20]. Some common elements can be extracted from these approaches, and the most significant factors for the formation of new care models are presented in Figure 1 [21,22].

Chronic diseases share three important features: i) acute and chronic phases alternate through time; ii) adequate diagnosis and state monitoring require multilevel system biomedical characterization; and iii) disease progress is significantly influenced by patient's behavior. Care for chronic diseases therefore must be: i) continuous between visits and hospitalization periods; ii) proactive and predictive; iii) provided by medical personnel as professional care and by patients as self-management; iv) influencing the patient's lifestyle; and v) dynamic, since all participants should learn and adapt during the care process.

Important aspects of chronic disease management are personalization, inclusion, and patient empowerment. These aspects are closely connected to care space evolution, ie, how the physical or virtual spaces of care change. In this regard, Information and Communication Technology (ICT) tools and biomedical technology act as enablers of mobile and remote health solutions and thus have an important role in care space evolution models.

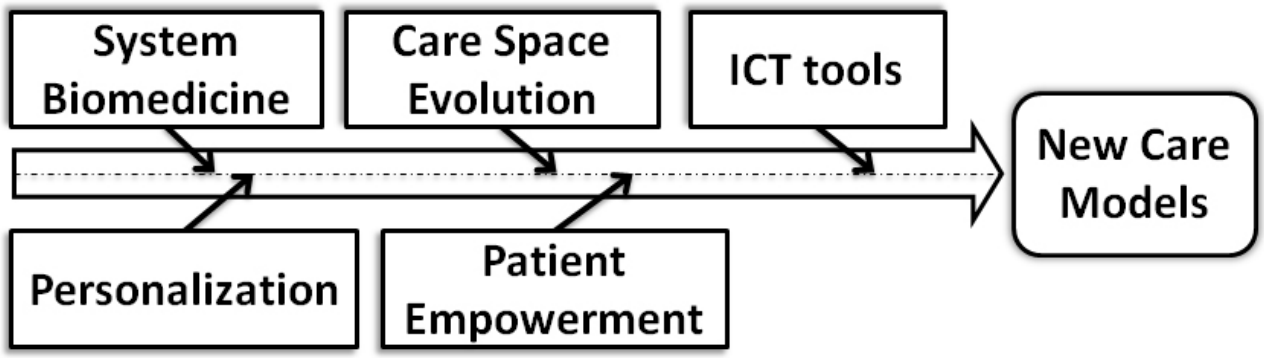
Recent results in eHealth explicitly split the care space into two interrelated spaces, the activity and the information space. Activities are implemented in a distributive way while the necessary data and information come from an integrated and unified information space. A shared access is available from different activity spaces into the information space. There are activity spaces of the patients, the care teams, and the care provider, as well as a virtual one of the intelligent agents. The information space and the activity spaces together generate the integrated care space, which promotes the effective application of info-communication technologies and unites the participants into a virtual organization [21,22].

This model of care allows patient mobility to decouple the actual activity space from the traditional physical care space of a professional health care system, placing emphasis on the sound development of the care space. New health care methods can create a growing volume of data on the one hand, and an ever growing demand for information, on the other hand. The data growth is connected with the development of diagnostic and monitoring methods and tools. The information demand is connected with new decision making methods and with modeling methods required for a better implementation of the necessary actions.

The separation of care spaces in chronic disease management allows streamlining of the roles of each stakeholder while at the same time providing them with tailored information for each role. Also, if managed correctly, it provides great opportunities for improved care at the point of need as well as organizational streamlining and potentials for cost savings.

In this case, in order to define new care models in the context presented, a challenge to overcome is the risk of the patient being isolated or alienated. If visits by care givers or family members are completely substituted with monitoring tools, patients can become isolated even in the most populated areas. It is imperative that a delicate balance is struck between closing the digital divide and closing patients in a virtual prison. Thus, inclusion enhancing methods must be incorporated in the new care models at all levels [23].

Figure 1. Significant factors influencing the new care models.



Advanced ICT Solutions in Diabetes Management: The REACTION Project

Due to demographic changes, European health care systems face two serious challenges: health care delivery may become inadequate for the perceived needs of the citizens or the cost may spiral out of control. Advanced ICT has the potential to provide service platforms able to improve health care. However, dramatic changes to health care provisioning and care models are needed.

The REACTION project [24] aims to address new care models for diabetes management through various clinical applications, such as monitoring of vital signs, feedback provision to the point of care, integrative risk assessment, event, and alarm handling. In addition, the project plans to integrate clinical and organizational workflows with external health information systems for attaining improvements in continuous blood glucose monitoring and tight/safe glycemic control [25]. The envisaged intelligent service platform for management and therapy of diabetes intends to reach different health care regimes across Europe.

Identifying Advanced Technology Services for Diabetes Management: User Preferences in Europe

User preferences and needs were explored through a series of focus groups conducted in four different European countries [26]. Rather than testing an already developed service within the project, the study explored preferences of end users, who were not involved in the project, about services that are in development phase. Also, the study offered an expanded view of technology management of diabetes that involves individualized care and cultural differences.

The focus groups examined what diabetic patients, nurses, doctors as well as health care professionals, and informal carers would expect from technology, in addition to identifying values, beliefs, hopes, concerns, and needs related to the use of telemonitoring services. Focus groups also highlighted how the use of information technology could potentially change the experience of living with diabetes. Understanding societal factors is a core prerequisite for addressing ethical and social issues at the design stage of technology development.

The focus groups were held in Greece, Italy, Cyprus, and France, with a range of participants including doctors, nurses, social scientists, technical personnel, patients, carers, nutritionists, and lawyers. The questions that guided the discussions focused on several topics including information and risk management of diabetes, security privacy and confidentiality issues, quality of living, monitoring and alert systems, device and sensors design, technical skills, daily activities, and end-users' concerns. Discussions were informal, encouraging all participants to express their opinion and relate their experiences with diabetes care. Recording devices were not used, and answers have not been attributed to specific participants guaranteeing confidentiality and privacy [27].

Analysis of the focus group discussions identified several topics involved in diabetes and new technologies. The following paragraphs present some of the views and opinions of the participants.

Autonomy and self-management: Technology can assist users to regain a measure of autonomy in managing their condition and prevent long-term risks. Focus group participants viewed technology as a potential personal assistant. An application on a mobile phone could be designed to provide personalized estimation of insulin needs, messages for motivation, and support as well as alerts. However, continuous monitoring may not be essential for improved management of diabetes and could even be seen as information overload. Participants expressed their opinion that technology can improve not only the physical management of the disease but also the social and psychological conditions. However, they also believe that widespread availability of technology at an affordable cost and effective integration into the national health systems remain open issues.

Privacy and confidentiality: Focus group participants expressed their willingness to disclose all information relevant to their condition to their physician, without particular attention to how this information is transmitted. However, if they had the option, they would like all transactions to include strong security mechanisms for any transmitted information. Participants would like to be informed about who is accessing their data, to discern whether their information is shared with trustworthy persons or with someone who may abuse their data. Participants with diabetes appeared less concerned about trust in the use of an Internet platform in conveying personal medical information. However, they were concerned about downtime of devices, inaccuracy of advice, and loss of personal information. Finally,

participants were willing to allow their data to be used anonymously for research purposes.

Diabetic person–care provider relationship: All participants expressed the conviction that diabetes management benefits from a close relationship between the person with diabetes and the health care provider team including the physician, nurses, nutritionist, psychology and others. Participants felt that technology may allow for a more accurate, faster response to crisis, as well as better overall management and prevention of complications. However, some participants felt that a careful balance between information and communication is important to avoid information overload and excess. Technology for diabetes management may bring profound changes in self-care and empowerment, as well as in-depth communication with the health providers.

Health management: Management of diabetes requires specific information in predefined time intervals that coupled with personalized algorithms can assist in insulin dosage decisions. A technology platform for diabetes management would require frequent measurements and regular entries of data. Participant willingness for data entry depended on the perceived benefit of the technology services and the time that data entry would require. However, people with diabetes felt that parameters that affect management of their condition, such as emotional and psychological stress, physical exercise, and variability of daily life, are more difficult to define, monitor, predict, and account for. A technology service that could take into consideration some of these parameters would be at a relative advantage compared to a service that uses only glucose measurements and food intake. In addition, participants felt that the amount of information required to be recorded and the ease of use of the technology services will play an important role in the acceptance and use of the services.

Results

Closed-Loop and Self-Management of Diabetes Using ICT

State-of-the-art technological solutions have the potential to assist health care professionals, patients, and informal carers to better manage diabetes insulin therapy in a variety of settings, help patients understand their disease, support self-management, and provide a safe environment by monitoring adverse and potentially life-threatening situations with appropriate crisis management.

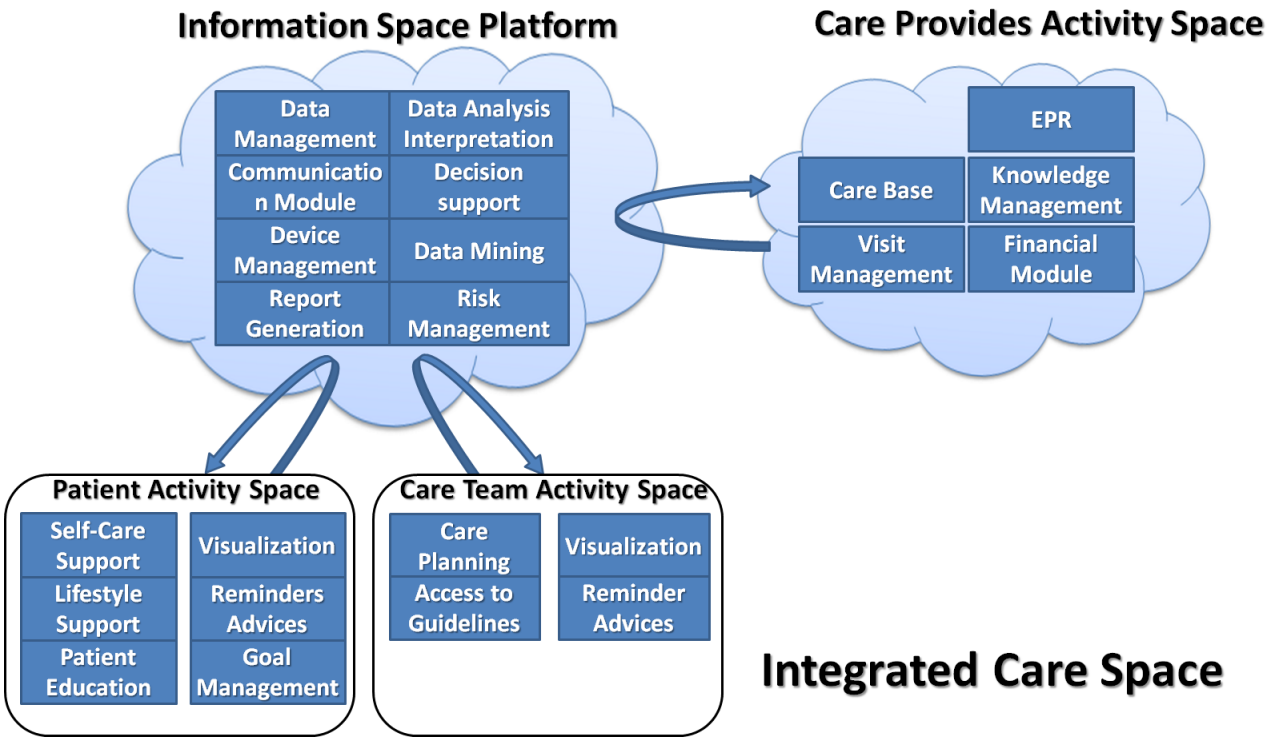
The REACTION project aims to support long-term management of diabetes based on the identified user requirements through modern advanced ICT solutions enabling wearable, continuous blood glucose monitoring, and automated closed-loop delivery of insulin. REACTION is developing a platform that will provide integrated, professional, management and therapy services to diabetes patients in different health care regimes, including professional decision support for in-hospital environments, and safety monitoring for dosage and compliance. This platform provides health care services to diabetes patients and caregivers using new chronic care models that support separation of activity spaces of the information care space (Figure 2) where a closed-loop system is implemented for managing and treating diabetes and for delivering insulin. These include wireless technologies for continuous blood glucose monitoring, clinical monitoring, and intervention strategies, monitoring and predictive risk analysis disease indicators.

The loop has to be closed to health professionals inside hospitals or to the patients (with the main focus on elderly patients in the case of type 2 diabetes) when outside the health care premises. Insulin will be delivered using the appropriate device. In case of automatic glucose management, the loop will be automatically closed on an insulin pump. At such purpose, the availability of glucose sensors with proper accuracy is requested for an effective and efficient functioning of the overall system.

Three clinical field trials are foreseen in the context of the REACTION project: Safe Glycemic Control (SGC) [25] in the hospital ward; chronic care and lifestyle management in the primary care sector; and Automatic Glycemic Control with closed-loop feedback.

While moving towards the full implementation of the platform, three major areas of research and development have been identified and consequently approached: the first one is related to the glucose sensor technology and wearability, the second is related to the platform architecture, and the third to the implementation of the end-user services.

Figure 2. The REACTION platform information and activity spaces and services.



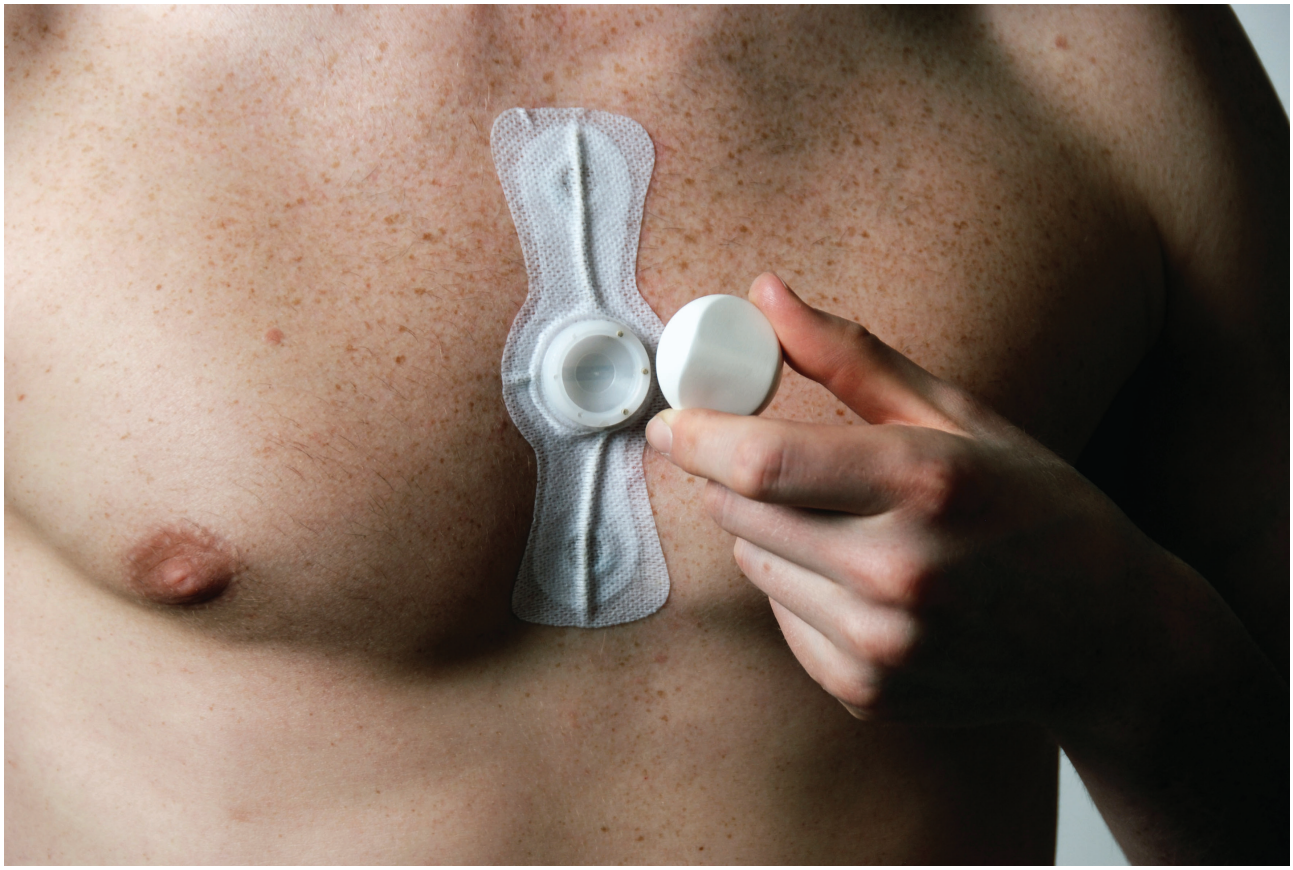
Evolution in Glucose Sensor Technology and Wearability

The standard of care for measuring glucose levels is by "finger-stick" blood glucose meters. For these, a drop of blood, usually drawn by piercing the skin of a finger, is brought in contact with a test strip. A chemical reaction, commonly mediated by glucose oxidase, glucose dehydrogenase or hexokinase enzymes, triggers an electrochemical sensor or a color reaction that is detected in a reader. The drawback of this method is that only a few measurements can be performed in the course of a day.

At this moment, there are no sensors commercially available that continuously measure blood glucose. Only a few commercially available sensors that allow continuous glucose monitoring in the interstitial fluid, which is a surrogate for blood glucose, have been approved for use. These sensors rely on electrochemical detection of an enzymatic reaction and are minimally invasive. A range of other sensor technologies are currently being tested for their suitability for glucose monitoring. The most promising technologies for continuous glucose monitoring can broadly be classified as follows: enzymatic (electrochemical), impedance spectroscopy / dielectric spectroscopy, optical in the infrared (IR) or near-infrared (NIR) range (IR/NIR absorption, mid-IR emission), polarimetric (eg, anterior chamber of the eye), refractive index, Raman spectroscopy (inelastic photon scattering), photoacoustic (pulsed light absorption dependent on glucose concentration), and others. Moreover, alternatives for invasive sampling are being investigated, but despite significant efforts these technologies are still in a development or evaluation phase and have yet to prove their reliability and accuracy.

Wearability for sensors for some time can help in the collection of the required measurements. For this purpose, the use of the ePatch technology has been considered (Figure 3) within the REACTION project. The ePatch sensor is a small body sensor, optimized for wearability and bio-compatibility, which senses physiological signals and is embedded in a skin-friendly adhesive. It can contain various types of miniaturized body sensors to measure physiological parameters (eg, ECG, while other vital signs are under development within the context of the REACTION project), microelectronics for data analysis, a wireless radio module for communication, and a battery power source. The basis for the adhesive will be hydrocolloid pressure-sensitive adhesives.

Figure 3. Wearable health monitoring ePatch system from DELTA.



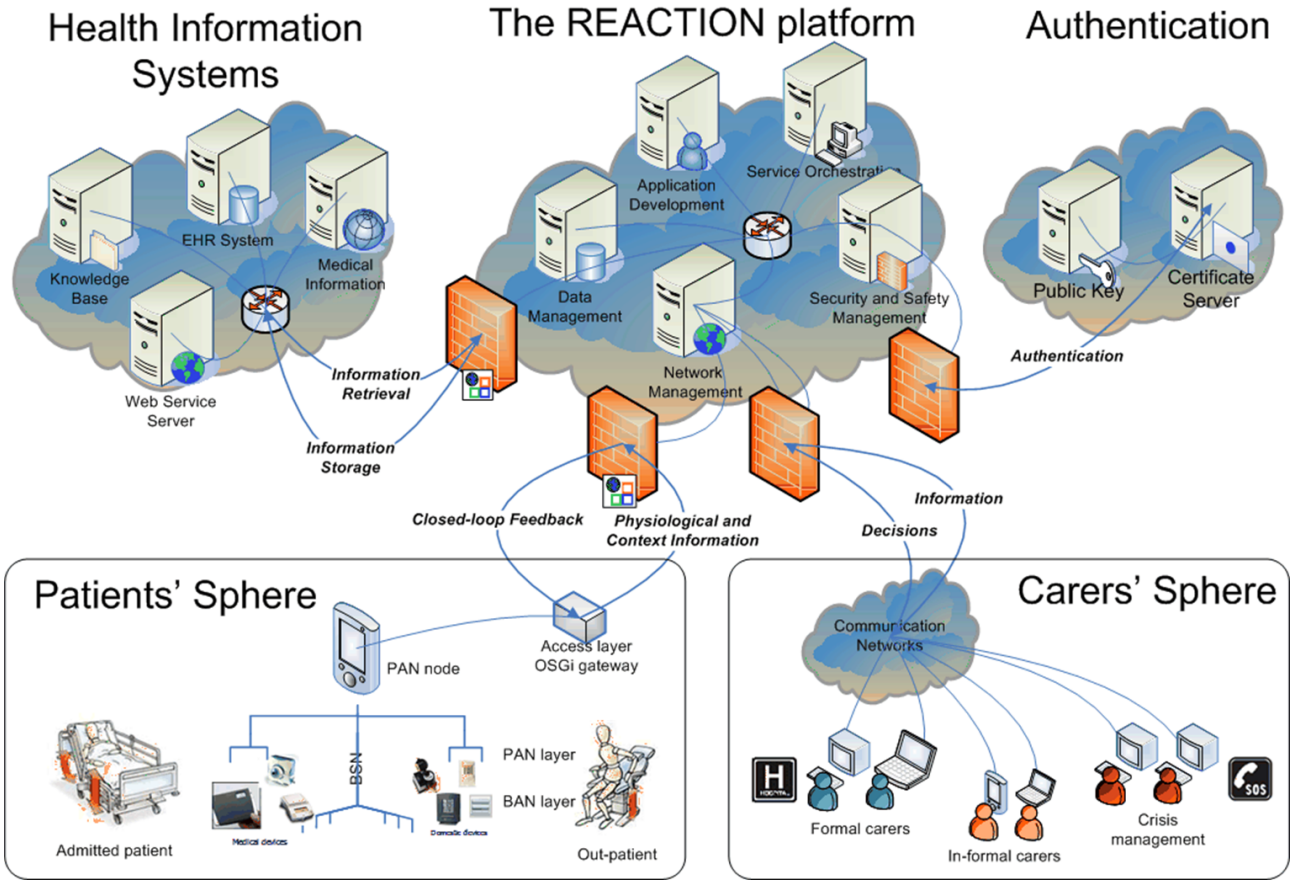
Platform Architecture

The REACTION platform (Figure 4) is a Service-Oriented Architecture (SOA) platform based on numerous individual services that can be developed and deployed to perform clinical monitoring, execute distributed decision support and security tasks, support workflow management, provide clinical feedback, and perform event handling and crisis management [28]. This approach aims to offer interoperable and loosely coupled services distributed in the network. In this context, a service is a function that is well defined, self-contained, and does not necessarily depend on the context or state of other services.

The REACTION platform, which builds on the LinkSmart [29] middleware for Internet of Things applications, represents each device as web services that can be invoked and consumed by other devices, services, or applications through an overlying peer-to-peer network.

The REACTION platform features a Device Connectivity Kit that enables integration and communication with IEEE 11073 device specializations and other medical devices based on proprietary protocols. It exposes the observations as HL7 v2.6 message format, which are transmitted from the patient sphere to a server where a Rule Engine executes a set of monitoring rules. The rules are expressed by clinicians using a graphical user interface. The platform allows monitoring of vital signs for thresholds as well as long-term trends. An Orchestration Manager allows combining and executing sequences of services and specifying actions such as sending alerts or rising alarms.

Figure 4. REACTION Platform architecture.



In-hospital Glucose Management System in Compliance With the Medical Device Directive (MDD)

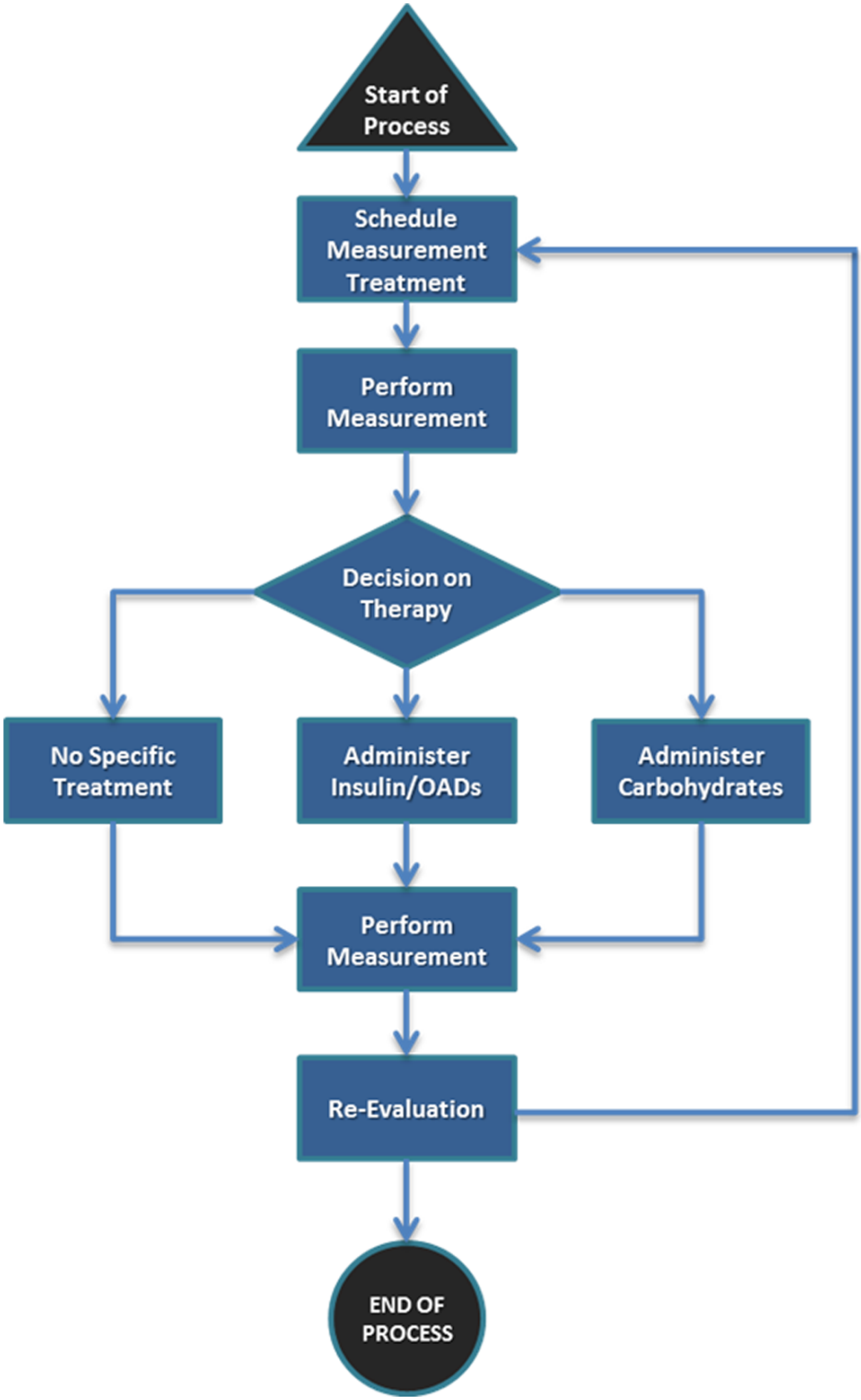
In the first phase of the REACTION project, the implementation of the end-user services has been considered for the hospital ward. A user-centered design approach has been chosen for the development of the first clinical application to be deployed in the endocrinology ward at the Medical University of Graz [30]. The following standards have been considered as relevant for the development of the appropriate software components for REACTION that fall within the scope of the MDD: ISO 13485, IEC 62304, ISO 14971, IEC 62366, and IEC 60601-1-6 [31].

Paper and software mock-ups have been used as a trigger in order to have a basis for discussion and testing. The design process has been supported by regular risk analysis sessions with all involved stakeholders. Derived risks have been collected and incorporated as change requests into the elicited requirements summarized in an extensive specification document. A detailed discussion of the design phase can be found in [32].

Due to maintainability and expandability requirements, it was decided to distinguish between an Android-based user interface and a platform-independent backend, which contains business logic for the electronic decision support system, as well as the data storage and interfaces to the hospital information system. The exchange of data between the backend and frontend components requires mutual authentication and is completely done via encrypted web services to provide data security. The behavior of the frontend application relates to the clinical workflow, which was identified together with end users in the design phase. A requirement issue tracker and task management tool for the documentation of each implementation step was adopted, for supporting the software life cycle processes (IEC 62304 standard), and ensuring an overview of open and already completed requirements, development tasks, and identified bugs.

In addition to the need for detailed documentation, IEC 62304 and IEC 62366 also demand verification and validation during the software life cycle. To verify the functionality after finishing implementation of each system unit, the correct behavior of the application has been tested with simulated user interactions. Already in the early phase of the requirement elicitation process, clinicians stated they would prefer a software system that offers only the required basic functionalities, but with an easy-to-use interface, tailored to the current workflow. In order to avoid poor usability and consequently fulfill the requirements according to IEC 62366, the physicians and nurses have been included in the design of the user interface based on the established clinical workflow, using paper mock-ups and functional prototypes. However, each ward in each hospital usually follows its own workflow; therefore, great attention has been given to the maintainability of the user interface.

Figure 5. Workflow description of glycemic management of a general ward.



The glucose management workflow system, created as a result of this process, offers decision support for insulin dosing to professional caregivers on a mobile tablet platform that fulfills the need of the users and is compliant with the MDD requirements. This workflow is embedded in the daily routine of the patients, the nursing, and the medical staff. The routine of the latter two groups is very structured and standardized, whereas the patients' workflows differs from day to day depending on the patients' health status as well as the planned examinations and their potential delays. These circumstances

have to be taken into account for safe glycemic control. The main in-hospital workflow related to glycemic management is shown in Figure 5.

The Glucose Management System application, developed within the REACTION project (Figure 6), is able to monitor and take into account a range of parameters from various sources including glucose level, nutritional intake, administered drugs, and the patient's insulin sensitivity. The data are contextualized and algorithms are used to calculate the required insulin doses for SGC. Results are presented to physicians and nurses at the point of care.

This system will be used, according to the REACTION work plan, for the first field trial, involving hospitalized patients with diabetes for the length of their hospital stay, in order to evaluate the feasibility and safety of the automated workflow and insulin dosing support in the ward for endocrinology of the Medical University of Graz.

Figure 6. Safe Glycemic Control decision support system.

Discussion

The trends in care model evolution are influenced by many factors such as biomedical and clinical R&D, financial incentives, technology development, and the socioeconomic environment. An important aspect of future chronic disease management is that personalization, inclusion, and empowerment of the patient have to be essential parts of the care model.

There is abundant evidence, for future diabetes management and therapy, that safe control of the blood glucose level is vital for good diabetes management and insulin therapy. Good glucose control requires frequent measurement of blood glucose levels and complicated algorithms for assessing the insulin dose needed to adjust for short-term variations in activity, diet, and stress. Good control of diabetes, as well as increased emphasis on control of lifestyle factors, may reduce the risk profile of most complications and contribute to health improvement overtime.

The REACTION project aims to respond to these challenges by providing integrated, professional, management and therapy services to diabetes patients in different health care regimes across Europe in an interoperable communication platform.

Acknowledgments

This work was supported in part by the FP7 Integrated Project REACTION (Remote Accessibility to Diabetes Management and Therapy in Operational healthcare Networks) partially funded by the European Commission under Grant Agreement 248590. The authors would like to thank all members of the REACTION consortium for their commitment and valuable contribution to the project and this paper.

Conflicts of Interest

None declared.

References

1. Zimmet P, Alberti KG, Shaw J. Global and societal implications of the diabetes epidemic. Nature 2001 Dec 13;414(6865):782-787. [doi: 10.1038/414782a] [Medline: 11742409]
2. Del Prato S, Marchetti P, Bonadonna RC. Phasic insulin release and metabolic regulation in type 2 diabetes. Diabetes 2002 Feb;51 Suppl 1:S109-S116 [FREE Full text] [Medline: 11815468]
3. Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Report of the expert committee on the diagnosis and classification of diabetes mellitus. Diabetes Care 2003 Jan;26 Suppl 1:S5-20. [Medline: 12502614]
4. Joint WHO/FAO Expert Consultation. Diet, nutrition and the prevention of chronic diseases. World Health Organ Tech Rep Ser 2003;916:i-149. [Medline: 12768890]
5. Spanakis EG, Chiarugi F. Diabetes Management: Devices, ICT Technologies and Future Perspectives. Wireless Mobile Communication and Healthcare, Mobihealth 2011;83:197-202.
6. Sakharova OV, Inzucchi SE. Treatment of diabetes in the elderly. Addressing its complexities in this high-risk group. Postgrad Med 2005 Nov;118(5):19-26, 29. [Medline: 16329527]
7. Das SK, Chakrabarti R. Non-insulin dependent diabetes mellitus: present therapies and new drug targets. Mini Rev Med Chem 2005 Nov;5(11):1019-1034. [Medline: 16307531]
8. Charles MA. Intensive insulin treatment in type 2 diabetes. Diabetes Technol Ther 2005 Oct;7(5):818-822. [doi: 10.1089/dia.2005.7.818] [Medline: 16241891]
9. Clement S, Braithwaite SS, Magee MF, Ahmann A, Smith EP, Schafer RG, American Diabetes Association Diabetes in Hospitals Writing Committee. Management of diabetes and hyperglycemia in hospitals. Diabetes Care 2004 Feb;27(2):553-591. [Medline: 14747243]
10. Gautier JF, Beressi JP, Leblanc H, Vexiau P, Passa P. Are the implications of the Diabetes Control and Complications Trial (DCCT) feasible in daily clinical practice? Diabetes Metab 1996 Dec;22(6):415-419. [Medline: 8985649]
11. Ellmerer M. Glucose Control in the Intensive Care Unit: The Rosy Future. 2008 Nov 13 Presented at: Diabetes Technology Meeting; 2008; Bethesda, Maryland, USA.
12. van den Berghe G, Wouters P, Weekers F, Verwaest C, Bruyninckx F, Schetz M, et al. Intensive insulin therapy in critically ill patients. N Engl J Med 2001 Nov 8;345(19):1359-1367 [FREE Full text] [doi: 10.1056/NEJMoa011300] [Medline: 11794168]
13. Van den Berghe G, Wilmer A, Hermans G, Meersseman W, Wouters PJ, Milants I, et al. Intensive insulin therapy in the medical ICU. N Engl J Med 2006 Feb 2;354(5):449-461 [FREE Full text] [doi: 10.1056/NEJMoa052521] [Medline: 16452557]
14. Meijering S, Corstjens AM, Tulleken JE, Meertens JH, Zijlstra JG, Ligtenberg JJ. Towards a feasible algorithm for tight glycaemic control in critically ill patients: a systematic review of the literature. Crit Care 2006 Feb;10(1):R19 [FREE Full text] [doi: 10.1186/cc3981] [Medline: 16469124]
15. Thestrup J, Gergely T, Beck P. Exploring New Care Models in Diabetes Management and Therapy with a Wireless Mobile eHealth Platform. Wireless Mobile Communication and Healthcare, Mobihealth 2011;83:203-210.
16. Intel Corporation. White Paper, Intel Corporation.: Intel Corporation; 2007. Chronic Care at the Crossroads: Exploring Solutions for Chronic Care Management URL: http://www.epractice.eu/files/media/media1900.pdf [accessed 2012-09-12] [WebCite Cache ID 6Ad0joE8C]
17. Adams J, Bakalar R, Boroch M, Knecht K, Mounib EL, Stuart N. IBM Global Business Services.: IBM Global Business Services; 2008. Healthcare 2015 and care delivery URL: http://www-03.ibm.com/industries/ca/en/healthcare/files/hc2015_full_report_ver2.pdf [accessed 2012-09-12] [WebCite Cache ID 6Ad10F9PO]
18. PricewaterhouseCoopers. PricewaterhouseCoopers. 2009. The new science of personalized medicine URL: http://www.pwc.com/us/en/healthcare/publications/personalized-medicine.jhtml [accessed 2012-09-12] [WebCite Cache ID 6Ad1Libxx]
19. PricewaterhouseCoopers. PricewaterhouseCoopers. 2005. HealthCast 2020: Creating a Sustainable Future URL: http://www.pwc.com/us/en/healthcare/publications/personalized-medicine.jhtml [accessed 2012-09-12] [WebCite Cache ID 6Ad1XsESV]
20. Georgeff M. E-Health and the Transformation of Healthcare. In: E-Health and the Transformation of Healthcare. Melbourne, Australia: Australian Centre for Health Research Limited; Apr 2007:1-18.
21. Deutsch T, Gergely T, Levay A. A New Care Model for Chronic Patient Care and its Intelligent Infocommunication System I-II. Informatics and Management in Healthcare 2009;8.
22. Gergely T. Integrated care space for chronic diseases. eHealth week 2011 Budapest 2011.
23. Mordini E, Wright D, de Hert P, Mantovani E, Wadhwa KR, Thestrup J, et al. Ethics, e-Inclusion and Ageing. Studies in Ethics, Law, and Technology 2009 Jun;3(1). [doi: 10.2202/1941-6008.1093]
24. The REACTION Project. URL: http://www.reaction-project.eu/news.php [accessed 2012-09-12] [WebCite Cache ID 6Ad2CFquY]

25. Krinsley JS, Preiser JC. Moving beyond tight glucose control to safe effective glucose control. Crit Care 2008;12(3):149 [FREE Full text] [doi: 10.1186/cc6889] [Medline: 18495050]
26. Kouroubali A, Chiarugi F. Developing Advanced Technology Services for Diabetes Management: User Preferences in Europe. In: Nikita K S, Lin J C, Fotiadis D I, Arredondo Weldmeyer M T. (Eds): Wireless Mobile Communication and Healthcare, Mobihealth 2011;83:69-74.
27. Miles MB, Huberman AM. Qualitative data analysis. Sage, London 1994.
28. Ahlsén M, Asanin S, Kool P, Rosengren P. Service-Oriented Middleware Architecture for Mobile Personal Health Monitoring. Wireless Mobile Communication and Healthcare, Mobihealth 2011;83:305-312.
29. LinkSmart. URL: http://sourceforge.net/projects/linksmart/ [accessed 2012-09-12] [WebCite Cache ID 6Ad2Y7mEy]
30. Spat S, Höll B, Beck P, Chiarugi F, Kontogiannis V, Spanakis M, et al. A Mobile Android-Based Application for In-Hospital Glucose Management in Compliance with the Medical Device Directive for Software. Wireless Mobile Communication and Healthcare, Mobihealth 2011;83:211-216.
31. Hall K. Developing Medical Device Software to IEC 62304. European Device Technology 2010;1(6).
32. Höll B, Spat S, Plank J, Schaupp L, Neubauer K, Beck P, et al. Design of a mobile, safety-critical in-patient glucose management system. Stud Health Technol Inform 2011;169:950-954. [Medline: 21893886]

Edited by G Eysenbach; submitted 31.05.12; peer-reviewed by M Tsiknakis, D Conforti; comments to author 01.08.12; revised version received 08.08.12; accepted 10.08.12; published 04.10.12.

Please cite as:
Spanakis EG, Chiarugi F, Kouroubali A, Spat S, Beck P, Asanin S, Rosengren P, Gergely T, Thestrup J
Diabetes Management Using Modern Information and Communication Technologies and New Care Models
Interact J Med Res 2012;1(2):e8
URL: http://www.jmir.org/2012/2/e8/
doi:10.2196/ijmr.2193
PMID:23612026

Review

Perceived Impact of Electronic Medical Records in Physician Office Practices: A Review of Survey-Based Research

Jesdeep Bassi[1], MSc; Francis Lau[1], PhD; Mary Lesperance[2], PhD, PStat

[1]School of Health Information Science, University of Victoria, Victoria, BC, Canada

[2]Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada

Corresponding Author:
Jesdeep Bassi, MSc
School of Health Information Science
University of Victoria
PO Box 3050 STN CSC
Victoria, BC, V8W 3P5
Canada
Phone: 1 250 472 4731
Fax: 1 250 472 4751
Email: jbassi@uvic.ca

Abstract

Background: Physician office practices are increasingly adopting electronic medical records (EMRs). Therefore, the impact of such systems needs to be evaluated to ensure they are helping practices to realize expected benefits. In addition to experimental and observational studies examining objective impacts, the user's subjective view needs to be understood, since ultimate acceptance and use of the system depends on them. Surveys are commonly used to elicit these views.

Objective: To determine which areas of EMR implementation in office practices have been addressed in survey-based research studies, to compare the perceived impacts between users and nonusers for the most-addressed areas, and to contribute to the knowledge regarding survey-based research for assessing the impact of health information systems (HIS).

Methods: We searched databases and systematic review citations for papers published between 2000 and 2012 (May) that evaluated the perceived impact of using an EMR system in an office-based practice, were based on original data, had providers as the primary end user, and reported outcome measures related to the system's positive or negative impact. We identified all the reported metrics related to EMR use and mapped them to the Clinical Adoption Framework to analyze the gap. We then subjected the impact-specific areas with the most reported results to a meta-analysis, which examined overall positive and negative perceived impacts for users and nonusers.

Results: We selected 19 papers for the review. We found that most impact-specific areas corresponded to the micro level of the framework and that appropriateness or effectiveness and efficiency were well addressed through surveys. However, other areas such as access, which includes patient and caregiver participation and their ability to access services, had very few metrics. We selected 7 impact-specific areas for meta-analysis: security and privacy; quality of patient care or clinical outcomes; patient–physician relationship and communication; communication with other providers; accessibility of records and information; business or practice efficiency; and costs or savings. All the results for accessibility of records and information and for communication with providers indicated a positive view. The area with the most mixed results was security and privacy.

Conclusions: Users sometimes were likelier than nonusers to have a positive view of the selected areas. However, when looking at the two groups separately, we often found more positive views for most of the examined areas regardless of use status. Despite limitations of a small number of papers and their heterogeneity, the results of this review are promising in terms of finding positive perceptions of EMR adoption for users and nonusers. In addition, we identified issues related to survey-based research for HIS evaluation, particularly regarding constructs for evaluation and quality of study design and reporting.

(Interact J Med Res 2012;1(2):e3) doi:10.2196/ijmr.2113

KEYWORDS

Health care surveys; evaluation studies; ambulatory care information systems

Introduction

The importance of office-based electronic medical records (EMRs) and related systems is being recognized internationally. For example, Canada Health Infoway [1] has an investment program to support the adoption and use of EMRs to help clinicians achieve increased clinical value. In its 2001 report entitled "Crossing the Quality Chasm," the Institute of Medicine discussed using information technology as one aspect of improving the health care delivery system in the United States [2]. Until now, adoption of EMRs in the ambulatory setting has been relatively slow [3,4]. According to the 2007 National Physician Survey, only 12.3% of Canadian family practitioners and general practitioners used EMRs exclusively, and 19.4% used a combination of EMRs and paper-based charts [5]. These figures rose to 21.5% and 27.5%, respectively, in the 2010 National Physician Survey [6], which indicates that adoption is on the rise. Ford et al [3] constructed a model using historical data to estimate that 86.6% of physicians in small practices will be using the systems in 2024 in the United States.

Given what appears to be a slow but increasing trend of EMR adoption, the next area that needs attention is the impact of such systems to both ensure that they are adopted and that they are helping practices to realize the expected benefits. Talmon et al [7] stated that "given the essential role of information technology (IT) systems on the delivery of modern health care, and the dependence of health professionals and organizations on them, it is imperative that they are thoroughly assessed through robust evaluations as with any other form of health process or technology" (p. 23). Evidence of positive effects of EMRs is still limited [8]. In terms of guiding impact assessment, the recently developed Clinical Adoption Framework by Lau et al [9] provides a comprehensive set of categories that address many areas for overall health information systems (HIS) adoption.

Taking a closer look at impact, we see that it can be evaluated objectively (for example, by using proxy measures such as reduction in medication errors), but there is also a subjective component for individuals involved with EMR adoption. EMRs are expected to have positive impacts in many areas, but do providers believe this? Ultimate acceptance and use of the system is up to the provider, so there is a need to understand their point of view. Based on the trends presented above, there are two general views to consider: nonuser/preimplementation and user/postimplementation. Those who already have an EMR are able to share their perceived experienced impact of use, whereas those who don't will have perceived expected impacts (ie, perceived benefits or concerns) that may hinder or drive adoption. One way to collect the views of users and nonusers is through the use of surveys. Surveys are commonly used in information systems evaluation [10-12].

In this review, we specifically address survey-based research studies. Surveys, or questionnaires, refer to the actual instruments used to gather data within a survey-based research study [10,13], which is an overall study design methodology. A survey instrument can be used as one data collection method within another methodology, but here we focused on studies where a survey was the primary means of gathering data. This review offers three contributions to researchers and practices planning to use this approach in future evaluation studies. First, it determined which areas of EMR impact have been evaluated using surveys and which areas have not. This provides an indication of what future survey research could address or where there are prior results available for comparison. Second, it describes a detailed approach, addressing the recognized challenge of reconciling results across heterogeneous studies [14,15], to synthesize the results and present summaries of views for the most prominent impact areas. Third, it contributes to the knowledge regarding survey-based research in the context of HIS evaluation and highlights quality issues to help inform the design of future studies. Therefore, the three key questions for this review are as follows. (1) What areas of EMR impact have been addressed most in survey-based papers (and subsequently, which areas have received little attention)? (2) For those areas that have been addressed most, what have the subjective views been so far regarding the perceived impact of EMRs among users and nonusers in some key areas? (3) How have survey-based studies been designed and used so far, and what are some common quality issues that should be considered?

Methods

Paper Selection

We briefly summarize the search strategy here, with selection flow details available in Multimedia Appendix 1. This review initially began as part of a larger systematic review on the impact of EMRs on physician office practices. In consultation with a librarian, we constructed queries for two online databases: Medline and CINAHL. As well, citations from systematic reviews on HIS were considered. Preliminary screening was carried out by one reviewer, followed by a full-text review done by two teams of two reviewers. Final selection decisions were made by consensus. Included papers had to evaluate the perceived impact of using an EMR system and its clinical functionality in an office-based practice, be based on original data, have providers as the primary end user, and report outcome measures that related to the positive or negative impact of the system. Since not all papers used the term *electronic medical record* and some used other terminology for such systems, we made decisions based on descriptions provided in papers and whether they discussed clinical functionality associated with an EMR. In this review we followed the definition provided by the Canadian Medical Protective Association [16] that an EMR generally refers to an electronic version of the paper record and is specifically used in ambulatory physician practices with added functionality to support clinical care. Please note that we use the term *electronic medical record* in this review for consistency when discussing papers regardless of the term used in the original paper. Papers were excluded if they were studies of outpatient EMRs integrated with inpatient information systems or conducted at hospital-based outpatient clinics, had patient respondents, reported on the same data as another paper, or had inadequate reported data on outcome measures, or if it was not possible to distinguish relevant results. At this stage, the set of selected papers included analytic, descriptive, and survey-based studies on the impact of EMRs on office-based practices. A

separate review was published on the set of analytic and descriptive papers [17], and the remaining survey-based studies were included in this review. For this review, we extended our search to include survey-based studies published between 2000 and 2012 (May).

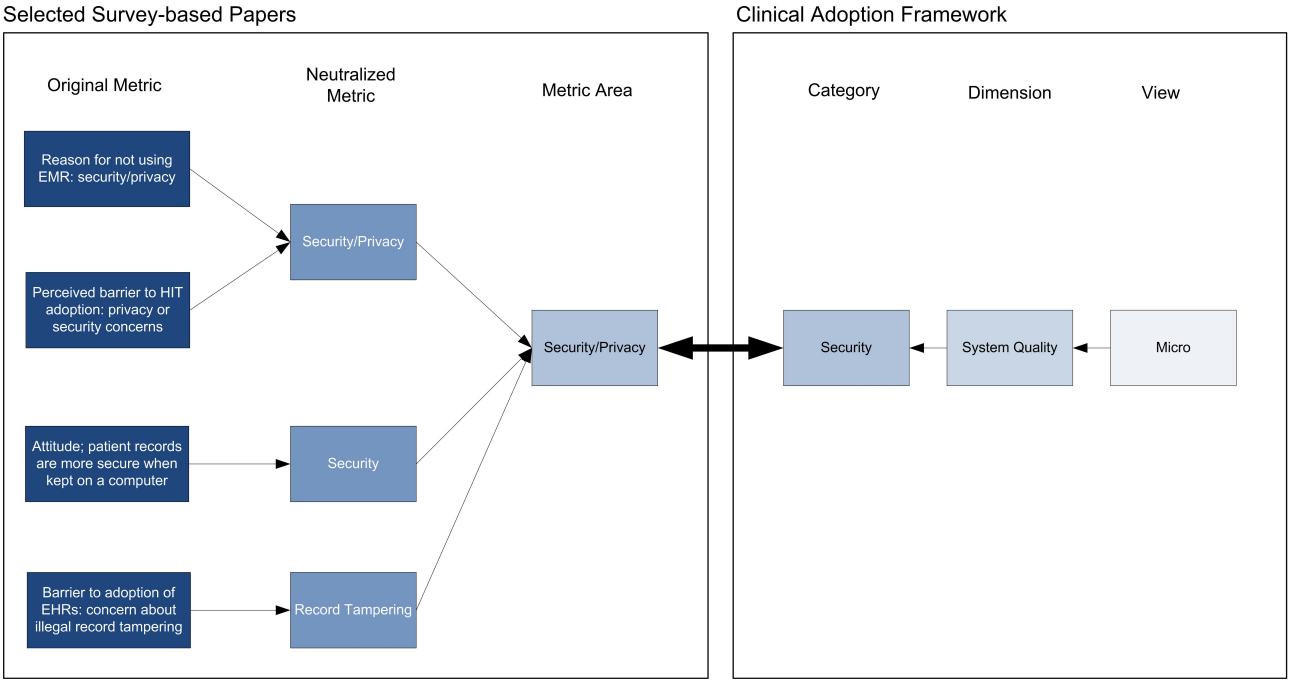
Quality Assessment

In terms of design quality, we considered methods and reporting quality as well as the constructs for evaluation. For study methods design quality, we used the set of 9 survey methodological attributes developed by Grover et al [18], where those with an asterisk can be assessed further for strength of study design: (1) report the approach used to randomize or select samples*, (2) report a profile of the sample frame, (3) report characteristics of respondents, (4) use a combination of personal, telephone, and mail data collection, (5) append the whole or part of the questionnaire, (6) adopt a validated instrument or perform a validity or reliability analysis*, (7) perform an instrument pretest, (8) report on response rate*, and (9) perform a statistical test to justify the loss of data from nonrespondents*. Although these attributes were originally designed for management information systems survey research, we believe they are applicable to the HIS context as well. To score papers, each attribute is given a score of 1 or 0 depending on whether it is present in the paper or not, respectively. The final score for each paper is then determined by adding all attribute scores. Like the creators of survey methodological attributes, we assumed that unreported attributes were not performed and assigned a score of 0. A limitation noted for the survey methodological attributes is that the dichotomous measures don't capture richness of some variables. We modified the scoring in our review to allow attributes to be given a score of 0.5 if they were somewhat present or weak. Scoring was done independently by two reviewers, and final scores for each paper were determined through consensus. All papers were included in the review for the insights they offered, but it is important for the reader to be aware of quality shortcomings. Ju et al [19] used survey methodological attributes to highlight the quality of research and considered a journal article to be adequate if the total survey methodological attributes score was greater than 4.5 out of a possible 9. We used the same interpretation. The intent here is to demonstrate how different papers addressed the 9 attributes and to highlight areas for improvement across the HIS evaluation field.

Data Extraction

For the first step in the data extraction process, we identified all survey items and questions from each paper, which we termed *metrics*. For example, privacy and confidentiality concerns as a potential barrier to EMR adoption were one metric. The next step was to organize all these metrics in a way that would facilitate meta-analysis. The challenge in combining data from several papers that address a common area is reconciling the constructs used for evaluation within each paper. According to Ju et al [19], "a critical process in the maturing of any discipline is the development of proper constructs and instruments to collect adequate and accurate data about phenomena of interest." To determine these constructs for synthesis, we mapped all extracted metrics to categories of the Clinical Adoption Framework (see Multimedia Appendix 2). The Clinical Adoption Framework and its predecessor, the Benefits Evaluation Framework, are based on DeLone and McLean's Information System Success Model [9]. The Clinical Adoption Framework groups evaluation categories into dimensions, which are then organized according to three views: meso, macro, and micro. The meso view is concerned with people, the organization, and implementation as a whole. The macro view considers environmental factors that have a direct influence on categories in the meso view such as standards, funding, and societal trends. Finally, the micro view is focused on the user level and net benefits in specific areas where the system is expected to have an impact. In 2003, Van Der Meijden et al [20] reviewed evaluations of patient care information systems and categorized them using DeLone and McLean's framework. They identified attributes from each paper and placed them into the framework categories. For our review, a metric area is equivalent to an attribute and is defined as the general aspect being evaluated, such as security. After extracting the reported metrics, we classified them into general metric areas and mapped them to Clinical Adoption Framework categories to allow for comparison and synthesis. Figure 1 shows how this was done. This mapping identified the main constructs for evaluation addressed in the papers. Finally, the last step was to split the corresponding results according to nonusers and users where possible. The aim was to extract only the data related to EMRs based on clinical functionality. Some papers included results on general information technology use such as Internet or email. We did not extract data pertaining to these except in cases where general functionality was embedded into clinical functionality.

Figure 1. Example mapping of metrics to the Clinical Adoption Framework. EHR = electronic health record, EMR = electronic medical record, HIT = health information technology.



Meta-Analysis

The goal of the meta-analysis was to identify the most commonly addressed areas and combine the reported results for these areas to determine users' and nonusers' overall views toward EMRs. We determined that the raw data presented in some papers needed to be transformed to make them comparable. The first step was to consider whether the metric was posed as negative or positive so that the reported results could be interpreted as either negative or positive.

The surveys collected two types of data: dichotomous (ie, proportions or percentages for agreement with statements) and categorical (eg, Likert-type scale scores), and they were not reported in the same manner in all papers. For the dichotomous data, if the result was not already expressed as a proportion, we calculated a proportion estimate based on the sample size reported in the paper. As well, some papers divided results into further groupings within the nonuser and user categories, so we pooled these where possible using 95% confidence intervals to confirm an overlap for pooling. We created a series of 2 × 2 tables to organize the reported results for each metric with respect to positive and negative views for users and nonusers. Using the tables, we calculated the estimated odds of a perceived positive view for users and nonusers and then, where possible, an estimated odds ratio for a positive view for users to nonusers. For the categorical data, we redefined the scales used in the papers where needed to make mean values comparable. Most papers used a 5-point scale, but it was sometimes reversed or used different values. We transformed each scale so that it ranged from 1 (strongly negative) to 5 (strongly positive). Mean scores were recorded for nonusers and users where possible. The resulting odds calculations and mean scores were interpreted and compared with reported findings in the papers to determine overall perceived views for each selected area. Positive views leaned toward more perceived benefits of the potential use of systems, whereas negative views represented more perceived concerns or barriers that could possibly hinder use.

Results

General Characteristics of Selected Papers

We selected 19 survey-based papers for inclusion in the review [21-39], presenting both user and nonuser views (see Table 1). Several papers presented results for both categories [21,22,25-27,29,33-35,37,38], 1 looked only at preimplementation [36], and a few were only postimplementation views [23,24,30-32,39]. The implementation state of respondents' EMR systems was not clear in 1 paper [28].

Table 1. Papers reporting results for the two categories of use status.

Author (year)	Preimplementation/ nonusers	Postimplementation/ users	Not specified
Chiang et al (2008) [21]	X	X	
DesRoches et al (2008) [22]	X	X	
Devine et al (2010) [23]		X	
El-Kareh et al (2009) [24]		X	
Gans et al (2005) [25]	X	X	
Johnston et al (2002) [26]	X	X	
Kemper et al (2006) [27]	X	X	
Leung et al (2003) [28]			X
Loomis et al (2002) [29]	X	X	
MacGregor et al (2006) [30]		X	
Mackenzie (2006) [31]		X	
Magnus et al (2002) [32]		X	
Menachemi et al (2007) [33]	X	X	
Russell and Spooner (2004) [34]	X	X	
Simon et al (2007) [35]	X	X	
Simon et al (2008) [36]	X		
Simon et al (2008) [37]	X	X	
Singh et al (2012) [38]	X	X	
Terry (2005) [39]		X	

Table 2 presents the general characteristics reported in the papers. In terms of respondents, the majority of surveys were administered to the physicians themselves [22-24,26,29,30-34], including specialists [33,39] such as pediatricians [27,34] and ophthalmologists [21]. In some cases, the respondents were the entire practice [25,38], nurses [23,31], or administrative office staff [37].

Many papers aimed to determine the perceived impact of adoption [21,22,24,25,27,35,36,39]. Others were specifically concerned with assessing perceived barriers to adoption [22,25,27,29,32-35,37,38], benefits [21,22,25,27,28,30,38], or overall attitudes toward adoption or use [22-24,26,27,29,31,32,34,36]. A few looked at physician satisfaction in general [21,22,39] and 2 focused on specific functionality [31,32]. (Note that these categories are not mutually exclusive.) Of the selected papers, 4 considered information and communication technology in general [26,30,36] but described features of EMRs, so we included them as well.

Table 2. General paper characteristics.

Author (year)	Country	Survey/study objective(s)	Respondents	Clinical context (ie, setting)	Survey method	Total sample	Response rate
Chiang et al (2008) [21]	United States	Assess the state of EHR[a] use by ophthalmologists, including adoption rate and user satisfaction	Ophthalmologists	Medical practices	Web-based survey (with 2 email reminders) and telephone survey	3796	15.6% (592)
DesRoches et al (2008) [22]	United States	Assess (1) physicians' adoption of outpatient EHRs, (2) satisfaction with such systems, (3) perceived effect of the systems on the quality of care, (4) perceived barriers to adoption	Physicians	Physicians providing direct ambulatory patient care	Mailed questionnaire (2 reminders by mail and phone); cash incentive	5000 (4484 eligible)	62% (2758)
Devine et al (2010) [23]	United States	Identify prescriber and staff (end user) characteristics that would predict attitudes and behaviors toward e-prescribing adoption in the context of a preexisting EHR	Prescribers (physicians, physician assistants, nurse practitioners) and staff (nurses and medical assistants)	3 primary care sites	Administered at the sites with 2 reminders sent via email	Total of 188 opportunities	Overall: 62% (117); prescribers: 82%; staff: 50%
El-Kareh et al (2009) [24]	United States	Measure changes in primary care clinician attitudes toward an EMR[b] during the first year following implementation	Physicians, nurse practitioners, physician assistants	Ambulatory health centers	Mailed questionnaire at 1, 3, 6, and 12 months postimplementation (2 mailings and reminder emails)	73 physicians; 10 nurse practitioners; 3 physician assistants	Month 1: 92% (79); month 2: 95% (81); month 3: 90% (76); month 12: 82% (69)[c]
Gans et al (2005) [25]	United States	Assess the rate and process of adoption of information technology and EHRs by medical group practices	Group practices	Group practices with 3 or more physicians practicing together with a common billing and medical record system	Web-based and mailed survey; a subset of nonresponders were surveyed by phone	17,195	21.1% (3628)
Johnston et al (2002) [26]	China	Identify prevailing attitudes among physicians to use of computers in the clinical setting and specifically those attitudes that may be associated with the adoption of computers in practice	Physicians	Individual practices	Mailed questionnaire	4850	18.5% (897)
Kemper et al (2006) [27]	United States	(1) Measure penetration and functionality of EMRs in primary care pediatric practice, (2) identify plans for adoption of EMRs, (3) understand common barriers to adoption, (4) evaluate attitudes toward EMRs among those with and without one	Pediatricians	Office-based practice	Separate mailed questionnaires to those with and without an EMR (3 mailings); cash incentive	1000 (901 eligible)	58% (526)
Leung et al (2003) [28]	China	Understand the contributory barriers and potential incentives associated with information technology implementation	Physicians	General physician population (individual and corporate settings)	Mailed survey (3 mailings and maximum of 7 phone calls)	949	77% (731)
Loomis et al (2002) [29]	United States	Investigate possible differences in attitudes and beliefs about EMRs between EMR users (early market) and nonusers (mainstream market)	Family physicians	Active members in the Indiana Academy of Family Physicians	Mailed survey (2 mailings)	1398	51.7% (618 usable)

Author (year)	Country	Survey/study objective(s)	Respondents	Clinical context (ie, setting)	Survey method	Total sample	Response rate
MacGregor et al (2006) [30]	Australia	(1) Examine perception of benefits derived from information technology adoption, (2) determine whether practice size, number of patients treated, gender of practitioner, or level of computer skills of the practitioner are associated with the perception of benefits	General practitioners	General practice	Mailed questionnaire	690	17.7% (122)
Mackenzie (2006) [31]	New Zealand	Nurses' and doctors' perceptions of the introduction and subsequent use of the Medtech 32 clinical module	Nurses, doctors	Family planning clinics	Paper questionnaire	132	57% (47 nurses and 28 doctors)
Magnus et al (2002) [32]	England	(1) Assess general practitioners' views on the relevance of information provided by computerized drug interaction alert systems, (2) determine the proportion of general practitioners who admit to frequently overriding alerts without properly checking them, (3) explore factors that might be associated with a tendency to override alerts	General practitioners	Primary care trust areas	Mailed questionnaire (2 mailings)	336	70% (236)
Menachemi et al (2007) [33]	United States	1. Examine rural–urban differences in the use of various information technology applications by physicians in the ambulatory setting	Physicians (family medicine, internal medicine, pediatrics, obstetrics and gynecology)	Ambulatory settings	Mailed questionnaire (2 mailings)	14,921	28.2% (4203)
Russell and Spooner (2004) [34]	United States	(1) Determine the use of EMRs in area practices, (2) identify physicians' attitudes adopting EMRs, particularly differences in attitudes between users and nonusers and between internal medicine and pediatric clinicians	Physicians (internal medicine and pediatrics)	Medical outpatient practices of internal medicine and pediatrics	Faxed and mailed survey (3 faxes and mailing); cash incentive	51 internal medicine, 24 pediatrics	Internal medicine: 51% (26); pediatrics: 63% (15)
Simon et al (2007) [35]	United States	(1) Determine the degree to which physicians used the various functions available in their EHR systems, (2) identify factors that correlate with use	Physicians	Office-based practice	Mailed survey (3 mailings with phone calls in between); cash incentive	1921 (1884 eligible)	71.4% (1345)
Simon et al (2008) [36]	United States	(1) Assess the degree to which the MAeHC[d] practices are representative of physician' practices statewide, (2) assess practice characteristics related to EHR adoption, prevailing office culture related to quality and safety, attitudes toward HIT[e], and perceptions of medical practice	Physicians	Physician office practices	Mailed survey with multiple reminders	MAeHC: 464; statewide: 1884	MAeHC: 77% (355); statewide: 71.4% (1345)[f]

Author (year)	Country	Survey/study objective(s)	Respondents	Clinical context (ie, setting)	Survey method	Total sample	Response rate
Simon et al (2008) [37]	United States	(1) Determine the state of EHR adoption and the degree to which doctors with EHRs are using the functionalities of those systems, (2) assess whether practices that had not yet adopted EHRs planned to adopt such systems and when, and what barriers impeded their progress	Office practice managers	Active medical and surgical practices (hospital and non-hospital based)	Mailed questionnaire (2 mailings and 2–6 phone calls)	1829	46% (847)
Singh et al (2012) [38]	United States	(1) Examine HIT and EMR adoption and use among primary care offices across the rural–urban spectrum, (2) assess perceived benefits and perceived barriers and facilitators to adoption	Offices (targeted office medical directors or owners)	Primary care offices	Mailed survey (reminder and second mailing); cash incentive	4669	21.4% (1001)
Terry (2005) [39]	United States	Determine EHR penetration, satisfaction, and use	Medical doctors and doctors of osteopathic medicine (including family practitioners, general practitioners, internists, obstetricians and gynecologists)	Office-based practice	Mailed survey	10,000	Not reported

[a] Electronic health record (term used in the paper).

[b] Electronic medical record.

[c] Only included month 12 data in analysis.

[d] Massachusetts eHealth Collaborative.

[e] Health information technology.

[f] Only included Massachusetts eHealth Collaborative data in analysis, as statewide data are reported in Simon et al [35].

Quality

Using the survey methodological attributes, we deemed more than half of the papers (12) to be of adequate quality (see Table 3). The items with the highest average scores were reporting a profile of the sample frame, with a response rate, and a profile of respondents. For sample frame we looked for inclusion and exclusion criteria that identified the target sample. In many cases the sample frame was all clinicians belonging to a membership or organization. Most papers reported a response rate and respondent characteristics in a table. A few papers provided demographic information at the practice level, rather than the individual respondent level, so we scored these as 0.5 for this item. The item that scored most poorly across all papers was analyzing the reliability or validity of item measurement or adopting a validated instrument. For this item we specifically looked for an indication of an analysis done to confirm the validity or reliability of instrument questions. Many papers reported developing instruments in consultation with experts, basing the questionnaire on previous work, or having an expert panel review it. However, only 1 paper [29] specifically reported having a test-retest reliability rate for each item. The other item that generally scored low was for the use of a combination of personal, telephone, or mail data collection. In most cases, survey data were collected solely through a mailed questionnaire. A few papers reported an opportunity for respondents to complete the survey by Web or telephone as well. We scored these as 1.

Table 3. Quality assessment using the survey methodological attributes.

	Criteria items[a]									Total
Author (year)	1	2	3	4	5	6	7	8	9	score[b]
Leung et al (2003) [28]	0.5	1	1	1	1	0.5	1	1	1	8
Chiang et al (2008) [29]	0.5	1	1	1	1	0.5	1	1	0	7
Singh et al (2012) [38]	1	1	0.75	0	1	0	1	1	1	6.75
DesRoches et al (2008) [22]	0.5	1	1	0	1	0	1	1	1	6.5
Gans et al (2005) [25]	1	1	0.5	1	1	0	0	1	1	6.5
Magnus et al (2002) [32]	1	1	1	0	0.5	0	1	1	1	6.5
Devine et al (2010) [23]	0.5	1	1	0.25	1	1	0.5	1	0	6.25
Loomis et al (2002) [29]	1	1	1	0	0	1	1	1	0	6
Menachemi et al (2007) [33]	0.5	1	1	0	0	0.5	0.5	1	1	5.5
Simon et al (2008) [37]	1	1	0	1	0	0	0	1	1	5
MacGregor et al (2006) [30]	1	0.5	1	0	1	0	0	1	0	4.5
Simon et al (2007) [35]	0.5	1	1	1	0	0	0	1	0	4.5
El-Kareh et al (2009) [24]	1	1	1	0	0	0	0	1	0	4
Kemper et al (2006) [27]	0.5	1	0.5	0	0	0	1	1	0	4
Russell and Spooner (2004) [34]	1	0.5	0.5	0	1	0	0	1	0	4
Simon et al (2008) [36]	1	1	0.5	0	0	0	0	1	0	3.5
Johnston et al (2002) [26]	0.5	1	1	0	0	0	0	1	0	3.5
Mackenzie (2006) [31]	0.5	1	1	0	0	0	0	0.5	0	3
Terry (2005) [39]	0	0	0	0	0	0	0	0.5	0	0.5

[a] 1 = sample selection approach, 2 = profile of sample frame, 3 = respondent characteristics, 4 = data collection methods, 5 = sample of questionnaire, 6 = validation of instrument, 7 = instrument pretest, 8 = response rate, 9 = test for nonrespondent.

[b] Out of a possible maximum score of 9.

Reported Metric Areas

During the data extraction phase of the review, we pulled reported metrics from the papers and grouped them into more general metric areas under the categories of the Clinical Adoption Framework. Only those metrics related to EMRs were extracted, which excluded general information technology. However, metrics related to use of other clinical information technology were extracted into the category of information and infrastructure. For example, Gans et al [25] asked respondents about use of other computer-based information systems, and Simon et al [37] reported whether having computerized claims or billing systems, computerized scheduling systems, or computerized prescribing systems was associated with adoption. Table 4 provides an overview of the mapping of metric areas addressed in the papers to categories of the Clinical Adoption Framework. The most-addressed categories were personal characteristics, structure and processes, stage, appropriateness and effectiveness, efficiency, and functionality. These appeared to have received the most attention in the surveys reported. Sometimes a paper had multiple metrics for a category. For a more detailed analysis, we split the metric areas into three groups: background, impact-specific, and other. Each group is described in the section below.

Several Clinical Adoption Framework categories did not have any metric areas identified: individual and groups, roles and responsibilities, added values, legislative acts, political trends, economic trends, use behavior pattern, intention to use, and participant and caregiver participation.

Background Areas

Most background areas corresponded to categories under the meso level, since surveys often had items pertaining to the background of respondents and practices such as practice size (number of staff), system use status, gender, future intention to use a system, and specialty. These areas and their metrics were often used not only to describe the sample, but also to determine whether there were any correlations within the reported findings. For instance, several papers found that system adoption and use was greater in larger practices [22,25,27,33,35]. We chose use status as the main categorical variable for our meta-analysis to help split results into nonuser/preimplementation and user/postimplementation categories.

Other Areas

We also extracted all other areas addressed that related to EMR adoption. These were not determined to be specifically impact related but were other associated aspects of EMR implementation that have been addressed through surveys. Table 4 shows that most tended to correspond to categories under the

macro level. Expense of implementation and functionality, either available or desired according to respondents, came up frequently. We included the use of features as a separate area, as availability of a feature doesn't necessarily correspond to use. A few papers [22,35,37] made this distinction.

Impact-Specific Areas

The third group of addressed areas were the ones of interest for this review. These areas specifically addressed the perceived potential or actual impacts of implementing and using an EMR. As shown in Table 4, most impact areas corresponded to categories at the micro level as expected, specifically under the net benefits dimension. No impact-related metrics mapped to the macro level, but a few did map to categories in the meso level. It is important to note there would not necessarily be impact-specific areas for every category of the Clinical Adoption Framework because it encompasses all aspects of system adoption.

Table 4. Mapping of metric areas to clinical adoption framework.

Level	Dimension	Category	Metric area	Type[a]	Papers (reference number)	Total number of metrics
Meso	People	Individuals and groups	(Determined by type of respondent survey is administered to)		All	0
		Personal characteristics	Age	B	23, 26, 28, 29, 31, 32, 33, 36, 39	9
			Gender	B	22, 23, 24, 26, 28, 29, 30, 32, 33, 35, 36	11
			Race and ethnic background	B	22	1
			Income	B	28	1
			Active in general practice and status	B	35	2
			Graduation year and years of practice	B	22, 24, 26, 34, 35, 36	6
			Specialty	B	22, 23, 26, 28, 33, 34, 35, 36, 37, 38, 39	11
			Computer skills and literacy	B	23, 26, 30, 31, 34, 36	9
			First to have new tests or treatments (general practice)	O	36	1
		Personal expectations	Comparison between paper based and electronic	I	27	1
			Feelings toward practice in general	O	35, 36	8
			Protecting physicians from personal liability for record tampering by external parties	I	22	1
		Roles and responsibilities				0
	Organization	Strategy	Actively improving quality (general practice)	O	36	1
			Local physician champion	O	38	1
			Physician recruitment	I	25	1
		Culture	Bad previous experience with an electronic record system	O	27	1
			Attitude toward the electronic record system	I	22, 24, 25, 26, 27, 29	4
			Physician and staff resistance	O	36, 37	9
			Isolation from colleagues (general practice)	O	35, 36	2
			Innovative staff (general practice)	O	36	2
		Information and infrastructure	Ability to interface and integrate with existing practice systems	O	21, 25, 27, 39	6
			Technical limits	O	36	1
			Use of other clinical information technology	O	25, 37, 38	4
		Structure and processes	Practice size (number of staff)	B	21, 22, 25, 26, 27, 28, 29, 30, 33, 34, 35, 36, 37, 38, 39	18
			Practice size (number of patients)	B	24, 29, 30, 35, 36	5
			Practice size (number of offices)	B	38	1
			Time spent caring for patients (hours)	B	24, 26, 28	3
			Practice type (eg, group)	B	26, 28, 33	3
			Remuneration patterns	B	26, 28	2
			Practice setting (eg, hospital or medical center)	B	22, 37	2
			Type of office	B	23, 38	4

Level	Dimension	Category	Metric area	Type[a]	Papers (reference number)	Total number of metrics
			Patient population	B	38	2
			Practice location	B	22, 29, 33, 36, 37, 38	7
			Communication with general practice business suppliers	O	30	1
		Return on value	Business expansion	I	30	1
			Expense of implementation	O	21, 22, 25, 26, 27, 28, 29, 33, 36, 37, 38	13
			Maintenance costs	O	21, 27, 26, 29, 33, 36	7
			Expected return on investment	I	22, 25, 27, 33, 34, 38, 39	7
	Implementation	Stage	Use status	B	21, 22, 25, 26, 27, 29, 32, 33, 34, 35, 36, 37, 38, 39	16
			Future intention to use	B	21, 22, 23, 27, 33, 34, 37, 38, 39	12
		Project	System development or selection	O	21, 22, 25, 27	5
			Time costs associated with computerization	I	21, 25, 26, 28, 33, 36	7
			Loss of productivity during transition	I	22, 33, 36, 38	5
			Entering historical data	O	25	1
		HIS[b]–practice fit	Staff requirements for implementation and maintenance	O	26, 27	2
			Meeting needs and requirements	O	22, 25, 27, 33, 37	5
			Capital available for practice expansion	O	36	1
Macro	Health care standards	HIS standards	Standardized medical terminology	O	21	1
			Transience of vendors	O	27	1
			Uniform data standards within the industry	O	25, 33, 36	3
		Performance standards	Evaluation of changes to improve quality (general practice)	O	36	1
			Quality problems (general practice)	O	36	1
			Procedures and systems to prevent errors (general practice)	O	36	1
		Practice standards	Adding to the skills of the practice	O	30	1
			Standardized questions to ask vendors	O	21, 25	2
			Model requests for proposal for contracts	O	21, 25	2
	Funding and incentive	Remunerations	Payment for having or using system	O	22, 36	3
			Payment for patient survey results or clinical quality	O	36	2
			Direct financial assistance	O	25, 38	2
		Added values				0
		Incentive programs	Financial incentives for purchase and implementation	O	21, 22, 25, 28, 35, 38	6
			Clarity of benefits	O	28	1
	Legislation, policy and governance	Legislative acts				0

Level	Dimension	Category	Metric area	Type[a]	Papers (reference number)	Total number of metrics
		Regulations and policies	Confidentiality	O	22, 27, 28, 29	4
			Access and sharing of to medical records	O	22, 29	2
			Intellectual property regulations	O	28	1
			Self-referral prohibitions regarding sharing of technology	O	25	1
			Government regulation requiring mandatory reporting of patient information	O	28	1
		Governance bodies	Vendor certification and accreditation	O	21, 25, 38	3
			Legal liability	O	22	1
	Societal, political and economic trends	Societal trends	Competitive peer pressure in terms of more practices becoming computerized	O	28	1
			Recommendations of colleagues	O	38	1
			Public or patient views for computerization	O	26, 28, 33	3
		Political trends				0
		Economic trends				0
Micro	System	Functionality	Features available and functions computerized	O	21, 22, 25, 26, 27, 35, 39	9
			Intention to computerize functions	O	26	1
			Features desired and functions that should be computerized	O	21, 26, 29, 31, 32	10
			Features used	O	22, 26, 35, 37, 38	5
			Features for patient use	O	22	5
		Performance	Reliability of system	I	22, 34	2
			System downtime	I	27, 33	2
			Frequency of potential drug interaction alerts	I	32	1
			How good system is in alerting for significant interactions	I	32	1
			Concern system would become obsolete	O	22	1
		Security	*Security and privacy*	I	22, 25, 26, 27, 29, 33, 34, 35, 36	11
	Information	Availability	Information storage and retrieval	I	30	1
			Reliability of information	I	32	1
			Accessibility of records and information	I	21, 22, 24, 25, 27, 35, 36, 38	11
		Content	Value of clinical records	I	26	1
			Accuracy of records	I	21, 25, 38	3
			Drug interaction alerts providing information that is irrelevant to the patient	I	32	3
			Amount of information provided	I	32	1
			Reason for overriding alert: more faith in other sources of information	I	32	3
			Grading interaction alerts according to severity	I	32	1
	Service	Responsiveness	Training	I	24, 29, 31, 34, 38	8
			Level of support	I	28, 31, 36, 37	4

Level	Dimension	Category	Metric area	Type[a]	Papers (reference number)	Total number of metrics
	Use	Use behavior and pattern				0
		Self-reported Use	Use of information technology for clinical management activities	O	27 (also see functionality)	1
			Overriding alerts	I	32	4
		Intention to use				0
	Satisfaction	Competency	Learning curve	O	21, 25, 27, 28, 33	6
		User satisfaction	Overall satisfaction	I	21, 22, 39	4
			Annoyance caused by drug interaction alert messages	I	32	1
			Usefulness in prescribing	I	23, 32	2
			Ease of use of system or clinical module	I	22, 23, 31, 33	5
		Ease of use	Data entry	I	25, 27, 29, 33, 38	5
			Interface and customization	I	39	1
	Net benefits	Quality: patient safety	Primary care and medical errors	I	27, 29	3
			Medication-related errors	I	22, 24, 25, 35, 36, 38	8
		Quality: appropriateness and effectiveness	Disease prevention or management	I	22, 30, 38	5
			Clinical decision making	I	22, 25	3
			Clinical functions	I	26	1
			Prescriptions	I	22, 25, 30	3
			Legibility	I	21	1
			Frequency of change in initial prescribing decision due to drug interaction alerts	I	32	1
			Awareness of information provided by drug interaction alerts	I	32	2
			Effect of computer use on patients' satisfaction with care received	I	34	1
			Patient–physician relationship and communication	I	21, 22, 24, 25, 26, 27, 28, 34, 35, 36	10
			Documentation	I	27	2
			Effect on medical practice; practice style	I	39	1
		Health outcomes	*Quality of patient care or clinical outcomes*	I	21, 24, 26, 27, 28, 29, 31, 35, 36	12
		Access: ability of patients and providers to access services	Remoteness in the provision of medical care	I	30	1
			Patient or customer base and area of coverage	I	30	1
		Access: patient and caregiver participation				0
		Productivity: efficiency	Accounting and billing or charge capture	I	21, 25, 27, 30	7
			Assistance in test ordering and management	I	22, 24	3
			Documentation time	I	21, 24	3

Level	Dimension	Category	Metric area	Type[a]	Papers (reference number)	Total number of metrics
			Business or practice efficiency	I	21, 27, 28, 30, 33, 34, 35, 36, 39	10
			Time for medication refills	I	38	1
			Time for patient care	I	24, 26, 30	3
			Workload	I	27, 30	4
		Productivity: care coordination	*Communication with other providers*	I	22, 24, 27, 30, 35, 36	8
			Workflow	I	21, 25, 27, 33, 37	5
		Productivity: net cost	*Costs or savings*	I	21, 25, 27, 28, 30, 35, 36	10

[a] B = background, O = other, I = impact-specific area.

[b] Health information system.

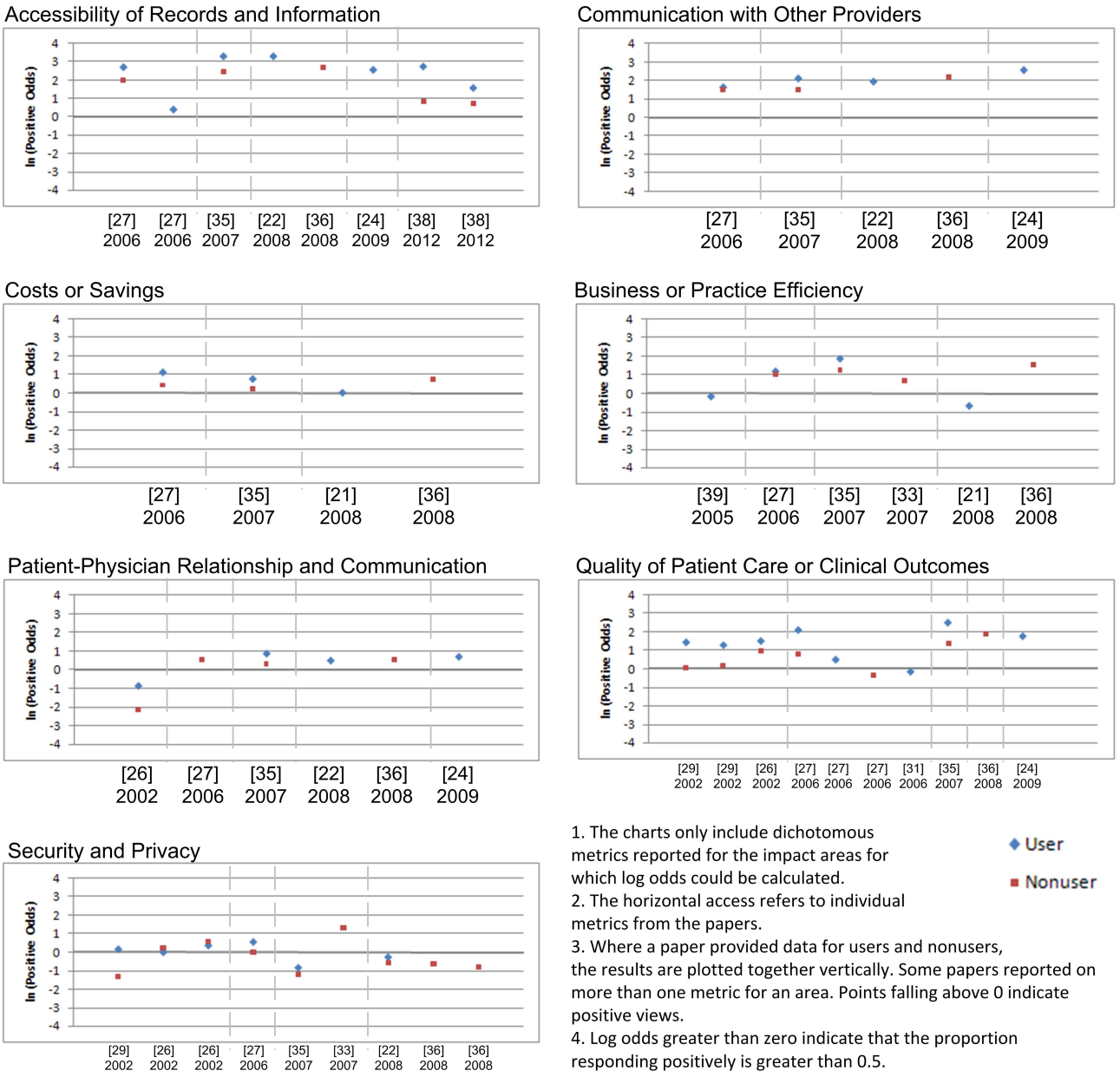
Our mapping of metrics to the Clinical Adoption Framework resulted in no metrics for the category patient and caregiver participation, and only 2 metrics for the ability of patients and providers to access services. Together, these make up the access category of net benefits. However, 1 paper [22] did report on available functions for patients, which could potentially be considered patient participation. However, since the survey reported on these only as available functionality within the system and not perceived impact, we classified them under other metrics for functionality. All metrics falling under functionality and competency were classified into either background or other areas.

We found appropriateness or effectiveness and efficiency to be the most-addressed areas of impact. The italicized areas in Table 4 were identified as the top impact-specific areas based on frequency of reported metrics and results in the original papers. These were security and privacy, quality of patient care or clinical outcomes, patient–physician relationship and communication, communication with other providers, accessibility of records and information, business or practice efficiency, and costs or savings.

Selected Area Findings

We synthesized reported data for the 7 top impact-specific areas using the meta-analysis approach described above. The estimated log odds for users and nonusers are graphed in Figure 2 (see Multimedia Appendix 3 for calculations and assumptions, including proportions, standard errors, and confidence intervals).

Figure 2. Estimated log odds for selected impact-specific areas.



Accessibility of Records and Information

Our meta-analysis showed that both users and nonusers viewed EMRs as having a positive impact on accessibility of information. For this area, 8 papers [21,22,24,25,27,35,36,38] reported on 10 different metrics. We were able to compare users with nonusers in 3 papers and found that users were more likely than nonusers to have a perceived positive view in all 3 papers. Looking only at users, we found many more positive views in 6 papers. For example, the odds of a positive view were over 3 in 3 papers [22,35,38]. The views of nonusers were also more positive in all the papers we reviewed. As well, mean scores reported in 2 papers [21,25] pointed to positive views for both users and nonusers. In both of these, improved access was the top-rated benefit.

El-Kareh et al [24] found no noted improvement in this measure from baseline to month 12 but that virtually all clinicians reported immediate improvement in this measure at study baseline.

Communication with Other Providers

Both users and nonusers perceived a positive effect on communication with other providers. A total of 6 papers [22,24,27,30,35,36] reported on 8 metrics for communication with other providers and health care professionals. We calculated odds ratios for 2 papers and found that, in both cases, users were likelier to hold a positive view. All the calculated odds, regardless of whether they were for users or nonusers, indicated more positive views with respect to the impact of EMRs on communication among providers. We calculated mean scores for 3 metrics reported in 1 paper [30]. All 3 scores indicated neutral to positive views. The authors in this paper also performed a series of linear regressions to determine any associated factors. They found that those who had a higher level of skill with the use of medical packages saw a greater benefit from communication with other medical organizations, and

larger practices saw a greater benefit for this and communications with fellow general practitioners. El-Kareh et al [24] performed a longitudinal study over the first year following implementation and concluded that within 1 year of implementation a vast majority of clinicians felt that it improved communication among clinicians.

Costs or Savings

Reduction in costs or savings was generally seen as an important impact of implementation in the majority of papers we reviewed. For the meta-analysis we only included metrics that reported on impact on practice costs after implementation for net benefits. Several others assessed views on costs to implement and maintain systems, but we categorized these under return on value. For the meta-analysis we looked at 10 metrics across 7 papers [21,25,27,28,30,35,36]. We found that those with a system were more likely to have positive views of the impact on costs or savings in 2 papers [27,35] that compared users versus nonusers. Chiang et al [21] asked respondents about the effect of implementation on practice costs after 6 months and found that the systems were associated with decreased or comparable practice costs. Gans et al [25] reported several mean scores for different types of costs for users, and all indicated that the system would be a benefit in reducing costs. The one unclear result was in the study by MacGregor et al [30], where reduction of costs was seen as almost neutral, ranking as the 12th benefit.

Business or Practice Efficiency

In the small number of papers reviewed, improvement in business or practice efficiency was seen as a benefit of implementation. There were a total of 10 metrics reported in 9 papers [21,27,28,30, 33-36,39] addressing business or practice efficiency. In the 2 papers providing dichotomous data for nonusers and users, users were more likely to have a positive view, and both groups individually had more positive views. MacGregor et al [30] gathered categorical data on the importance of efficiency and operation as a benefit of adoption, and the results leaned toward the positive side. The authors found that those with a higher perceived level of ability with medical software packages saw higher benefits. In the study of Leung et al [28], improved efficiency was the most attractive incentive for computerizing clinical practice. In 2 papers [21,39], users indicated more negative views, but we assumed that no change in productivity was negative in one case [39] and that same, decreased, or unsure was negative in the other [21].

Patient–Physician Relationship and Communication

Of all papers reporting on patient–physician relationships and communication, views appeared to be generally positive, with two exceptions. For this area, 9 papers [22,24-28,34-36,] reported on 9 related metrics. The odds ratios for 2 papers produced mixed results. One paper [35] showed positive results for both users and nonusers and a higher likelihood of a positive view for users. The other paper [26] reported more negative views for users and nonusers. However, users were about 3.5 times more likely to have a less-negative view, which aligned with the authors' conclusion that clinical users (ie, those with one or more clinical functions computerized) were less negative about potential interference with the doctor–patient encounter. The mean scores reported for users supported more positive views.

For nonusers, we calculated negative views in 2 papers. As mentioned above, in the study of Johnston et al [26], nonusers were more negative. As well, in Russell and Spooner's study [34], nonusers felt an EMR would have a negative impact on doctor-patient interaction.

Quality of Patient Care or Clinical Outcomes

A total of 8 papers [24,26-29,31,35,36] reported on 11 metrics related to perceived impact on quality of care or outcomes. We found that improvement to quality of care was generally seen as a benefit of implementation in these, except for 2 papers with results indicating otherwise. In 5 papers, the odds of a perceived positive view of users versus nonusers indicated a higher likelihood of a positive view for users. The individual odds of a positive view calculated for users and nonusers were all positive as well in these papers. Looking at only user views, only 1 paper reported that users had a negative view [31].

Nonusers had more positive views for 6 out of 7 metrics according to our calculations. The exception was in the study of Kemper et al [27], where the respondents were pediatricians, and more than half of those without a system cited lack of belief that these systems improve care as a barrier to adoption. But this paper also reported that for users, a common reason for adopting was to improve quality of care. Loomis et al [29] reported that for nonusers there was a considerable lack of belief that EMRs would improve quality, but our calculations still found the odds of a positive view to be greater than 0.

One paper [28] presented a mean score indicating that respondents (users and nonusers) found higher quality of care to be the one of the most attractive incentives for computerizing their clinical practice.

Security and Privacy

Privacy and security appeared to be an area of mixed perceptions regarding the impact of EMRs. A total of 9 papers [22,25-27,29,33-36] addressed security and privacy through 11 metrics, which produced mixed results for positive and negative views. Out of the odds ratio estimates calculated for 6 metrics, 4 showed a higher likelihood of positive views for users over nonusers. However, this actually meant that users were more likely to have a less-negative view, because the individual odds showed more negative views in some cases. In 1 paper [35] where we found a higher likelihood of positive views for users but where each group individually had more negative views, the authors concluded that "more than 40% of responding physicians, regardless of whether they used EMRs, said that computers may have a negative effect on patient privacy." For DesRoches et al [22], we found that those with and without an EMR system had more negative views about illegal record tampering. However, the paper interpreted this to mean that protecting physicians from personal liability for record tampering by external parties could be a major facilitator of adoption. A total of 3 papers [27,29,34] specifically reported that more users had more positive perceptions of EMRs than nonusers. Conversely, in 1 paper [26] we compared users with

nonusers, where users were not more likely to have positive views. However, when looking at users and nonusers separately, we found that each group had more positive views.

When looking at only the odds of positive views for users, our calculated results for 4 out of 6 metrics were positive. For nonusers, 3 out of our 9 calculated results were positive. Mean scores were reported in 2 papers, and when these were transformed into comparable scales, they reflected positive views for both users and nonusers. For example, Russell and Spooner [34] found that neither users nor nonusers felt patient privacy was harder to ensure with an EMR but nonusers showed more concern about disadvantages regarding privacy. Gans et al [25] ranked the top barriers to implementing a system, and security and privacy were of least concern.

Discussion

The 19 papers reviewed provided valuable insights into the state of evaluation of perceived EMR impacts through survey research methods. It is clear that evaluation from the user's perspective is needed alongside objective measures of impact.

Areas (Most and Least) Addressed in Survey-Based Papers

The first aim of this review was to determine which areas of EMR implementation in office practices have been addressed in survey-based research studies. The majority of background areas corresponded to the meso level, and other areas looking at aspects of implementation corresponded to the macro level. A possible explanation for the lack of metrics for individuals and groups and for roles and responsibilities is that these can be considered the basis for selecting respondents and would therefore not have specific metrics related to them. In most cases, the researchers predetermined that respondents would be physicians who may be users of the EMR. Added values, legislative acts, political trends, and economic trends are all in line with the macro level according to the Clinical Adoption Framework and certainly do affect EMR use but didn't seem to be the main objective for surveys evaluating more localized views of implementation in the office. It may be no surprise that the expense of implementation is a major consideration for office practices, and so this was a common area addressed.

The impact-specific areas we focused on were mostly contained within the net benefits dimension at the micro level of the framework. While functionality was frequently addressed in the surveys, the questions seemed to mainly pertain to availability of features rather than impact. Future surveys may wish to ask not only what is available and desired but also how it had an impact on practice. For example, this could drill further down into the efficiency areas in that improvements in efficiency may be associated with specific functionality, such as electronic transfer of laboratory results into the record, which may eliminate paper or fax transmission and manual entry time. One might expect user satisfaction to be the most-addressed category from the Clinical Adoption Framework, as surveys do generally obtain views on satisfaction. However, to understand the specific areas of satisfaction, we teased out the aspects of user satisfaction into more specific categories so that this particular category only included overall satisfaction. For example, user satisfaction with respect to the system's effect on their efficiency was mapped to productivity. Use behavior or pattern and intention to use may be encompassed by functions used and use status. Appropriateness or effectiveness and efficiency seemed to be well addressed through surveys, but there were areas of net benefits that would be expected to have had more metrics than were found. The reasons for the lack of addressed areas found for patient and caregiver participation aren't apparent. Either this specific aspect of EMR use hasn't been studied in depth or perhaps there is a degree of overlap between this category and others such as care coordination. This particular category may warrant further exploration.

Perceived Views Among Users and Nonusers for Most-Addressed Areas

For the second contribution of this review, to compare the perceived impacts between users and nonusers, we looked at the 7 most-addressed areas of impact in the set of papers: security and privacy, quality of patient care or clinical outcomes, patient-physician relationship and communication, communication with other providers, accessibility of records and information, business or practice efficiency, and costs or savings. For these areas, the views of users were generally more positive than those of nonusers, but even when looking at the two groups separately, we found mostly positive views for most of the impact-specific areas. In reviewing computer-based patient record systems (including EMRs), Delpierre et al [40] also concluded that user satisfaction was mainly positive. We found only positive views for impact on communication with other providers and accessibility of information. The other metrics had mostly positive results with some exceptions as noted. As mentioned in the Methods section, this review focused on the survey-based studies, while another systematic review was completed on the set of analytic and descriptive studies [17]. Interestingly, that review found that only 23 of 44 studies showed overall positive impacts. While a different set of areas were examined in that paper, this suggests a gap between the perceived impacts of users and nonusers reviewed here and the actual impacts seen. For example, in terms of impact on patient-physician relationship and communication, we found that all but 1 paper indicated positive views. The systematic review [17] found that it was one area that was least improved. These findings are consistent with other reviews that have found mixed results as well. Delpierre et al [40] found that in some papers, computer-based patient record systems were perceived as a physical barrier that could have a negative impact on the patient-physician relationship. And in their review of the barriers to acceptance of EMRs, Boonstra and Broekhuis [41] found that the traditional doctor-patient relationship will be changed by EMRs, but whether it is a problem is not clear. Shachak and Reis [42] conducted a narrative literature review specifically looking at factors affecting the impact of EMRs on patient-doctor communication and concluded that EMRs have a positive influence on the information-sharing aspect of communication but negative impacts on patient centeredness, emotional and psychological communication, and establishing rapport between physicians and patients.

Privacy and security was an inconclusive area, which returned the most mixed results for users and nonusers and may therefore warrant further exploration. According to Boonstra and Broekhuis [41], many researchers agree that EMRs may have a negative effect on patients' privacy and is a concern among both users and nonusers. The authors commented on a lack of clear security standards. To address security and privacy concerns, Loomis et al [29] suggested making systems compliant with the US Health Insurance Portability and Accountability Act and educating physicians about the security and confidentiality risks of paper records, as well as the safeguards built into EMRs.

Given that there were some noted differences in perceptions between users and nonuser with respect to most of the areas we looked at, it may be prudent to look at some of the associated background and other factors we identified in our mapping, which may account for some of the differences. Delpierre et al [40] noted in their review that system and user characteristics influenced user perceptions of the system and, in another review specifically on EMRs, the authors stated the characteristics of a practice can affect the extent of certain barriers to use [41]. A few papers in our review commented on differences seen in relation to practice size, but the results were somewhat mixed. MacGregor et al [30] saw some differences in perceived benefits related to practice size, but Gans et al [25] observed no consistent differences in benefits experienced by size of practice for practices that have electronic records. Boonstra and Broekhuis [41] discovered differences in their review and commented that further study is needed to analyze the reasons for the gap.

An interesting point noted in some of the papers we reviewed mirrors the classic chicken-and-egg puzzle. That is, did the more-positive views seen in users exist before they adopted the systems or did they develop them as a result of adoption? El-Kareh et al [24] commented that for cross-sectional studies this is not clear, and both Simon et al [35] and Russell and Spooner [34] mention it as a limitation to their studies. Russell and Spooner [34] explain that "if adopters were inclined toward EMRs to begin with then they haven't changed attitudes because of using the EMRs but if adopters were no different from non-adopters to begin with then they have developed positive attitudes because of use." Either way, the results support positive views among users, but this may suggest a need for more longitudinal evaluations such as that by El-Kareh et al [24].

State of Survey-Based Research for HIS Evaluation and Quality Issues

Lastly, in assessing the state of knowledge regarding survey-based research in HIS evaluation, there appears to be a lack of clear methodological guidance. Regarding design quality, the papers varied in terms of methods used and how they were reported. The items with the highest-quality scores were reporting a profile of the sample frame, a response rate, and a profile of respondents. The item that scored most poorly across all papers was performing a reliability or validity analysis of item measurement or adopting a validated instrument. Only 1 paper specifically reported having a test–retest reliability rate for each item. The other item that generally scored low was the use of a combination of personal, telephone, and mail data collection. In most cases, survey data were collected solely through a mailed questionnaire.

In terms of constructs for evaluation, we identified an issue related to neutrality for our review. We aimed to determine whether there was an overall positive or negative perceived impact with respect to each selected area but found that, in many cases, the individual survey items seemed to lean in one direction or the other—for example, a survey item asking respondents whether security is a barrier versus an item asking whether the respondent sees a benefit with regard to security. Both items address the construct of security, but the responses elicited by each may be affected by how they are posed. Therefore, it is possible that the constructs for evaluation used in each study may have had an impact on the negative and positive responses collected, which in turn affects how the results are reconciled across papers. In designing surveys, Trochim [43] asks the reader to consider what assumptions the question makes and whether the wording is loaded or slanted. In this case, where respondents are being asked for their opinion regarding an evaluation item, it may be better to ask the question in a more neutral manner, such as "What do you think about the system in regard to security?" Respondents can then rank their opinion on a Likert scale ranging from negative to positive. Or, for dichotomous data, the question can be posed as "What effect do you think the system would have on security?" The choices can be positive or negative so that the overall percentage of respondents indicating each type can be calculated and reported. An example of a paper in our review that did present neutral results well was that by Simon et al [35]. Table 2 in that paper provides percentages of respondents indicating a positive, negative, or no perceived effect of computers according to 8 dimensions. The results are presented for electronic record adopters (broken down further by high- and low-use adopters) and nonadopters. *P* values are included in the table to show significant differences between adopters and nonadopters. Although the data are presented as percentages, sample sizes are provided for each group so that proportions can be estimated.

Limitations

We experienced several challenges common to meta-analyses of survey-based research. At the paper selection stage, a major challenge was determining whether the paper discussed EMR use, which relied on descriptions of functionality provided in the papers. This review is based on a small set of 19 papers, and we had to further narrow down the set of papers for the meta-analysis based on the type of data reported. The biggest limitation in our review is related to the heterogeneity of all the papers included. Both Rao et al [14] and Jamal et al [15] describe this challenge as well in their respective reviews. In mapping the metrics and areas, we based categorizations on what was reported, which was sometimes nothing more than a simple term. Several assumptions were made for calculations as noted, and perhaps a key consideration to note is that the estimates produced don't necessarily reflect statistically significant results. They are best estimates based on the reported data, intended to help produce indications rather than absolute measures. As mentioned above regarding constructs for evaluation, we had to make assumptions about one-sided

(negative or positive) results reported in order to be able to compare and synthesize them. For example, a nonnegative result was assumed to be positive. Sometimes there were several metrics for an area, but they all came from the same paper. We relied solely on the information reported in the papers and didn't seek out all survey tools to analyze at the individual question level because one of the goals was to determine what the focus of prior survey research has been through what has been reported. Finally, there is the possibility of publication bias in our source papers, which may have contributed to the number of positive results seen. Despite these many challenges, we were still able to devise and follow a systematic, rigorous approach to synthesize the data from our set of survey-based papers to produce some interesting results.

Conclusion

Although based on a small set of papers and estimated calculations, the results of this review are promising in terms of clinicians' views for adoption of systems and suggest that clinicians are beginning to see benefits in certain areas. However, there are additional factors (eg, organizational and system) that influence perceptions, so it is important to consider a wide range of contextual factors when adopting an EMR. Our mapping identified areas corresponding to categories of the Clinical Adoption Framework that have been addressed most and other areas that haven't been looked at yet for evaluation through surveys. Although practices with electronic record systems already in use may have more positive views of impact, our review found that those without the systems still generally have a perceived positive impact with respect to some key areas, with the exception of mixed views toward privacy and security. The findings of this review have the potential to highlight areas of concern or benefit for adoption and should be considered in future implementations and evaluations. One hope is that nonusers can look to the areas where users saw more positive perceptions as areas where they can expect to see potential benefits for adoption. As well, associations between the most-addressed areas and least-addressed areas may help practices determine where they can focus effort during implementation planning, taking into account some of the key background and other areas we've identified.

Survey-based research studies are a valuable way to collect users' views for HIS evaluation. They offer data and findings that can make a significant contribution to the field. However, careful effort should be made to ensure methodological rigor and consider potential future syntheses. Our review demonstrated an approach for reconciling results presented in different ways across heterogeneous survey-based studies, which is a recognized challenge. In terms of design quality, researchers should ensure that important survey-based research study design elements are present and clearly reported using a guide such as the survey methodological attributes. As well, the constructs for evaluation should draw from an established framework or tool when possible and be expressed in a neutral manner to elicit peoples' views.

Acknowledgments

We wish to thank Heidi Bell, Morgan Price, Jeanette Boyd, and Colin Partridge for their earlier work and contributions to initial paper selection. Funding support for this review was provided by the Canadian Institutes for Health Research and Canada Health Infoway through the eHealth Chair Award. Funding sources had no role in the conduct of this review.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Paper selection flow diagram.

[JPG File, 445KB - ijmr_v1i4e3_app1.jpg]

Multimedia Appendix 2

The Clinical Adoption Framework.

[JPG File, 1MB - ijmr_v1i4e3_app2.jpg]

Multimedia Appendix 3

Estimated odds and mean calculations for selected impact-specific areas.

[XLS File (Microsoft Excel File), 88KB - ijmr_v1i4e3_app3.xls]

References

1. Canada Health Infoway. 2012. EMR and Integration URL: https://www.infoway-inforoute.ca/index.php/programs-services/investment-programs/emr-and-integration [accessed 2012-06-28] [WebCite Cache ID 68lfij121]

2. Institute of Medicine. 2001 Mar. Crossing the Quality Chasm: A New Health System for the 21st Century Report Brief URL: http://www.iom.edu/~/media/Files/Report%20Files/2001/Crossing-the-Quality-Chasm/Quality%20Chasm%202001%20%20report%20brief.pdf [accessed 2012-07-23] [WebCite Cache ID 69NQtvriZ]
3. Ford EW, Menachemi N, Phillips MT. Predicting the adoption of electronic health records by physicians: when will health care be paperless? J Am Med Inform Assoc 2006;13(1):106-112 [FREE Full text] [doi: 10.1197/jamia.M1913] [Medline: 16221936]
4. Versel N. Stalking the elusive 'killer app'. Modern Physician 2002;6(5):21.
5. National Physician Survey/Sondage national des médecins. 2012. 2007 National Physician Survey URL: http://www.nationalphysiciansurvey.ca/nps/2007_Survey/2007nps-e.asp [accessed 2012-06-28] [WebCite Cache ID 68lkwOtEL]
6. National Physician Survey/Sondage national des médecins. 2012. 2010 National Physician Survey URL: http://www.nationalphysiciansurvey.ca/nps/2010_Survey/2010nps-e.asp [accessed 2012-06-28] [WebCite Cache ID 68lkzpvP1]
7. Talmon J, Ammenwerth E, Brender J, de Keizer N, Nykänen P, Rigby M. STARE-HI -statement on reporting of evaluation studies in health informatics. Yearb Med Inform 2009:23-31. [Medline: 19855867]
8. Anonymous. Current and expected use of EMRs. Can Fam Physician 2010 Jan;56(1):39 [FREE Full text] [Medline: 20090082]
9. Lau F, Price M, Keshavjee K. From benefits evaluation to clinical adoption: making sense of health information system success in Canada. Healthc Q 2011;14(1):39-45. [Medline: 21301238]
10. Aydin CE. Survey Methods for assessing social impacts of computers in healthcare organizations. In: Anderson JG, Aydin CE, editors. Anderson JG, Aydin CE, editors. Evaluating the Organizational Impact of Healthcare Information Systems. New York, NY: Springer; 2005:75-128.
11. Chen W, Hirschheim R. A paradigmatic and methodological examination of information systems research from 1991 to 2001. Inf Syst J 2004;14:197-235. [doi: 10.1111/j.1365-2575.2004.00173.x]
12. Calder J. Survey research methods. Med Educ 1998;32(6):636-652. [doi: 10.1046/j.1365-2923.1998.00227.x]
13. Pinsonneault A, Kraemer KL. Survey research methodology in management information systems: an assessment. J Manag Inf Syst 1993;10(2):75-105.
14. Rao SR, Graubard BI, Schmid CH, Morton SC, Louis TA, Zaslavsky AM, et al. Meta-analysis of survey data: application to health services research. Health Serv Outcomes Res Methodol 2008;8(2):98-114. [doi: 10.1007/s10742-008-0032-0]
15. Jamal A, McKenzie K, Clark M. The impact of health information technology on the quality of medical and health care: a systematic review. HIM J 2009;38(3):26-37. [Medline: 19875852]
16. Canadian Medical Protective Association. 2009. Electronic Records Handbook: Implementing and Using Electronic Medical Records (EMRs) and Electronic Health Records (EHRs) URL: http://www.cmpa-acpm.ca/cmpapd04/docs/submissions_papers/pdf/com_electronic_records_handbook-e.pdf [accessed 2012-07-23] [WebCite Cache ID 69NSJQJwK]
17. Lau F, Price M, Boyd J, Partridge C, Bell H, Raworth R. Impact of electronic medical record on physician practice in office settings: a systematic review. BMC Med Inform Decis Mak 2012;12:10 [FREE Full text] [doi: 10.1186/1472-6947-12-10] [Medline: 22364529]
18. Grover V, Lee CC, Durand D. Analyzing methodological rigor of MIS survey research from 1980-1989. Inform Manag 1993;24(6):305-318. [doi: 10.1016/0378-7206(93)90028-R]
19. Ju TL, Chen Y, Sun S, Wu C. Rigor in MIS survey research: in search of ideal survey methodological attributes. J Comput Inf Syst 2007;47(2):112-123.
20. Van Der Meijden MJ, Tange HJ, Troost J, Hasman A. Determinants of success of inpatient clinical information systems: a literature review. J Am Med Inform Assoc 2003;10(3):235-243 [FREE Full text] [doi: 10.1197/jamia.M1094] [Medline: 12626373]
21. Chiang MF, Boland MV, Margolis JW, Lum F, Abramoff MD, Hildebrand PL, American Academy of Ophthalmology Medical Information Technology Committee. Adoption and perceptions of electronic health record systems by ophthalmologists: an American Academy of Ophthalmology survey. Ophthalmology 2008 Sep;115(9):1591-1597. [doi: 10.1016/j.ophtha.2008.03.024] [Medline: 18486218]
22. DesRoches CM, Campbell EG, Rao SR, Donelan K, Ferris TG, Jha A, et al. Electronic health records in ambulatory care--a national survey of physicians. N Engl J Med 2008 Jul 3;359(1):50-60 [FREE Full text] [doi: 10.1056/NEJMsa0802005] [Medline: 18565855]
23. Devine EB, Patel R, Dixon DR, Sullivan SD. Assessing attitudes toward electronic prescribing adoption in primary care: a survey of prescribers and staff. Inform Prim Care 2010;18(3):177-187. [Medline: 21396241]
24. El-Kareh R, Gandhi TK, Poon EG, Newmark LP, Ungar J, Lipsitz S, et al. Trends in primary care clinician perceptions of a new electronic health record. J Gen Intern Med 2009 Apr;24(4):464-468. [doi: 10.1007/s11606-009-0906-z] [Medline: 19156468]
25. Gans D, Kralewski J, Hammons T, Dowd B. Medical groups' adoption of electronic health records and information systems. Health Aff 2005;24(5):1323-1333. [doi: 10.1377/hlthaff.24.5.1323] [Medline: 16162580]
26. Johnston JM, Leung GM, Wong JF, Ho LM, Fielding R. Physicians' attitudes towards the computerization of clinical practice in Hong Kong: a population study. Int J Med Inform 2002 Apr;65(1):41-49. [Medline: 11904247]

27. Kemper AR, Uren RL, Clark SJ. Adoption of electronic health records in primary care pediatric practices. Pediatrics 2006 Jul;118(1):e20-e24 [FREE Full text] [doi: 10.1542/peds.2005-3000] [Medline: 16818534]
28. Leung GM, Yu PL, Wong IO, Johnston JM, Tin KY. Incentives and barriers that influence clinical computerization in Hong Kong: a population-based physician survey. J Am Med Inform Assoc 2003;10(2):201-212 [FREE Full text] [Medline: 12595409]
29. Loomis GA, Ries JS, Saywell RM, Thakker NR. If electronic medical records are so great, why aren't family physicians using them? J Fam Pract 2002 Jul;51(7):636-641. [Medline: 12160503]
30. MacGregor R, Hyland P, Harvei C, Lee BC, Dalley A, Ramu S. Benefits of ICT adoption and use in regional general medical practices: a pilot study. HIM J 2006;35(3):23-35. [Medline: 18195425]
31. Mackenzie H. The introduction of the clinical module of a primary care patient management system into a family planning association. Stud Health Technol Inform 2006;122:632-635. [Medline: 17102339]
32. Magnus D, Rodgers S, Avery AJ. GPs' views on computerized drug interaction alerts: questionnaire survey. J Clin Pharm Ther 2002 Oct;27(5):377-382. [Medline: 12383140]
33. Menachemi N, Langley A, Brooks RG. The use of information technologies among rural and urban physicians in Florida. J Med Syst 2007 Dec;31(6):483-488. [Medline: 18041281]
34. Russell SC, Spooner SA. Barriers to EMR adoption in internal medicine and pediatric outpatient practices. Tenn Med 2004 Oct;97(10):457-460. [Medline: 15536973]
35. Simon SR, Kaushal R, Cleary PD, Jenter CA, Volk LA, Orav EJ, et al. Physicians and electronic health records: a statewide survey. Arch Intern Med 2007 Mar 12;167(5):507-512. [doi: 10.1001/archinte.167.5.507] [Medline: 17353500]
36. Simon SR, Kaushal R, Jenter CA, Volk LA, Burdick E, Poon EG, et al. Readiness for electronic health records: comparison of characteristics of practices in a collaborative with the remainder of Massachusetts. Inform Prim Care 2008;16(2):129-137. [Medline: 18713529]
37. Simon SR, McCarthy ML, Kaushal R, Jenter CA, Volk LA, Poon EG, et al. Electronic health records: which practices have them, and how are clinicians using them? J Eval Clin Pract 2008 Feb;14(1):43-47. [doi: 10.1111/j.1365-2753.2007.00787.x] [Medline: 18211642]
38. Singh R, Lichter MI, Danzo A, Taylor J, Rosenthal T. The adoption and use of health information technology in rural areas: results of a national survey. J Rural Health 2012 Jan;28(1):16-27. [doi: 10.1111/j.1748-0361.2011.00370.x] [Medline: 22236311]
39. Terry K. Doctors and EHRs. Med Econ 2005 Jan 21;82(2):72-84. [Medline: 15727343]
40. Delpierre C, Cuzin L, Fillaux J, Alvarez M, Massip P, Lang T. A systematic review of computer-based patient record systems and quality of care: more randomized clinical trials or a broader approach? Int J Qual Health Care 2004 Oct;16(5):407-416 [FREE Full text] [doi: 10.1093/intqhc/mzh064] [Medline: 15375102]
41. Boonstra A, Broekhuis M. Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. BMC Health Serv Res 2010;10:231 [FREE Full text] [doi: 10.1186/1472-6963-10-231] [Medline: 20691097]
42. Shachak A, Reis S. The impact of electronic medical records on patient-doctor communication during consultation: a narrative literature review. J Eval Clin Pract 2009 Aug;15(4):641-649. [doi: 10.1111/j.1365-2753.2008.01065.x] [Medline: 19522722]
43. Trochim WMK. Web Center for Social Research Methods. 2006 Oct 20. Research Methods Knowledge Base: Survey Research URL: http://www.socialresearchmethods.net/kb/survey.php [accessed 2012-06-27] [WebCite Cache ID 68ll9huIe]

Abbreviations

EMR: electronic medical record
HIS: health information system

Edited by G Eysenbach; submitted 21.03.12; peer-reviewed by N Archer; comments to author 23.05.12; revised version received 03.07.12; accepted 06.07.12; published 28.07.12.

Please cite as:
Bassi J, Lau F, Lesperance M
Perceived Impact of Electronic Medical Records in Physician Office Practices: A Review of Survey-Based Research
Interact J Med Res 2012;1(2):e3
URL: http://www.jmir.org/2012/2/e3/
doi:10.2196/ijmr.2113
PMID:23611832

Original Paper

Exploring Nurses' Intention to Use a Computerized Platform in the Resuscitation Unit: Development and Validation of a Questionnaire Based on the Theory of Planned Behavior

Christian Malo[1], MD; Xavier Neveu[2], MSc; Patrick Michel Archambault[1,3], MSc, MD, FRCPC; Marcel Émond[1,2], MSc, MD, FRCPC; Marie-Pierre Gagnon[4], MSc, PhD

[1]Département de médecine familiale et de médecine d'urgence, Faculté de Médecine, Université Laval, Québec, QC, Canada

[2]Traumatologie – Urgence – Soins intensifs, Centre de recherche FRQS du Centre hospitalier affilié universitaire de Québec, Québec, QC, Canada

[3]Centre de santé et de services sociaux Alphonse-Desjardins (CHAU de Lévis), Département d'urgence, Québec, QC, Canada

[4]Centre de recherche du Centre hospitalier universitaire de Québec (CRCHUQ), Québec, QC, Canada

Corresponding Author:
Christian Malo, MD
Département de médecine familiale et de médecine d'urgence
Faculté de Médecine
Université Laval
Pavillon Vandry
1050 avenue de la Médecine
Québec, QC, G1V 0A6
Canada
Phone: 1 418 283 2662
Fax: 1 418 525 4194
Email: christian.malo.1@ulaval.ca

Abstract

Background: In emergency department resuscitation units, writing down information related to interventions, physical examination, vital signs, investigations, and treatments ordered is a crucial task carried out by nurses. To facilitate this task, a team composed of emergency physicians, nurses, and one computer engineer created a novel electronic platform equipped with a tactile screen that allows systematic collection of critical data. This electronic platform also has medical software (ReaScribe+) that functions as an electronic medical record and a clinical decision support system.

Objective: To develop and validate a questionnaire that can help evaluate nurses' intention to use a novel computerized platform in an emergency department resuscitation unit, based on Ajzen's theory of planned behavior (TPB).

Methods: The sample for this study was composed of 87 nurses who worked in the resuscitation unit of a tertiary trauma center. We held three focus groups with nurses working in the resuscitation unit to identify the salient modal beliefs regarding their intended use of a new electronic medical charting system for the care of trauma patients. The system included a clinical decision support tool. We developed a questionnaire in which salient modal beliefs were used as items to evaluate the TPB constructs. We also added 13 questions to evaluate nurses' computer literacy. The final questionnaire was composed of 46 questions to be answered on a 7-point Likert scale. All nurses in the resuscitation unit and present during a regular work shift were individually contacted by the principal investigator or a research assistant (phase 1). A subsample of the nurses who completed the questionnaire was invited to complete it a second time 2 weeks later (phase 2).

Results: In phase 1, we received 62 of the 70 questionnaires administered (89% response rate). Of the 27 questionnaires administered in phase 2 (retest phase), 25 were completed (93% response rate). The questionnaire showed very good internal consistency, as Cronbach alpha was higher than .7 for all constructs. Temporal stability was acceptable with intraclass correlations between .41 and .66. The intention to use the electronic platform to chart the resuscitation of trauma patients was very high among the respondents. In the logistic regression model, the only construct that predicted nurses' intention to adopt the computerized platform was the professional norm (odds ratio 3.31, 95% confidence interval 1.41–7.78).

Conclusions: We developed and validated a questionnaire that can now be used in other emergency departments prior to implementation of the computerized platform. The intention to adopt was very high among the respondents, which suggests that the implementation of this innovation could be successful at our institution.

(Interact J Med Res 2012;1(2):e5) doi:10.2196/ijmr.2150

KEYWORDS

Primary care nurses; adoption of new behavior; intention; theory of planned behaviour; emergency department; trauma care; electronic health record; clinical decision support system

Introduction

In the resuscitation room, particularly for a patient with multiple injuries, the documentation of critical patient data is a challenge for nurses. Nurses are required to write down information concerning interventions, physical examinations, vital signs, investigations, and treatments that are ordered. Given the stress in the resuscitation environment, handwriting all these data may result in information loss or errors in the documentation. For example, in the case of a tertiary trauma center, Kind et al found that 20% of vital signs were absent from a patient's chart in the 15 minutes following endotracheal intubation [1]. To minimize these difficulties, a team composed of clinicians (emergency physicians and nurses) and a computer engineer created an electronic platform equipped with a tactile screen that allows for the systematic collection of critical data in the resuscitation unit. The electronic platform is an electronic medical charting system that integrates a clinical decision support tool (ReaScribe+; ReaEvolution inc, Québec, Canada) for the care of trauma patients. It allows automatic recording of vital signs, charting of administered medication, intravenous line use, and general information related to patient care. Furthermore, it informs the radiology department of investigations required immediately.

The impact of a positive or negative attitude on the part of staff in relation to the integration of new technology at work has been fully documented [2-5]. Moreover, in the last 30 years, numerous studies have evaluated nurses' attitudes toward computer use [6,7]. In a systematic review, Huryk found that several factors influence nurses' attitudes toward the implementation of new technologies in health care. Nurses between 30 and 39 years of age, those 60 years of age and over, and those with a higher level of computer experience are more likely to adopt a positive attitude toward electronic patient records. Older nurses and those with more computer and nursing experience, who hold higher staff positions, or have higher levels of education had more positive attitudes toward clinical management systems [7]. Furthermore, the provision of adequate training and support during implementation and the inclusion of interested nurses on the implementation committee are all actions that help develop favorable attitudes toward health care information technology. Stress related to the use of new technology and the general impression that a new tool is going to reduce time available for bedside care contribute to nurses' negative attitudes toward technology [8]. According to Lee [9], success with the implementation of technology is also related to the individual's decision to adopt it or not. Nurses' acceptance of computerized documentation was influenced by their perception of its advantages, whether they could see when it was being used by others, the complexity of the system, and the compatibility of its use with their values and experience [10].

Theoretical Background

Numerous theories have been proposed that may help to explain the mecanisms involved in the adoption of a new behavior. Ajzen's [11] theory of planned behavior (TPB) has been succesfully applied to study a range of behaviors in health care professionnals. In their systematic review, Godin et al demonstrated that the TPB explained 59% of the explained variance of intention [12]. According to Ajzen, the adoption of a new behavior is predicted by the person's intention to engage in that behavior. In turn, intention depends on three main determinants: attitudes, subjective norms, and perceived behavioral control. Attitudes can be represented by the sum of the advantages and disadvantages related to adoption of the new behavior. Subjective norms consist of the person's internalization of the reference groups' opinions about the realization of the behavior. Finally, perceived behavioral control relates to the person's impression that he or she has the required resources and capacities to adopt the new behavior. Some authors (eg, Gagnon et al [13]; Daneault et al [14]) include an additional dimension in Ajzen's model to adapt it to health care workers: the professional norm. This dimension is related to a person's integration of the specific normative pressures of one's professional group. Ajzen also identifies three types of beliefs that may influence behavioral determinants: behavioral, normative, and control-related beliefs. The salient modal beliefs are the most common beliefs reported in a specific group of people concerning the adoption of a certain behavior. Consequently, it is imperative to identify those salient modal beliefs in the specific population under study in order to understand the factors explaining the intention to adopt a new behavior in a particular situation.

Shoham and Gonen [6] evaluated the intention of 411 nurses to use a computer at work, based on the TPB. In their study, 72% of nurses demonstrated a positive attitude toward the use of a computer at work. They also found that the department's work environment, nurses' work experience, stress generated by the use of a new computerized tool, and self-perception of computer skills influenced nurses' intention to use a computer. This study included nurses who generally worked in a department other than emergency and evaluated nurses' intention to use a computer at work rather than their intention to use specialized software.

A study of primary care nurses' intention to use an electronic health record conducted in Quebec, Canada [15], found that the TPB explained 58% of the variance in intention. The main determinants of nurses' intention to use the electronic record were normative beliefs, attitude, and perceived behavioral control.

Our review of the literature found no other study investigating the determinants of nurses' adoption of computerized systems based on the TPB in a resuscitation situation. Also, there exist

no questionnaires that evaluate the intention of nurses working in a resuscitation unit to use a computerized platform to chart critical patient information during care.

Objectives

The aim of this study was thus to apply Ajzen's theory to develop and validate a questionnaire that evaluates nurses' intention to use an electronic medical charting system that includes a clinical decision support tool for the care of trauma patients in the resuscitation room.

Methods

Participants

All nurses working in the resuscitation unit of the emergency department of a tertiary trauma center in Quebec City, Canada, were identified by the chief nurse. The sample of this study comprised 87 nurses working in the resuscitation unit. We excluded nurses who were absent from their usual working shifts (eg, on vacation or absent for medical problems) during the recruitment period.

Focus Groups

We held three focus groups to identify the salient modal beliefs of nurses working in the resuscitation unit in relation to the use of a new computerized platform for trauma care. Overall, 12 nurses participated (4 in each group). At the beginning of each focus group, the technician who had developed the software presented the key concepts and functions of the new platform. Then, 6 open questions were asked to identify their beliefs about adopting this system: (1 and 2) What would be the advantages or disadvantages of your use of the platform for trauma patient care in the resuscitation unit?, (3 and 4) What groups could approve or disapprove of your use of the platform in the resuscitation unit?, and (5 and 6) What would you consider to be a facilitator or an obstacle to the routine use of the platform for trauma patient care in the resuscitation unit? Each focus group lasted about 60 minutes and was audio recorded. Each participant received a Can $50 compensation.

Questionnaire Design

Two investigators individually reviewed the content of each focus group with a standardized data extraction form and identified the most common themes reported by the participants. We then pooled the results and kept themes that had been reported at least 3 times (representing 25% of the participants) as the salient modal beliefs. A questionnaire (see Multimedia Appendix 1) was developed in which each of these salient modal beliefs was used as an item to evaluate the TPB constructs. We included 30 items in total: 8 to evaluate the advantages and disadvantages related to platform use (attitude), 6 to evaluate the people or groups that would approve or disapprove of use of the platform (social norm), 10 to evaluate the factors that could facilitate or limit platform use (perceived control), 3 to evaluate the intention to use the platform, and 3 to evaluate the professional norm. Each of these questions (question numbers 4 to 33 of the questionnaire) were assessed using a 7-point Likert scale (eg, total disagreement = 1, strong disagreement = 2, partial disagreement = 3, neutral = 4, partial agreement = 5, strong agreement = 6, total agreement = 7). We also included 13 questions to evaluate the nurses' computer literacy in order to consider this variable in our analyses. These questions were translated and adapted (with the authors' permission) from a validated questionnaire developed by Gassert and McDowell [16]. We used the same answer choices as those that were initially developed by Gasser and McDowell. We also included 3 sociodemographic questions in the questionnaire. The final questionnaire totaled 46 closed questions. The questionnaire was face validated by 3 research nurses, and we used their comments to improve the final version of the questionnaire (see Multimedia Appendix 1).

A cover letter was attached to the questionnaire to present the study's objectives and a brief overview of the key characteristics of the platform. A notification at the bottom of the first page clearly indicated that returning the completed questionnaire signified consent to participate in the study. This study was approved by the hospital's ethics review board.

Data Collection

Participants were recruited in the emergency department between April 25 and May 6, 2011. All the nurses working in the resuscitation unit and present during their regular work shifts were individually contacted by the principal investigator or a research assistant. Each questionnaire was identified by a unique identification number. Participants were instructed to return their completed questionnaire to an identified secure box. Participants who accepted to complete the questionnaire were invited to complete it a second time 2 weeks later to measure the questionnaire's reliability. Participants received a Can $5 gift card for each questionnaire completed (one for phase 1 and one for phase 2).

Statistical Analyses

First, we assessed the temporal stability of the measurement of theoretical constructs in the questionnaire by the intraclass correlation coefficient (ICC) for answers to the first and second questionnaires (test-retest reliability) [17-19]. We considered test-retest reliability to be fair for values of ICC between .4 and .59, good for values between .60 and .74, and excellent for values higher than .75 [20-22]. Then, we calculated Cronbach alpha coefficients to assess the internal consistency of the theoretical constructs [17-19]. We explored construct validity by examining Spearman correlations between items and theoretical constructs. Correlations between items forming the same construct were also explored to detect collinearity problems.

Second, we examined descriptive analyses of sociodemographic characteristics and theoretical variables. Associations between these variables were explored with chi-square tests for the univariate analysis, and a logistic regression was performed to identify the determinants of nurses' high versus moderate intention to use the platform. Using the median intention score as the cut-off point, we defined high intention as a score of >6 (total agreement) on the Likert scale and moderate intention as a score ≤6. All analyses were done with SAS version 9.2 (SAS Institute, Cary, NC, USA).

Results

In phase 1, we received 62 of the 70 questionnaires administered (89% response rate). Of the 27 questionnaires administered in the second phase (2 weeks later), 25 were completed (93% response rate). We chose to administer 27 questionnaires in the second phase to obtain results for the retest phase from at least 30% of our initial sample. This number corresponds to the number of participants recommended for calculating temporal stability in a psychosocial questionnaire based on the TPB [23]. Descriptive statistics of the sample are presented in Table 1. More women than men participated in the study, which is consistent with the gender distribution in the population. Of the respondents, 66% (40/61) were 40 years of age or less and 30% (18/61) had more than 15 years of experience. As the study questionnaire was entirely anonymous, it was not possible to identify nurses who did not return their questionnaire. However, to assess the possibility of nonresponse bias, we compared respondents' characteristics with those of Quebec nurses [24]. The proportion of nurses aged less than 40 years and the proportion of men in our sample were higher than those of Quebec nurses. These differences can be explained by the fact that nurses from our sample work exclusively in the emergency department, and for numerous reasons (eg, shift work, stress), emergency department nurses tend to be younger, and there is a higher proportion of men.

Table 1. Demographic characteristics of respondents and high versus low intention to use the platform.

Characteristic of sample (n = 62)	Frequency	χ^2 (high versus low intention to use the platform)[a]	*df*	*P* value
Gender [b]		1.3	1	.26
Male	16 (26%)			
Female	45 (74%)			
Age (years) [b]		0.9	2	.63
<31	21 (34%)			
31–40	19 (31%)			
>40	21 (34%)			
Experience (years)		3.5	3	.32
<5	17 (27%)			
6–10	13 (21%)			
11–15	14 (23%)			
>15	18 (29%)			

[a] High intention: score of >6 on Likert scale; low intention: score of ≤6 on Likert scale.

[b] n = 61.

Internal Consistency

Table 2 shows the Cronbach alpha coefficients for each construct of the questionnaire. All constructs had a Cronbach alpha higher than .7, which is considered excellent [25,26].

Table 2. Internal consistency of the constructs.

Construct	Number of items	Cronbach alpha	95% CI[a]
Attitude	8	.85	.77–.90
Perceived control	10	.82	.52–.89
Professional norm	3	.84	.71–.93
Social norm	6	.85	.80–.90
Intention	3	.94	.80–.99
Computer literacy	13	.89	.85–.92

[a] Confidence interval.

Reliability

Table 3 shows ICC results for each construct. Temporal stability was shown to be good for the professional norm and social norm constructs and fair for the other constructs.

Table 3. Test–retest stability of constructs (n = 25).

Construct	ICC[a]	95% CI[b]
Attitude	.41	.03–.69
Perceived control	.41	.04–.69
Professional norm	.66	.37–.83
Social norm	.60	.28–.80
Intention	.59	.27–.80

[a] Intraclass correlation coefficient.

[b] Confidence interval.

Descriptive Statistics for the Constructs

For each construct, we evaluated the mean score of nurses' responses on the 7-point Likert scale (Table 4). Overall, the participants had a high intention to use the computerized platform when it becomes available (mean 6.35 out of a maximum of 7).

Table 4. Means and standard deviations of constructs.

	Attitude	Perceived control	Professional norm	Social norm	Intention	Computer literacy[a]
Mean	5.81	5.94	6.10	5.63	6.35	2.58
SD	0.96	0.81	1.13	0.96	1.06	1.12

[a] Score on a 4-point Likert scale.

Bivariate Correlations Between Constructs

We explored correlations between constructs as shown in Table 5. According to Kline [27], multicollinearity is present when the correlation between two independent variables is greater than .85. The coefficient of correlation between each construct was acceptable, which indicates that a multicollinearity problem was not present.

The mean score was 2.58 (out of a maximum of 4) for computer literacy. This corresponds to some experience with the various computer software packages or devices evaluated (Table 6). The internal consistency for computer literacy was excellent (Cronbach alpha = .89, 95% confidence interval [CI] .85–.92). Temporal stability was acceptable with an ICC of .68 (95% CI .40–.84). A large majority (59/62, 95%) of respondents indicated that they had their own computer at home. The nurses indicated that they had less experience with two types of software: databases (eg, Microsoft Access) and computerized statistical analysis software (eg, SAS, SPSS, Maple).

Table 5. Spearman correlation coefficients (and *P* values) between each construct pair (n = 62).

Construct	Intention	Attitude	Perceived control	Professional norm	Social norm	Computer literacy
Intention	1.00	.42 (<.001)	.35 (.01)	.62 (<.001)	.31 (.02)	.30 (.02)
Attitude		1.00	.38 (.003)	.39 (.002)	.43 (<.001)	.38 (.002)
Perceived control			1.00	.35 (.01)	.38 (.003)	.19 (.14)
Professional norm				1.00	.36 (.004)	.19 (.14)
Social norm					1.00	.26 (.04)
Computer literacy						1.00

Table 6. Assessment of computer literacy in the study sample[a].

Computer skill	Score	
	Mean	SD
Microcomputer use (PC or Mac)	2.97	0.85
Keyboard/typing skills	3.42	0.64
Word processing	2.98	0.98
Spreadsheet (eg, Microsoft Excel)	2.16	0.93
Database (eg, Microsoft Access)	1.95	0.86
Email	3.52	0.62
Internet	3.61	0.52
Bibliographic database searching (eg, CINAHL, Medline, PubMed)	2.02	0.98
Computerized statistical programs	1.37	0.61
Presentation software (eg, PowerPoint)	2.08	1.00
Personal digital assistant (eg, Palm, iPod, iPhone)	2.50	1.22
Use of cell phone with Web capability	2.42	1.14
Do you have your own personal computer at home (yes), n (%)	59/62	95%

[a] Adapted from Gassert and McDowell [16] with their permission.

Logistic Regression Model

We tested a logistic regression model with the full sample (n = 62) to identify the main determinants of nurses' low versus high intention to use the platform. The only construct that predicted the nurses' intention to adopt the computerized platform was the professional norm (odds ratio 3.31, 95% CI 1.41–7.78; see Table 7). The Nagelkerke *R2* obtained by this model is .41.

Table 7. Logistic regression model for dichotomized intention construct (high vs low intention)[a].

Construct	Estimated (β)	OR[b]	95% CI[c]	*P* value
Attitude	0.511	1.67	0.65–4.30	.29
Perceived control	0.266	1.31	0.44–3.89	.63
Professional norm	1.197	3.31	1.41–7.78	.01
Social norm	0.022	1.02	0.49–2.15	.95
Computer literacy	0.628	1.87	0.60–5.89	.28

[a] Regression model equation: logit[probability((score intention >6) = 1)] = $\beta_0 + \beta_1$.attitude + β_2.perceived control + β_3.professional norm + β_4.social norm + β_5.computer skills.

[b] Odds ratio.

[c] Confidence interval.

Discussion

To the best of our knowledge, this is the first study that applies the TPB to assess nurses' intention to use a computerized platform as a support to resuscitation work in an emergency department. We believe this computerized platform is an important innovation because it is an electronic medical record and a clinical decision support system created for nurses working in the resuscitation unit. We developed a questionnaire based on the TPB that showed good internal consistency and fair temporal stability. Furthermore, we translated into French and validated a 13-question instrument developed by Gassert and McDowell [16] to evaluate nurses' computer literacy. Because the response rate was high (89%), it is unlikely our study was subject to a self-selection bias.

Of particular interest is that only the professional norm significantly influenced the intention to use the platform in our logistic regression model. This can be partially explained by the fact that in the year preceding the present study, the emergency department completed a health services quality accreditation process, conducted by Accreditation Canada. The accreditation of the emergency department was based on five key elements. The maintenance of accessible and efficient clinical information systems, and the creation of prepared and proactive teams were two of these five principles. Participation in this accreditation process may have sensitized nursing personnel to current professional norms.

From the literature, it could be expected that nurses with higher computer literacy would be more likely to have a higher intention to use the platform [7]. Nurses in our study evaluated

themselves as having some to moderate computer skills (58/62, 94%) with a mean score of 2 or 3), and 95% (59/62) claimed to have a personal computer at home. However, in our study we only found a trend toward an association (odds ratio 1.87, 95% CI 0.60-5.89). Our regression model indicated that the TPB variables explained 41% of the variance in nurses' intention to use the computerized platform, which is comparable with what is reported in the TPB literature [12].

Age, gender, and number of years of experience did not have any influence on nurses' intention to use the platform. These results are different from those reported in previous studies [7,28,29]. Previous studies about nurses' adoption of computerized information systems have shown that perceived behavioral control, normative beliefs, and attitudes influenced their intention to adopt an electronic health record [15]. Perceived behavioral control and attitude were also related to nurses' intention to use a computer in a hospital setting in the study by Shoham and Gonen [6]. The lack of impact of these determinants in our study may be related to the small number of participants and the fact that intention was very high among them. Thus, if most nursing staff already support the adoption of the platform, the intention of nurses is less likely to be influenced by social pressures. The same explanation can be provided for the attitudinal beliefs: everyone shares the same beliefs about the advantages of using the computerized information system in the resuscitation unit, so attitude is not likely to have an impact on nurses' intention. The lack of influence of the behavioral control beliefs could indicate that the perceived facilitators for use of the platform were very similar among the respondents. The fact that this platform was developed in the same resuscitation unit by clinicians who already work with them may have led nurses to perceive few barriers to their adoption of the platform. Furthermore, it is possible that they felt some kind of peer pressure to answer positively to the questionnaire even though the questionnaires were answered anonymously. Indeed, it could be useful to administer this questionnaire in other centers not involved in the development of this electronic platform.

As participants during phase 1 were aware that there would be a second phase to evaluate temporal stability, it is possible that the participants tried to remember their answers in phase 1 in order to write the same answers in phase 2. We believe it is unlikely that this bias influenced the results because phase 1 and 2 were separated by 2 weeks and because the questionnaire included 46 questions.

Some limitations should be noted in relation to our results. First, our population was small, which explains our small sample size, even though we had a good response rate. Furthermore, our study was conducted in only one emergency department. Generalization of the results to other populations and settings is not possible, but the instrument that was developed can be applied in other similar settings. Second, clinicians who worked in the same emergency department participated in the development of the software. This could have influenced the results by introducing a social desirability bias. Third, nurses' intention to use the computerized platform was high (mean of 6.35 out of 7), thus leading to very low variance in the intention scores, which could have limited the performance of the theoretical model. Fourth, for the perceived control construct, 2 items of the questionnaire did not correlate well with the other items of the construct. These 2 items were (1) the fact that some people are resistant to change, and 2) the eventual presence of computer bugs. These were the only two obstacles we identified in the focus groups. It was not possible to create a new construct with those 2 items because it would have been in contradiction to experts' recommendations in relation to measurement of psychosocial constructs [23,30]. We decided to keep these items in the questionnaire because we felt they may have an impact on nurses' intention to use the platform, but the results indicated that it this was not the case. Finally, because our questionnaire contained 46 questions, our results may have been influenced by respondent fatigue (the weariness effect).

Conclusion

This study evaluated nurses' intention to use a computerized platform in the resuscitation unit of an emergency department. The respondents' intention was very high, which suggests that the implementation of this innovation is likely to be successful in our emergency department. Once it is implemented, it will be interesting to assess whether this high intention has translated into a high adoption rate of the computerized platform. The questionnaire developed proved to have good internal consistency and good temporal stability. The platform developers hope to implement this computerized platform in other emergency departments in Quebec. The validated questionnaire could be used in other emergency departments prior to implementation of the computerized platform to orient the actions of managers toward improving the likelihood of a successful implementation. Evaluation of this questionnaire in other trauma centers with different settings could help to assess its validity and could help to identify the potential barriers and facilitators to implementation of this novel charting and decision support platform in other trauma centers in the province.

Acknowledgments

We would like to thank Ms Carole A Gassert and Ms Dorothea E MacDowell for allowing us to use their questionnaire to evaluate nurses' computer literacy. We would also like to thank Mr Duncan Sanderson for his help with the review of this article. Finally, we thank nurses who agreed to participate in the present study.

Funding: We received funding from the Fonds de recherche en santé du Québec (FRQ-S #22323). The funders did not have any role in the review or approval of the manuscript for publication.

Conflicts of Interest

One of the authors (ME) is the president of ReaEvolution, the company that created and sells ReaScribe+. The other authors declare no conflicts of interests.

Multimedia Appendix 1

The Questionnaire developed to evaluate nurses' intention to use a computerized platform in the resuscitation unit.

[PDF File (Adobe PDF File), 71KB - ijmr_v1i5e5_app1.pdf]

References

1. Kind S, Émond M. Evaluation of emergency physicians (EP) intubation practice before and after the introduction of videolaryngoscopy (VDL). CJEM 2009;11(3):287.
2. Dowling AF. Do hospital staff interfere with computer system implementation? Health Care Manage Rev 1980;5(4):23-32. [Medline: 7000734]
3. Eaton N. Computer use in a school of nursing. Nurs Stand 1990;5(7):35-37. [Medline: 2123682]
4. Shumway JM, Jacknowitz AI, Abate MA. Analysis of physicians', pharmacists', and nurses' attitudes toward the use of computers to access drug information. Methods Inf Med 1990 Mar;29(2):99-103. [Medline: 2342434]
5. Burkes M. Identifying and relating nurses' attitudes toward computer use. Comput Nurs 1991;9(5):190-201. [Medline: 1933661]
6. Shoham S, Gonen A. Intentions of hospital nurses to work with computers: based on the theory of planned behavior. Comput Inform Nurs 2008;26(2):106-116. [doi: 10.1097/01.NCN.0000304777.48155.e3] [Medline: 18317262]
7. Huryk LA. Factors influencing nurses' attitudes towards healthcare information technology. J Nurs Manag 2010 Jul;18(5):606-612. [doi: 10.1111/j.1365-2834.2010.01084.x] [Medline: 20636510]
8. Smith K, Smith V, Krugman M, Oman K. Evaluating the impact of computerized clinical documentation. Comput Inform Nurs 2005;23(3):132-138. [Medline: 15900170]
9. Lee FW. Adoption of electronic medical records as a technology innovation for ambulatory care at the Medical University of South Carolina. Top Health Inf Manage 2000 Aug;21(1):1-20. [Medline: 11010366]
10. Lee TT. Nurses' adoption of technology: application of Rogers' innovation-diffusion model. Appl Nurs Res 2004 Nov;17(4):231-238. [Medline: 15573331]
11. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991;50:179-211. [doi: 10.1016/0749-5978(91)90020-T]
12. Godin G, Bélanger-Gravel A, Eccles M, Grimshaw J. Healthcare professionals' intentions and behaviours: a systematic review of studies based on social cognitive theories. Implement Sci 2008;3:36 [FREE Full text] [doi: 10.1186/1748-5908-3-36] [Medline: 18631386]
13. Gagnon MP, Godin G, Gagné C, Fortin JP, Lamothe L, Reinharz D, et al. An adaptation of the theory of interpersonal behaviour to the study of telemedicine adoption by physicians. Int J Med Inform 2003 Sep;71(2-3):103-115. [Medline: 14519403]
14. Daneault S, Beaudry M, Godin G. Psychosocial determinants of the intention of nurses and dietitians to recommend breastfeeding. Can J Public Health 2004;95(2):151-154. [Medline: 15074909]
15. Leblanc G, Gagnon MP, Sanderson D. Determinants of Primary Care Nurses' Intention to Adopt an Electronic Health Record in Their Clinical Practice. Comput Inform Nurs 2012 May 15. [doi: 10.1097/NXN.0b013e318257db17] [Medline: 22592453]
16. Gassert CA, McDowell D. Evaluating graduate and undergraduate nursing students' computer skills to determine the need to continue teaching computer literacy. Medinfo 1995;8 Pt 2:1370. [Medline: 8591450]
17. Nunnaly JC, Bernstein IH. Psychometric Theory. 3rd edition. McGraw-Hill Series in Psychology. New York, NY: McGraw-Hill; 1994.
18. Efron B. Better bootstrap confidence intervals. J Am Stat Assoc 1987;82:171-200.
19. Efron B, Tibshirani RJ. An Introduction to the Bootstrap. New York, NY: Chapman & Hall; 1993.
20. Fermanian J. [Measuring agreement between 2 observers: a quantitative case]. Rev Epidemiol Sante Publique 1984;32(6):408-413. [Medline: 6531504]
21. Cicchetti D. Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. Psychol Assess 1994;6(4):284-290. [doi: 10.1037/1040-3590.6.4.284]
22. Rosner B. Fundamentals of Biostatistics. 6th edition. Belmont, CA: Thomson/Brooks/Cole; 2006:615-616.
23. Gagné C, Godin G. Les théories sociales cognitives: Guide pour la mesure des variables et le développement de questionnaire. Ste-Foy, QC: Groupe de recherche sur les aspects psychosociaux de la santé. École des sciences infirmières, Université Laval; 1999.

24. Rapport statistique sur l'effectif infirmier 2010-2011: Le Québec et ses régions. Ordre des infirmières et des infirmiers du Québec. 2012 Apr 30 URL: http://www.oiiq.org/sites/default/files/uploads/pdf/l_ordre/qui_sommes_nous/statistiques/292S-Rapport_statistique_2010-2011.pdf [accessed 2012-09-04] [WebCite Cache ID 6ARF2cIgQ]
25. DeVon HA, Block ME, Moyle-Wright P, Ernst DM, Hayden SJ, Lazzara DJ, et al. A psychometric toolbox for testing validity and reliability. J Nurs Scholarsh 2007;39(2):155-164. [doi: 10.1111/j.1547-5069.2007.00161.x] [Medline: 17535316]
26. Watson R, Thompson DR. Use of factor analysis in Journal of Advanced Nursing: literature review. J Adv Nurs 2006 Aug;55(3):330-341. [doi: 10.1111/j.1365-2648.2006.03915.x] [Medline: 16866827]
27. Kline RB. Principles and Practice of Structural Equation Modeling. New York, NY: Guilford Press; 1998.
28. Lorenzi NM, Riley RT, Blyth AJ, Southon G, Dixon BJ. Antecedents of the people and organizational aspects of medical informatics: review of the literature. J Am Med Inform Assoc 1997;4(2):79-93 [FREE Full text] [Medline: 9067874]
29. Getty M, Ryan AA, Ekins ML. A comparative study of the attitudes of users and non-users towards computerized care planning. J Clin Nurs 1999 Jul;8(4):431-439. [Medline: 10624260]
30. Ajzen I, Fishbein M. Understanding Attitudes and Predicting Social Behavior. Englewood Cliffs, NJ: Prentice-Hall; 1980.

Abbreviations

CI: confidence interval
ICC: intraclass correlation coefficient
TPB: theory of planned behavior

Edited by G Eysenbach; submitted 30.04.12; peer-reviewed by TT Lee, S Jetté, G Chen; comments to author 19.06.12; revised version received 28.06.12; accepted 30.07.12; published 13.09.12.

Please cite as:
Malo C, Neveu X, Archambault PM, Émond M, Gagnon MP
Exploring Nurses' Intention to Use a Computerized Platform in the Resuscitation Unit: Development and Validation of a Questionnaire Based on the Theory of Planned Behavior
Interact J Med Res 2012;1(2):e5
URL: http://www.jmir.org/2012/2/e5/
doi:10.2196/ijmr.2150
PMID:23611903

Original Paper

Using Machine-Learned Bayesian Belief Networks to Predict Perioperative Risk of Clostridium Difficile Infection Following Colon Surgery

Scott Steele[1], MD; Anton Bilchik[2], MD; John Eberhardt[3]; Philip Kalina[3]; Aviram Nissan[4], MD; Eric Johnson[1], MD; Itzhak Avital[5], MD; Alexander Stojadinovic[6], MD

[1]Madigan Army Medical Center, Department of Surgery, Tacoma, WA, United States
[2]California Oncology Research Institute (CORI), Los Angeles, CA, United States
[3]DecisionQ Corporation, Washington, DC, United States
[4]Department of Surgery, Rabin Medical Center, Petah Tikvah, Israel
[5]Surgery Branch, National Institutes of Health, Bethesda, DC, United States
[6]Department of Surgery, Division of Surgical Oncology, Walter Reed National Military Medical Center, Bethesda, DC, United States

Corresponding Author:
Scott Steele, MD
Madigan Army Medical Center
Department of Surgery
Bldg 9040 Jackson Avenue
Tacoma, WA, 98431
United States
Phone: 1 (253) 691 7976
Fax: 1 (253) 968 2200
Email: harkersteele@mac.com

Abstract

Background: Clostridium difficile (C-Diff) infection following colorectal resection is an increasing source of morbidity and mortality.

Objective: We sought to determine if machine-learned Bayesian belief networks (ml-BBNs) could preoperatively provide clinicians with postoperative estimates of C-Diff risk.

Methods: We performed a retrospective modeling of the Nationwide Inpatient Sample (NIS) national registry dataset with independent set validation. The NIS registries for 2005 and 2006 were used for initial model training, and the data from 2007 were used for testing and validation. International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM) codes were used to identify subjects undergoing colon resection and postoperative C-Diff development. The ml-BBNs were trained using a stepwise process. Receiver operating characteristic (ROC) curve analysis was conducted and area under the curve (AUC), positive predictive value (PPV), and negative predictive value (NPV) were calculated.

Results: From over 24 million admissions, 170,363 undergoing colon resection met the inclusion criteria. Overall, 1.7% developed postoperative C-Diff. Using the ml-BBN to estimate C-Diff risk, model AUC is 0.75. Using only known a priori features, AUC is 0.74. The model has two configurations: a high sensitivity and a high specificity configuration. Sensitivity, specificity, PPV, and NPV are 81.0%, 50.1%, 2.6%, and 99.4% for high sensitivity and 55.4%, 81.3%, 3.5%, and 99.1% for high specificity. C-Diff has 4 first-degree associates that influence the probability of C-Diff development: weight loss, tumor metastases, inflammation/infections, and disease severity.

Conclusions: Machine-learned BBNs can produce robust estimates of postoperative C-Diff infection, allowing clinicians to identify high-risk patients and potentially implement measures to reduce its incidence or morbidity.

(Interact J Med Res 2012;1(2):e6) doi:10.2196/ijmr.2131

KEYWORDS

Clostridium difficile; Bayesian belief network; pseudomembranous colitis; colectomy; NIS

Introduction

Clostridium difficile (C-Diff) infection has continued to be associated with a steady rise in incidence, increasing over 200% in the United States alone from 2000 to 2005 [1,2]. Stool cultures have demonstrated the underlying gram-positive rod bacterium in approximately 3% of healthy adults, whereas incidence rates are as high as 16% to 35% in hospitalized patients [3,4]. These rates are even higher following prolonged exposure to antibiotics and in patients with underlying cancer or immunosuppression [4]. With this increase, there has been a concomitant rise in one particularly virulent C-Diff strain, 027/B1/North American pulsed-field type 1 (NAP1), that is associated with increased spore formation, higher resistance to fluoroquinolones, up to 23-fold increase in toxin production, and overall worse outcomes [5-7]. This emerging epidemic has not been isolated to the United States, with Canadian reports showing an increase from 0.7 cases per 1000 in 1999-2002 to 14.9 cases per 1000 in 2003-2005 [8]. Additional reports of polymerase chain reaction (PCR) ribotype 027 strains of C-Diff outbreaks across North America and Europe highlight the need for increased vigilance and risk-reducing interventions to prevent its onset [9,10].

Multiple factors contribute to the risk of developing C-Diff colitis. Some of these are well known and easy to monitor, such as antibiotic use and bowel preparations that alter the normal gastrointestinal flora (although controversial) [11-13]. Others are more difficult to pinpoint. The secondary development following elective colonic resection has been shown to be associated with an increased length of stay, higher complication rates, and a nearly 4-fold increase in mortality [14].

Ideally, recognition of patients early in the course of the disease, even with limited data, would allow physicians to initiate treatment in a timely fashion and reduce the likelihood of poor outcomes. A persistent challenge in the treatment of patients who develop C-Diff postoperatively is the absence of a prognostic tool to identify patients who are at high risk of failing standard medical therapy. Identification of patients at an increased risk for C-Diff colitis prior to surgery and implementation of prophylactic strategies could potentially prevent this significant secondary infection altogether. Clinical decision-support systems (CDSSs) have fulfilled an important unmet need to allow for more accurate estimates and predictions where multiple different variables influence disease patterns. CDSSs typically are comprised of a knowledge base that interprets patient-specific information along with a user interface that enables clinicians to interact with the system. The concept is to use specific patient information to make individualized decisions about that patient's care based on thousands of prior similar scenarios. In other words, "to get the right information needed, to make the right decision, for the right patient, at the right time" [15]. Along with other advances in technology, these have become an essential component of clinical practice in multiple disease processes [16,17].

One such CDSS employs machine-learned Bayesian belief networks (ml-BBNs). These are directed acyclic graphs of conditional probabilities that allow users to understand how different features are conditionally independent of each other. In this study, our objective was to determine if ml-BBNs could preoperatively identify predisposing factors and provide actionable postoperative estimates of C-Diff colitis development following colectomy.

Methods

Data Selection and Curation

Data for this study came from the Nationwide Inpatient Sample (NIS), an administrative database provided by the US Department of Health and Human Services and a product of the Healthcare Cost and Utilization Project sponsored by the Agency for Healthcare Research and Quality (AHRQ). This study was performed in accordance with the NIS data user agreement and approval was obtained through a local institutional review board. The NIS is the largest inpatient, all-payer database in the United States, accounting for approximately 8 million hospital admissions each year. It contains information on patient demographics, comorbidities, admission and discharge diagnoses, and multiple outcome measures totaling 220 distinct variables per hospitalization in 2007 alone. Among the data fields are 15 slots for the *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM) diagnosis codes and 15 slots for ICD-9-CM procedure codes. By utilizing a stratified sampling frame and discharge weights, NIS is able to create accurate national estimates from a 20% sample of all nationwide discharges. The states excluded each year per group were different from year to year. The NIS also contains multiple validated severity adjustment measures to estimate patient disease severity used for clinical comparisons. The NIS is described in detail at http://hcup-us.ahrq.gov/nisoverview.jsp (archived at http://www.webcitation.org/6AWhEpjDg).

Patients included in the study were identified within the NIS dataset for the period of 2005-2007 using ICD-9-CM procedure and diagnostic codes. Initial inclusion criteria were patients who underwent colonic resection during hospital admission. All records containing any ICD-9-CM procedure codes beginning with 45.7 (partial excision of large intestine) or 45.8 (total intra-abdominal colectomy) were pulled for analysis because these codes indicate some form of colon resection. A complete list of the corresponding codes, including a summary for each, can be found in Table 1.

Table 1. Colon resection ICD-9-CM procedure codes.

Procedure code	Procedure
45.7	Open and other partial excision of large intestine
45.71	Open and other multiple segmental resection of large intestine
45.72	Open and other cecectomy, resection of cecum and terminal ileum
45.73	Open and other right hemicolectomy, ileocolectomy, right radical colectomy
45.74	Open and other resection of transverse colon
45.75	Open and other left hemicolectomy
45.76	Open and other sigmoidectomy
45.79	Other and unspecified partial excision of large intestine
45.8	Total intra-abdominal colectomy, excision of cecum, colon, and sigmoid
45.81	Laparoscopic total intra-abdominal colectomy
45.82	Open total intra-abdominal colectomy
45.83	Other and unspecified total intra-abdominal colectomy

Patients were then identified as having an infection with C-Diff during their admission by searching the NIS NDX-1 secondary diagnosis fields (DX2-DX15) for the ICD-9-CM diagnosis code 008.45 (the code for C-Diff). The primary diagnosis field (DX1) was excluded from this search in order to identify only those hospitalizations in which C-Diff colitis developed following colon resection versus undergoing a colectomy for primary C-Diff colitis [14].

Definition of Variables

Demographic variables examined included age (years), gender, race, expected payer (ie, Medicare, Medicaid, private insurance, self-pay, or other), type of resection (see Table 1), and median income in the patient's ZIP code. We also included information on the hospital, such as bed size, control/ownership, region, and teaching status.

The AHRQ comorbidity software, provided by the NIS, was used to examine pre-existing medical conditions. This software assigns variables to identify comorbidities from hospital discharge records using ICD-9-CM diagnosis codes. Comorbidity variables included in our analysis were congestive heart failure (CM_CHF), diabetes (CM_DM), hypertension (CM_HTN), chronic pulmonary disease (CM_CHRNLUNG), renal failure (CM_RENFAIL), peripheral vascular disease (CM_PERIVASC), obesity (CM_OBESE), and malnutrition (CM_WGHTLOSS).

Patient disease severity was accounted for using two validated variables contained within the NIS (provided by the Medstat Disease Staging software, version 5.21): (1) disease staging: principal stage (DS Stage); and (2) disease staging: mortality scale (DS Mtr S). Both variables use several patient-specific parameters present at time of admission to provide a measure of severity for clinical comparison. We used principal stage in our model, which is an assigned numerical value reflective of the level of severity for the principal admitting diagnosis only. For further characterization, we recoded the NIS disease stage variable into 3 basic levels: (1) disease with no complications, (2) disease with local complications, and (3) disease involving multiple sites or systemic complications.

"InflamAndOtherInfection"

Due to the relationship between antibiotic use and the development of C-Diff colitis, this risk factor was critical for the data analysis. Although the NIS database includes a rich array of information, it does not explicitly identify which antibiotics were administered during the patient's hospitalization.

To identify risk factors for inflammation and infection, we reviewed the entire list of multilevel clinical classifications software (CCS) categories [18] and used consensus opinion to determine which variables to associate with this category. This culminated in the following infection groupings (and codes): tuberculosis (Tuberculosis), streptococcal septicemia (StreptococcalSepticemia), staphylococcal septicemia (StaphylococcalSepticemia), *Escherichia coli* septicemia (EColiSepticemia), other gram-negative septicemia (OtherGramNegSepticemia), other specified septicemia (OtherSpecSepticemia), unspecified septicemia (UnspecSepticemia), sexually transmitted infection not human immunodeficiency virus or hepatitis (SexTransInfectNotHIVorHep), other bacterial infection (BacterialInfectionOther), any bacterial infection (BacterialInfectionAny), and inflammation/other infection (InflamAndOtherInfection) for CCS codes that did not fit one of the other groups.

Additional Data Curation

Using an iterative modeling process, the first round of preliminary modeling provided insights on variables and resulted in basic data recoding, such as changing occurrences of -8 and -88 (undefined variables) to nulls to standardize all missing data within these fields. Additional fields that were conditionally independent of developing C-Diff colitis were censored to simplify the structure of the network and reduce confounders in the model. This was done using structural analysis of the model, such that nodes (ie, variables) that were conditionally independent of C-Diff colitis were removed when they were on the edge of the network.

Clinically, chronic conditions are known when the patient is admitted. To reduce complexity and improve model robustness, the 4 chronic condition variables in NIS (CHRON1, CHRONB1, CHRON2, and CHRONB2*)* were recoded to consolidate their information into 2 variables (CHRONB1mod and CHRONB2mod). The new CHRONB1mod encodes the body system associated with the principal diagnosis, but only if that condition is chronic and the CHRONB2mod variable encodes the body system associated with the second diagnosis only if that condition is chronic.

Main Outcome Measures

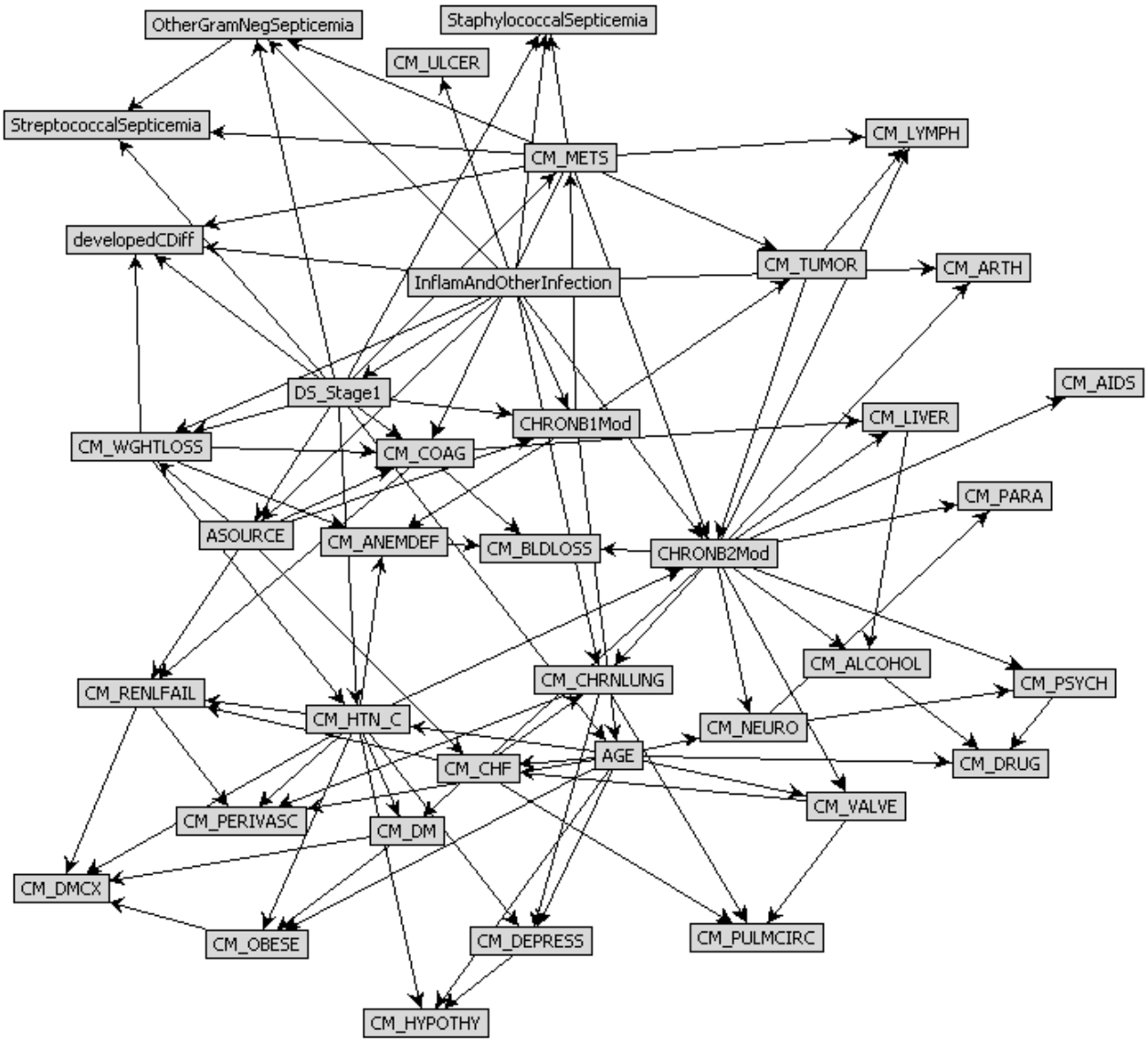
The primary variable in this study was the presence of C-Diff infection following colectomy (ICD-9-CM code 008.45). Again, the NIS diagnosis DX1 was excluded because this would likely indicate admission for primary C-Diff infection and not infection after colon surgery.

Machine-Learned Bayesian Belief Networks

Machine-learned Bayesian belief networks (ml-BBNs) were trained using commercially available machine-learning algorithms (FasterAnalytics, DecisionQ Corporation, Washington, DC) and a training dataset (NIS 2005 and 2006) to learn network structure and prior probability distributions. The FasterAnalytics software uses heuristic algorithms to allow computers to learn natively from data and discover the most likely structure of conditional dependence between variables in order to specify a BBN. The BBNs are graphs of conditional probabilities that allow users to understand how different features are conditionally independent of each other and to understand how different pieces of information can be used to estimate the likelihood of an outcome. In the present study, this translates to the risk of developing C-Diff infection subsequent to colon resection. We can, for example, identify which data features are first-degree associates of an outcome of interest, or directly conditionally dependent, as indicated by an arc in the BBN graph (Figure 1). Furthermore, because the BBN contains estimates of prior probability distributions as well as joint probability distributions of associated features, by entering observed knowledge into the BBN, it can calculate an estimate of the posterior probability of an event.

More importantly, accurate individual estimates can be made in a multitude of different clinical scenarios, even when all the data points are not known. The object in training ml-BBNs was to focus on postoperative estimates of the risk of developing C-Diff colitis that could be determined preoperatively when given some combination of known demographics, diagnosis and procedure codes, and hospital-level information.

Figure 1. Final pruned focused ml-bbn model.



Training and Validation Data

To develop this model, datasets were obtained from the NIS from the US Department of Health and Human Services for the years 2005-2007. Data were conformed to a common specification as described previously. Data from 2005 and 2006 were used to train the models and data from 2007 were withheld to provide an independent validation set of the model. The objective in using a subsequent year to validate the model (versus k-fold cross-validation or classical statistical analysis) was to provide an independent estimate of model robustness. In essence, it answers the question, "If the model were to be used to assess a new patient population for risk of C-Diff, how would the model perform?" A further benefit of independent set validation is that it tends to produce negatively biased testing results.

To assess model robustness and accuracy, the 2007 NIS independent test set was used to plot receiver operating characteristic (ROC) curves for the final pruned model. ROC curves use the posterior probabilities generated by the model to rank each estimate and compare the trade-off between sensitivity and false positive fraction. Using these curves, we were also able to calculate area under the curve (AUC), a metric of overall classification performance of the model. An AUC < 0.50 is not predictive, whereas an AUC = 1 is perfectly predictive.

As AUCs increase between this range, the model has an overall improved ability to predict outcomes. An AUC between 0.7 and 0.8 is considered "fair" in terms of its ability to predict outcomes. ROC curves can also be used to select the optimal calling threshold by selecting a threshold that optimizes for sensitivity and/or specificity. As such, we ran two threshold cases, a high sensitivity case (70%+) and a high specificity case (70%+), to determine which model was optimized.

Modeling and Classification

A stepwise model training and feature selection process was used to train the ml-BBNs developed in this study. This consisted of several stages of recursive modeling and data curation, intended to maximize robustness through the selection

of an appropriate cohort of data features. This iterative process consisted of (1) preliminary modeling, (2) naïve modeling, (3) global modeling, and (4) focused modeling. At each stage, data features were pruned using a combination of expert knowledge and assessment of model structure, specifically focused on the identification of pathways of conditional dependence with the development of C-Diff infection after colectomy. Additionally, data features were repeatedly assessed to evaluate data quality, resulting in additional curation as needed to further semantically normalize and correct errors within the data. Because clinical registries often have high amounts of unknown or missing data, a passive imputation algorithm was used to impute values for those features in which missing data represented less than 50% of total record count and for which there was no adequate substitute feature.

The first step of the process consisted of a naïve Bayesian model specifying development of C-Diff (developedCDiff) as the dependent variable. Because the NIS dataset is so extensive, the naïve model was used to select features that may be conditionally associated with C-Diff from the overall cohort in order to reduce the complexity of the potential solution set and make the remaining models easier to understand without sacrificing predictive power. Using the naïve model structure, a subset of features was identified to be independently associated with development of C-Diff. This naïve model helped as a guide to decide which variables to retain in the full ml-BBN models. Features suggested by the naïve model, together with the new inflammation/infection features (InflamAndOtherInfection), and the features of the pruned preliminary model were used to define the feature set for a full ml-BBN model.

After selecting a reduced set of features based on naïve analysis, a full ml-BBN was trained in the preliminary modeling step. The objective of this step was to identify confounding effects due to data coding or quality issues, and resulted in additional recoding as discussed previously, such as the combined chronic condition variables and recoding of unknown/missing data to nulls for imputation.

Once additional curation was completed, a set of global models was trained to evaluate individual data features for either pruning or inclusion in the final feature set. Additional variables were pruned from the global model using a combination of expert knowledge and structural evaluation of the classifier. Expert knowledge was applied in 3 areas: (1) to identify those features that were proxies for other variables (identical information under a different name), (2) to identify features that were analogs (not identical to other features, but with highly associated distributions), and (3) to identify features that act as confounders in the model.

Since these three types of features increase the complexity of the model and increase computational time while either reducing or not enhancing robustness, these features were pruned from the final feature list. Additionally, we pruned features that were not included in the ml-BBN using “goodness of fit scoring” from the final feature list. The final list of features was trained as an ml-BBN focused model.

Results

Our final training cohort consisted of 56,717 colon resection cases in the NIS during 2005 and 54,480 cases in 2006. Rates of C-Diff infection were 1.58% (n = 895) and 1.65% (n = 934) in 2005 and 2006, respectively. Demographics of the training cohort are described in Table 2. The 2007 NIS set consisted of 57,166 cases of colon resection with a rate of C-Diff infection of 1.86% (n = 1064).

Table 2. Training population descriptive statistics.

General characteristics	Colon resections without C-Diff (n = 111,368)	Colon resections with C-Diff (n = 1829)	*P*	Total (N = 113,197)
Average length of stay (days)	10.5	23.6	<.001	10.8
In-hospital mortality, n (%)	5432 (4.9)	336 (18.4)	<.001	5768 (5.1)
Mean age (years)	62.9	68.3	<.001	63.0
Sex, n (%)				
Female	59,520 (53.4)	1007 (55.1)	.19	60,527 (53.5)
Male	51,848 (46.6)	822 (44.9)	.19	52,670 (46.5)
Disposition (DISPUNIFORM), n (%)				
1 = Routine	69,339 (62.3)	409 (22.4)		69,748 (61.6)
2 or 5 = Other facility	17,787 (16.0)	710 (38.8)		18,497 (16.3)
6 = Home health	18,616 (16.7)	373 (20.4)		18,989 (16.8)
20 = Died in hospital	5464 (4.9)	336 (18.4)		5800 (5.1)
Other	162 (0.1)	1 (0.1)		163 (0.1)
Disease stage [a] (DS_Stage1), n (%)				
0	59 (0.1)	0 (0.0)		59 (0.1)
1	38,973 (35.0)	438 (23.9)		39,411 (34.8)
2	49,452 (44.4)	647 (35.4)		50,099 (44.3)
3	22,884 (20.5)	744 (40.7)		23,628 (20.9)
Mean	1.9	2.2	<.001	1.9
Admission source, n (%)				
1 = Emergency department	34,555 (31.0)	922 (50.4)		35,477 (31.3)
2 = Another hospital	1968 (1.8)	96 (5.2)		2064 (1.8)
3 = Another facility including long-term care	936 (0.8)	61 (3.3)		997 (0.9)
4 = Court/law enforcement	35 (0.0)	0 (0.0)		35 (0.0)
5 = Routine/birth/other	73,372 (65.9)	738 (40.3)		74,110 (65.5)
Unknown	502 (0.5)	12 (0.7)		514 (0.5)
Comorbidities, n (%)				
CM_AIDS (Acquired immune deficiency syndrome)	93 (0.1)	3 (0.2)	.24	96 (0.1)
CM_ALCOHOL (Alcohol abuse)	2163 (1.9)	51 (2.8)	.009	2214 (2.0)
CM_ANEMDEF (Deficiency anemias)	16,846 (15.1)	348 (19.0)	<.001	17,194 (15.2)
CM_ARTH (Rheumatoid arthritis/collagen vascular diseases)	2038 (1.8)	47 (2.6)	.02	2085 (1.8)
CM_BLDLOSS (Chronic blood loss anemia)	5197 (4.7)	97 (5.3)	.20	5294 (4.7)
CM_CHF (Congestive heart failure	9490 (8.5)	364 (19.9)	<.001	9854 (8.7)
CM_CHRNLUNG (Chronic pulmonary disease)	18,120 (16.3)	476 (26.0)	<.001	18,596 (16.4)
CM_COAG (Coagulopathy)	3825 (3.4)	212 (11.6)	<.001	4037 (3.6)
CM_DEPRESS (Depression)	5482 (4.9)	105 (5.7)	.11	5587 (4.9)
CM_DM (Diabetes, uncomplicated)	14,787 (13.3)	224 (12.2)	.20	15,011 (13.3)
CM_DMCX (Diabetes with chronic complications)	1479 (1.3)	50 (2.7)	<.001	1529 (1.4)
CM_DRUG (Drug abuse)	866 (0.8)	11 (0.6)	.39	877 (0.8)
CM_HTN_C (Hypertension combine uncomplicated and complicated)	48,181 (43.3)	738 (40.3)	.01	48,919 (43.2)

General characteristics	Colon resections without C-Diff (n = 111,368)	Colon resections with C-Diff (n = 1829)	*P*	Total (N = 113,197)
CM_HYPOTHY (Hypothyroidism)	8384 (7.5)	116 (6.3)	.06	8500 (7.5)
CM_LIVER (Liver disease)	1733 (1.6)	39 (2.1)	.048	1772 (1.6)
CM_LYMPH (Lymphoma)	666 (0.6)	20 (1.1)	.007	686 (0.6)
CM_LYTES (Fluid and electrolyte disorders)	25,651 (23.0)	881 (48.2)	<.001	26,532 (23.4)
CM_METS (Metastatic cancer)	16,488 (14.8)	178 (9.7)	<.001	16,666 (14.7)
CM_NEURO (Other neurological disorders)	3849 (3.5)	125 (6.8)	<.001	3974 (3.5)
CM_OBESE (Obesity)	5940 (5.3)	53 (2.9)	<.001	5993 (5.3)
CM_PARA (Paralysis)	1085 (1.0)	37 (2.0)	<.001	1122 (1.0)
CM_PERIVASC (Peripheral vascular disorders)	3894 (3.5)	103 (5.6)	<.001	3997 (3.5)
CM_PSYCH (Psychoses)	1841 (1.7)	46 (2.5)	.004	1887 (1.7)
CM_PULMCIRC (Pulmonary circulation disorders)	684 (0.6)	22 (1.2)	.002	706 (0.6)
CM_RENLFAIL (Renal failure)	5042 (4.5)	230 (12.6)	<.001	5272 (4.7)
CM_TUMOR (Solid tumor without metastasis)	2983 (2.7)	62 (3.4)	.06	3045 (2.7)
CM_ULCER (Peptic ulcer disease excluding bleeding)	61 (0.1)	1 (0.1)	.99	62 (0.1)
CM_VALVE (Valvular disease)	5444 (4.9)	135 (7.4)	<.001	5579 (4.9)
CM_WGHTLOSS (Weight loss)	7000 (6.3)	381 (20.8)	<.001	7381 (6.5)
Infectious variables, n (%)				
BacterialInfectionAny	12,693 (11.4)	777 (42.5)	<.001	13,470 (11.9)
BacterialInfectionOther	4531 (4.1)	139 (7.6)	<.001	4670 (4.1)
EColiSepticemia	329 (0.3)	12 (0.7)	.005	341 (0.3)
OtherGramNegSepticemia	472 (0.4)	33 (1.8)	<.001	505 (0.4)
OtherSpecSepticemia	147 (0.1)	51 (2.8)	<.001	198 (0.2)
SexTransInfectNotHIVnorHep	18 (0.0)	0 (0.0)	.59	18 (0.0)
StaphylococcalSepticemia	746 (0.7)	79 (4.3)	<.001	825 (0.7)
StreptococcalSepticemia	367 (0.3)	30 (1.6)	<.001	397 (0.4)
Tuberculosis	140 (0.1)	3 (0.2)	.65	143 (0.1)
UnspecSepticemia	6875 (6.2)	500 (27.3)	<.001	7375 (6.5)
InflamAndOtherInfection	68,447 (61.5)	1554 (85.0)	<.001	70,001 (61.8)

[a] Thomson-Reuters/Medstat: stage of principal disease category

Using the 2007 dataset, the final BBN model had an AUC of 0.746, reflecting an acceptable level of predictive capacity. Posterior probability thresholds of 1.2% and 1.5% were used for the high sensitivity and high specificity scenarios, respectively. The high sensitivity case resulted in a sensitivity of 81.0%, specificity of 50.1%, positive predictive value (PPV) of 2.6%, and negative predictive value (NPV) of 99.4%. In contrast, the high specificity analysis had a sensitivity of 55.4%, specificity of 81.3%, PPV of 3.5%, and a NPV of 99.1%.

As a further validation, we determined that some of the comorbidities might be unknown at the time of colon resection. As such, we wanted to estimate robustness and predictive power of the ml-BBN in the absence of features that may not be known at the time of resection.

These features are as follows (with codes): congestive heart failure (CM_CHF), coagulopathy (CM_COAG), hypothyroidism (CM_HYPOTHY), liver disease (CM_LIVER), neurological disorders (CM_NEURO), paralysis (CM_PARA), peripheral vascular disorders (CM_PERIVASC), pulmonary circulation disorders (CM_PULMCIRC), kidney failure (CM_RENLFAIL), peptic ulcer (CM_ULCER), inflammation and other infections (InflamAndOtherInfection), other gram-negative septicemia (OtherGramNegSepticemia), staphylococcal septicemia (StaphylococcalSepticemia), and streptococcal septicemia (StreptococcalSepticemia).

Excluding these potentially unknown or *ex post facto* features, we re-estimated the posterior probability of C-Diff colitis development for each case in our independent validation set

and recalculated our validation results. The resulting AUC was 0.743—approximately the same as for the entire set with all variables included.

While using the same posterior probability thresholds, the high sensitivity case had a sensitivity of 77.6%, specificity of 52.0%, PPV of 3.7%, and NPV of 99.1%. The high specificity scenario resulted in a sensitivity of 55.9%, a specificity of 78.9%, a PPV of 6.2%, and NPV of 98.9%. In each situation, this is approximately the same as the full dataset. The similar results are likely due to the highly recursive nature of the model structure.

The final focused ml-BBN structure is described in Figure 1, which represents the conditional independence between associate variables. By reading the structure of the network, we can observe that our outcome of "developed C-Diff" (developedCDiff) has 4 first-degree associates: (1) comorbid metastatic cancer (CM_METS), (2) presence of other, non-C-Diff infections (InflamAndOtherInfection), (3) disease staging (DS_Stage1), and (4) patient weight loss (CM_WGHTLOSS). By incorporating the presence or absence of each of these variables, we were able to develop posterior estimates of the probability of C-Diff infection (Table 3).

Table 3. Estimates of *Clostridium difficile* (C-Diff) infection based on presence of risk factors.

Case frequency (%)	Drivers		Target			
	Metastatic tumor	Weight loss	Disease stage	Inflammation/ other infection	Developed C-Diff (%)	
					No	Yes
1.0	No	Yes	Systemic complications	Yes	88.6	11.4
0.0	Yes	Yes	Local complications	No	90.0	10.0
0.1	No	Yes	Systemic complications	No	91.1	8.9
0.0	Yes	Yes	No complications	No	92.3	7.7
0.0	Yes	Yes	No complications	Yes	92.5	7.5
5.1	No	No	Systemic complications	Yes	93.0	7.0
1.0	No	Yes	No complications	Yes	93.7	6.3
2.4	No	Yes	Local complications	Yes	96.3	3.7
0.9	Yes	Yes	Systemic complications	Yes	96.9	3.1
0.2	Yes	No	No complications	No	97.1	2.9
1.1	No	No	Systemic complications	No	97.1	2.9
0.4	Yes	Yes	Systemic complications	No	97.4	2.6
0.0	Yes	Yes	Local complications	Yes	97.7	2.3
0.2	No	Yes	No complications	No	97.8	2.2
0.3	Yes	No	No complications	Yes	97.9	2.1
0.4	Yes	No	Local complications	Yes	98.0	2.0
0.5	No	Yes	Local complications	No	98.3	1.7
24.1	No	No	Local complications	Yes	98.5	1.5
5.3	Yes	No	Systemic complications	Yes	98.5	1.5
21.4	No	No	No complications	Yes	98.8	1.2
0.2	Yes	No	Local complications	No	99.3	0.7
16.6	No	No	Local complications	No	99.4	0.6
11.8	No	No	No complications	No	99.5	0.5
7.1	Yes	No	Systemic complications	No	99.5	0.5

In addition to first-degree associates, the variable developedCDiff, through its first-degree associates also has second-degree associates that can be used to estimate the first-degree associates in patients when there is an absence of information about the first-degree variables (ie, unknown if patient has cancer, weight loss, infection, or staging). With BBN modeling, the user can derive a posterior estimate for the likelihood of C-Diff even with incomplete information. These 15 second-degree associates are as follows (with codes): (1) comorbid lymphoma (CM_LYMPH), (2) comorbid tumor

without metastasis (CM_TUMOR), (3) comorbid chronic disease in diagnosis 1 (CHRONB1mod), (4) comorbid chronic disease in diagnosis 2 (CHRONB2mod), (5) comorbid coagulopathy (CM_COAG), (6) comorbid coagulopathy (CM_COAG), (7) streptococcal septicemia (StreptococcalSepticemia), (8) other gram-negative septicemia (OtherGramNegSepticemia), (9) staphylococcal septicemia (StaphylococcalSepticemia), (10) comorbid peptic ulcer (CM_ULCER), (11) comorbid chronic pulmonary disease (CM_CHRNLUNG), (12) admission source (ASOURCE), (13) age (AGE), (14) comorbid hypertension (CM_HTN_C), and (15) comorbid anemia (CM_ANEMDEF) (see Figure 1).

Discussion

Four first-degree associates that influence the probability of C-Diff development were identified: weight loss, tumor metastases, inflammation/infections, and disease severity. Furthermore, ml-BBNs can produce robust estimates of postoperative C-Diff infection.

The incidence of C-Diff colitis is steadily rising, with most institutions citing rates among all hospitalized patients approaching 1% [19] and as high as 10% in general medical ward patients hospitalized for at least 2 days.

The combination of recent hospitalizations and frequent antibiotic use has led to a near epidemic of chronic carrier states in long-term care facilities. The impact on the health care system is also significant, not only in terms of increased morbidity, but also in terms of escalating costs due to the requirement for patient isolation, personnel protective equipment, and overall care [20]. Although the majority of these patients remain as asymptomatic carriers or only experience mild diarrhea, more fulminant disease may ensue [21]. Yet, C-Diff colitis can also present following elective colonic resection for various disease states ranging from diverticulitis and cancer to inflammatory bowel disease. Among our select cohort of patients undergoing colonic resection, we found a secondary rate of C-Diff colitis of 1.86% for 2007. This is consistent with a slow rise in the years preceding our study, in which the estimated incidence was 14.9 cases per 1000 postoperative hospitalized patients between 2003 and 2005 [8]. Compounding the impact of this recent surge is the accompanying increase in disease recurrence (particularly with the NAP1/B1/027 strain), refractory infections, and the increased clinical severity of cases, with particularly high treatment-related mortality for severe, complicated C-Diff colitis [22]. Previous factors associated with higher morbidity and mortality from C-Diff colitis include low serum albumin, intensive care unit admission, older age [23,24], and poor immunologic response to toxins released by the bacteria [25]. Because each of these factors results in higher rates and disease that is more virulent, identifying those patients at risk and preventing its onset is of paramount importance.

Importantly, current classification systems for C-Diff colitis often understage disease severity, and underscore the need for better models [26].

Estimating the risk of disease-specific outcomes can decidedly improve the management of patients undergoing colonic resection. The goal of this study was, therefore, to create predictive models to provide information on how readily available clinical and disease-specific factors can, in a codependent manner, collectively influence postoperative outcomes through preoperative risk assessment. Machine-learned BBNs have previously been demonstrated to be effective in other areas of medicine, such as estimating risk and prognosis of cancer in patients included in various cancer registries [27,28]. Furthermore, ml-BBNs have the added advantage of providing more accurate estimates when not all the data are known. Although the AUCs of 0.74 and 0.75 predict a “fair” level of predictive capacity, we calculated both high sensitivity and high specificity scenarios to optimize the model. Given the superb model robustness demonstrated through cross-validation in the present study, along with the high degree of variance that can be derived in terms of estimates *a posteriori*, these models provide the basis for an easily usable, personalized medical CDSS even when confronted with limited data.

Given detailed information, the model can also be used as an individualized patient-specific calculator. Table 3 illustrates one mode of using the trained and validated ml-BBN. It uses the 4 first-degree associates—comorbid metastatic cancer (CM_METS), presence of other, non-C-Diff infections (InflamAndOtherInfection), disease staging (DS_Stage1), and patient weight loss (CM_WGHTLOSS)—to estimate the posterior probability of C-Diff given knowledge of these 4 factors. This table represents all possible cases (total 24) within the first-degree associates and their related estimated frequency and posterior probability of C-Diff.

Those cases that exceed the 1.5% threshold (the high specificity threshold) represent an estimated 13.5% of cases, whereas the below-threshold cases represent 86.5% of cases. The value of Table 3 lies in its ability to illustrate how the model can be used to develop estimates of outcome when individual factors are considered collectively. Thus, although weight loss, metastatic cancer, and complications have individual contributions to the likelihood of developing C-Diff colitis, they also have a specific influence on probability when acting together. These estimates are derived from the observed rate of outcome within each subpopulation. In the context of this analysis, variables such as weight loss and systemic complications occur fairly frequently (1 per 100 patients), whereas a case with metastatic disease, weight loss, and local complications is extremely rare (only a handful of patients in our training set). This partly explains our low PPV. Yet, with relatively infrequent incidence in the bigger picture, it is ideal to have a higher NPV, as demonstrated with our model. When expanded to a Web-based application, several other variables (included in Figure 1) could be present in “drop-down” menus in which the provider could place known values, individualizing the patient-specific estimate of disease even further.

A unique aspect of our study is that it evaluates the incidence of C-Diff colitis development following resection for other primary diagnoses rather than focus on surgical therapy of C-Diff colitis itself. This is of particular relevance at a time when the rate of moderate to severe C-Diff colitis is an area of active study covering aspects from vancomycin enemas and

fecal transplants, to diverting ileostomy with colonic lavage, or total abdominal colectomy and end ileostomy [29,30].

More pertinent, we were able to identify factors that are often known prior to surgery which increase the risk of the development of C-Diff colitis. These can serve as focal points for intervention such as improving nutrition (for weight loss), treating infection, and optimizing management of systemic disease. With these central efforts, pathways can be implemented to attempt to prevent the onset of C-Diff colitis altogether. Expanding this to an online CDSS will give physicians 24-hour access to input all known data including all the variables (ie, both first- and second-degree associates) to estimate the probability of C-Diff infection following surgery. Decisions could then be made whether to pursue surgery or direct further care prior to surgery. Even beyond the morbidity and mortality, C-Diff infection during hospitalization results in a US $77,000 additional cost per admission, and increases the length of stay by 16 days [31].

We acknowledge some limitations to our study. As in any registry study, there are many issues with data consistency and completeness, as discussed in the Methods section, that required clinical judgment applied to data preparation for analysis. The study team attempted to address these shortcomings through a combination of data curation and censoring, but ultimately these issues cannot be perfectly resolved and we had to rely on the use of ml-BBN independent set validation to assess the impact of database inconsistencies on model accuracy. Also, although the NIS provides a large sample size, it lacks details specific to patients' hospital courses, including specific antibiotic use, status of chronic carrier states, and degree/severity of comorbid conditions, that could help draw definitive conclusions regarding our endpoints. Finally, the retrospective nature of this analysis likely introduces bias that would not be present in a prospective study.

Despite these limitations, this study does provide useful models that can be easily and readily used to derive case-specific estimates of the development of C-Diff infection for use in identification of high-risk patients and adjusting treatment planning to minimize the onset of C-Diff postoperatively.

Conclusion

In a large cohort of patients undergoing colonic resection, we have found secondary development of C-Diff colitis to be associated with significant morbidity and mortality. Machine-learned BBN can be used to create robust classifiers capable of estimating the probability of C-Diff colitis preoperatively in patients undergoing colectomy. By identifying high- and low-risk cohorts, physicians can be more aware of patients at additional risk and implement strategies to minimize the probability of secondary C-Diff infection.

Acknowledgments

This clinical research effort was supported, in part, by the United States Military Cancer Institute (USMCI), Washington, DC, the Henry M. Jackson Foundation (HJF) for the Advancement of Military Medicine, Rockville, MD, and the California Oncology Research Institute (CORI), Los Angeles, CA. Our team is comprised partly of military service members and employees of the US Government. This work was prepared as part of our official duties. Title 17 USC 105 provides that "Copyright protection under this title is not available for any work of the United States Government." Title 17 U.S.C. 101 defines a US Government work as a work prepared by a military service member or employee of the US Government as part of that person's official duties. The views expressed in this presentation are those of the authors and do not reflect the official policy of the Department of the Army, the Department of Defense, or the United States Government.

This paper was presented as an e-poster presentation at the 2012 Pacific Coast Surgical Association Meeting in Napa, California, on February 18, 2012.

Authors' Contributions

Conception and design: Stojadinovic, Eberhardt, Kalina, Steele; acquisition of data: Eberhardt, Kalina; analysis and interpretation of data: Stojadinovic, Eberhardt, Kalina, Nissan, Avital, Bilchik, Steele; drafting of manuscript: Stojadinovic, Eberhardt, Kalina, Nissan, Johnson, Bilchik, Steele; critical revision: Stojadinovic, Johnson, Eberhardt, Nissan, Avital, Bilchik, Steele; statistical expertise: Eberhardt, Kalina; and supervision: Stojadinovic, Steele.

Conflicts of Interest

None declared.

References

1. Senior K. Concern over Clostridium difficile in the USA. The Lancet Infectious Diseases 2008;8(6):352. [doi: 10.1016/S1473-3099(08)70115-4]
2. Khanna S, Pardi DS, Aronson SL, Kammer PP, Orenstein R, St Sauver JL, et al. The epidemiology of community-acquired Clostridium difficile infection: a population-based study. Am J Gastroenterol 2012 Jan;107(1):89-95. [doi: 10.1038/ajg.2011.398] [Medline: 22108454]
3. Johnson S, Clabots CR, Linn FV, Olson MM, Peterson LR, Gerding DN. Nosocomial Clostridium difficile colonisation and disease. Lancet 1990 Jul 14;336(8707):97-100. [Medline: 1975332]

4. Kyne L, Sougioultzis S, McFarland LV, Kelly CP. Underlying disease severity as a major risk factor for nosocomial Clostridium difficile diarrhea. Infect Control Hosp Epidemiol 2002 Nov;23(11):653-659. [doi: 10.1086/501989] [Medline: 12452292]
5. Gerding DN, Muto CA, Owens RC. Treatment of Clostridium difficile infection. Clin Infect Dis 2008 Jan 15;46 Suppl 1:S32-S42 none [FREE Full text] [doi: 10.1086/521860] [Medline: 18177219]
6. Warny M, Pepin J, Fang A, Killgore G, Thompson A, Brazier J, et al. Toxin production by an emerging strain of Clostridium difficile associated with outbreaks of severe disease in North America and Europe. Lancet 2005;366(9491):1079-1084. [doi: 10.1016/S0140-6736(05)67420-X] [Medline: 16182895]
7. Efron PA, Mazuski JE. Clostridium difficile colitis. Surg Clin North Am 2009 Apr;89(2):483-500. [doi: 10.1016/j.suc.2008.09.014] [Medline: 19281896]
8. Wilkinson E. C difficile risk with prophylactic antibiotics during surgery. The Lancet Infectious Diseases 2008;8(7):412. [doi: 10.1016/S1473-3099(08)70144-0]
9. Balassiano IT, Yates EA, Domingues RM, Ferreira EO. Clostridium difficile: a problem of concern in developed countries and still a mystery in Latin America. J Med Microbiol 2012 Feb;61(Pt 2):169-179. [doi: 10.1099/jmm.0.037077-0] [Medline: 22116982]
10. Clements AC, Magalhães RJ, Tatem AJ, Paterson DL, Riley TV. Clostridium difficile PCR ribotype 027: assessing the risks of further worldwide spread. Lancet Infect Dis 2010 Jun;10(6):395-404. [doi: 10.1016/S1473-3099(10)70080-3] [Medline: 20510280]
11. Ricciardi R, Rothenberger DA, Madoff RD, Baxter NN. Increasing prevalence and severity of Clostridium difficile colitis in hospitalized patients in the United States. Arch Surg 2007 Jul;142(7):624-31; discussion 631. [doi: 10.1001/archsurg.142.7.624] [Medline: 17638799]
12. Wren SM, Ahmed N, Jamal A, Safadi BY. Preoperative oral antibiotics in colorectal surgery increase the rate of Clostridium difficile colitis. Arch Surg 2005 Aug;140(8):752-756. [doi: 10.1001/archsurg.140.8.752] [Medline: 16103284]
13. Krapohl GL, Phillips LR, Campbell DA, Hendren S, Banerjee M, Metzger B, et al. Bowel preparation for colectomy and risk of Clostridium difficile infection. Dis Colon Rectum 2011 Jul;54(7):810-817. [doi: 10.1007/DCR.0b013e3182125b55] [Medline: 21654247]
14. Lesperance K, Causey MW, Spencer M, Steele SR. The morbidity of Clostridium difficile infection after elective colonic resection-results from a national population database. Am J Surg 2011 Feb;201(2):141-148. [doi: 10.1016/j.amjsurg.2010.09.017] [Medline: 21266214]
15. Osheroff JA. Improving Medication Use and Outcomes with Clinical Decision Support: A Step-by-Step Guide. Chicago, IL: The Healthcare Information and Management Systems Society Publishers; 2009.
16. Knops AM, Goossens A, Ubbink DT, Legemate DA, DECAID study group. Regarding "pilot testing of a decision support tool for patients with abdominal aortic aneurysms". J Vasc Surg 2011 Jun;53(6):1757. [doi: 10.1016/j.jvs.2010.12.071] [Medline: 21609808]
17. Kim S, Moon SK, Jung DC, Hwang SI, Sung CK, Cho JY, et al. Pre-operative prediction of advanced prostatic cancer using clinical decision support systems: accuracy comparison between support vector machine and artificial neural network. Korean J Radiol 2011;12(5):588-594 [FREE Full text] [doi: 10.3348/kjr.2011.12.5.588] [Medline: 21927560]
18. Healthcare Cost and Utilization Project (HCUP). 2011. Appendix C - Clinical Classification Software - Expanded Diagnosis Categories (January 1980 through September 2012) URL: http://www.hcup-us.ahrq.gov/toolssoftware/ccs/AppendixCMultiDX.txt [accessed 2012-02-12] [WebCite Cache ID 6ASHpBK0j]
19. Fry DE. A systems approach to the prevention of surgical infections. Surg Clin North Am 2009 Apr;89(2):521-37, x. [doi: 10.1016/j.suc.2008.09.010] [Medline: 19281898]
20. Fletcher KR, Cinalli M. Identification, optimal management, and infection control measures for Clostridium difficile-associated disease in long-term care. Geriatr Nurs 2007;28(3):171-81; quiz 182. [doi: 10.1016/j.gerinurse.2007.02.003] [Medline: 17561015]
21. McFarland LV, Mulligan ME, Kwok RY, Stamm WE. Nosocomial acquisition of Clostridium difficile infection. N Engl J Med 1989 Jan 26;320(4):204-210. [doi: 10.1056/NEJM198901263200402] [Medline: 2911306]
22. Bendle JS, James PA, Bennett PM, Avison MB, Macgowan AP, Al-Shafi KM. Resistance determinants in strains of Clostridium difficile from two geographically distinct populations. Int J Antimicrob Agents 2004 Dec;24(6):619-621. [doi: 10.1016/j.ijantimicag.2004.06.013] [Medline: 15555889]
23. Kenneally C, Rosini JM, Skrupky LP, Doherty JA, Hollands JM, Martinez E, et al. Analysis of 30-day mortality for clostridium difficile-associated disease in the ICU setting. Chest 2007 Aug;132(2):418-424. [doi: 10.1378/chest.07-0202] [Medline: 17573523]
24. Fernandez A, Anand G, Friedenberg F. Factors associated with failure of metronidazole in Clostridium difficile-associated disease. J Clin Gastroenterol 2004;38(5):414-418. [Medline: 15100520]
25. McPherson S, Rees CJ, Ellis R, Soo S, Panter SJ. Intravenous immunoglobulin for the treatment of severe, refractory, and recurrent Clostridium difficile diarrhea. Dis Colon Rectum 2006 May;49(5):640-645. [doi: 10.1007/s10350-006-0511-8] [Medline: 16525744]

26. Fujitani S, George WL, Murthy AR. Comparison of clinical severity score indices for Clostridium difficile infection. Infect Control Hosp Epidemiol 2011 Mar;32(3):220-228. [doi: 10.1086/658336] [Medline: 21460506]
27. Forsberg JA, Eberhardt J, Boland PJ, Wedin R, Healey JH. Estimating survival in patients with operable skeletal metastases: an application of a bayesian belief network. PLoS One 2011;6(5):e19956 [FREE Full text] [doi: 10.1371/journal.pone.0019956] [Medline: 21603644]
28. Stojadinovic A, Nissan A, Eberhardt J, Chua TC, Pelz JO, Esquivel J. Development of a Bayesian Belief Network Model for personalized prognostic risk assessment in colon carcinomatosis. Am Surg 2011 Feb;77(2):221-230. [Medline: 21337884]
29. Neal MD, Alverdy JC, Hall DE, Simmons RL, Zuckerbraun BS. Diverting loop ileostomy and colonic lavage: an alternative to total abdominal colectomy for the treatment of severe, complicated Clostridium difficile associated disease. Ann Surg 2011 Sep;254(3):423-7; discussion 427. [doi: 10.1097/SLA.0b013e31822ade48] [Medline: 21865943]
30. Kassam Z, Hundal R, Marshall JK, Lee CH. Fecal transplant via retention enema for refractory or recurrent Clostridium difficile infection. Arch Intern Med 2012 Jan 23;172(2):191-193. [doi: 10.1001/archinte.172.2.191] [Medline: 22271132]
31. Zerey M, Paton BL, Lincourt AE, Gersin KS, Kercher KW, Heniford BT. The burden of Clostridium difficile in surgical patients in the United States. Surg Infect (Larchmt) 2007 Dec;8(6):557-566. [doi: 10.1089/sur.2006.062] [Medline: 18171114]

Abbreviations

AHRQ: Agency for Healthcare Research and Quality
AUC: area under the curve
BBN: Bayesian belief network
CCS: clinical classifications software
C-Diff: Clostridium difficile
CDSS: clinical decision-support system
ml-BBN: machine-learned Bayesian belief network
NIS: National Inpatient Sample
NPV: negative predictive value
PPV: positive predictive value
ROC: receiver operating characteristic

Edited by G Eysenbach; submitted 18.04.12; peer-reviewed by S Keith; comments to author 08.06.12; revised version received 11.06.12; accepted 11.06.12; published 19.09.12.

Please cite as:
Steele S, Bilchik A, Eberhardt J, Kalina P, Nissan A, Johnson E, Avital I, Stojadinovic A
Using Machine-Learned Bayesian Belief Networks to Predict Perioperative Risk of Clostridium Difficile Infection Following Colon Surgery
Interact J Med Res 2012;1(2):e6
URL: http://www.jmir.org/2012/2/e6/
doi:10.2196/ijmr.2131
PMID:23611947

Original Paper

Improving Hospital Care and Collaborative Communications for the 21st Century: Key Recommendations for General Internal Medicine

Robert C Wu[1,2,3], MD, MSc; Vivian Lo[2], MASc; Peter Rossos[2,3,4,5], MD MBA FRCPC; Craig Kuziemsky[6], PhD; Kevin J O'Leary[7], MD; Joseph A Cafazzo[5,8,9], PhD; Scott Reeves[10,11,12,13,14], PhD; Brian M Wong[3,15,16], MD; Dante Morra[1,2,3,17], MD MBA FRCPC

[1]Toronto General Hospital, Division of General Internal Medicine, University Health Network, Toronto, ON, Canada
[2]Toronto General Hospital, Centre for Innovation in Complex Care, University Health Network, Toronto, ON, Canada
[3]Department of Medicine, University of Toronto, Toronto, ON, Canada
[4]Shared Information Management Services, University of Health Network, Toronto, ON, Canada
[5]Centre for Global eHealth Innovation, University Health Network, Toronto, ON, Canada
[6]Telfer School of Management, University of Ottawa, Ottawa, ON, Canada
[7]Feinberg School of Medicine, Department of Medicine, Northwestern University, Chicago, IL, United States
[8]Department of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada
[9]Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada
[10]Centre for Faculty Development, Li Ka Shing International Health Care Education Centre, St Michael's Hospital, Toronto, ON, Canada
[11]Li Ka Shing Knowledge Institute, Keenan Research Centre, St Michael's Hospital, Toronto, ON, Canada
[12]Wilson Centre for Research in Education, University Health Network, Toronto, ON, Canada
[13]Department of Psychiatry, University of Toronto, Toronto, ON, Canada
[14]Centre of Innovation in Interprofessional Health Education, University of California, San Francisco, San Francisco, CA, United States
[15]Sunnybrook Health Sciences Centre, Department of Medicine, Toronto, ON, Canada
[16]Centre for Patient Safety, University of Toronto, Toronto, ON, Canada
[17]Centre for Interprofessional Education, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Robert C Wu, MD, MSc
Toronto General Hospital
Division of General Internal Medicine
University Health Network
TGH 14EN-222
200 Elizabeth Street
Toronto, ON, M5G 2C4
Canada
Phone: 1 416 340 4567
Fax: 1 416 595 5826
Email: robert.wu@uhn.on.ca

Abstract

Background: Communication and collaboration failures can have negative impacts on the efficiency of both individual clinicians and health care system delivery as well as on the quality of patient care. Recognizing the problems associated with clinical and collaboration communication, health care professionals and organizations alike have begun to look at alternative communication technologies to address some of these inefficiencies and to improve interprofessional collaboration.

Objective: To develop recommendations that assist health care organizations in improving communication and collaboration in order to develop effective methods for evaluation.

Methods: An interprofessional meeting was held in a large urban city in Canada with 19 nationally and internationally renowned experts to discuss suitable recommendations for an ideal communication and collaboration system as well as a research framework for general internal medicine (GIM) environments.

Results: In designing an ideal GIM communication and collaboration system, attendees believed that the new system should possess attributes that aim to: a) improve workflow through prioritization of information and detection of individuals' contextual situations; b) promote stronger interprofessional relationships with adequate exchange of information; c) enhance patient-centered care by allowing greater patient autonomy over their health care information; d) enable interoperability and scalability between and within institutions; and e) function across different platforms. In terms of evaluating the effects of technology in GIM settings, participants championed the use of rigorous scientific methods that span multiple perspectives and disciplines. Specifically, participants recommended that consistent measures and definitions need to be established so that these impacts can be examined across individual, group, and organizational levels.

Conclusions: Discussions from our meeting demonstrated the complexities of technological implementations in GIM settings. Recommendations on the design principles and research paradigms for an improved communication system are described.

(Interact J Med Res 2012;1(2):e9) doi:10.2196/ijmr.2022

KEYWORDS

hospital care communication; technology; knowledge transfer; interprofessional collaboration

Introduction

Interprofessional communication between clinicians has traditionally relied on numeric paging systems that are riddled with numerous problems [1-4]. These include difficulty in identifying and contacting the right clinician, limited capability as a one-way receiver of information, and frequent interruptions. These may contribute to medical error and often result in increased frustration amongst clinicians [5,6]. There is a significant impact—poor communication practices with the resulting breakdowns on health care delivery. In a review of 14,000 hospital admissions, poor communication and collaboration practices were identified as the most common cause of preventable clinical errors [7]. Communication failures were also rated as one of the top most preventable causes for all known clinical errors [8]. Communication inefficiencies that resulted in wasted time for clinicians and increased length of stay for patients could cost North American hospitals as much as 12 billion dollars per year [9].

Recognizing the costs and problems associated with ineffective clinical communication, health care organizations have begun to look at emerging communication technologies to address some of these inefficiencies [10-12]. This coincides with an increased uptake of new mobile communications devices by clinicians, with an estimated 81% of North American physicians who currently own or use smartphone technology [13]. New communication technology that combines mobile phone, text, and email functions has the potential to improve clinical communication. Various hospitals have attempted different technological solutions to enhance their communication processes. These solutions have included the use of wireless email, two-way alphanumeric paging, smartphone communication, and a web-based communication tool that queues non-urgent messages [14].

Yet despite the increased adoption of communication technologies in health care, there is very little research that evaluates the effectiveness of these information and communication systems. Furthermore, rapidly changing technology is making evaluations and interpretations of these implementations challenging. Recent systematic reviews of smartphones concluded that there is limited evidence demonstrating that the use of technology leads to direct improvements in either clinician efficiency or patient care [14,15]. Consequently, many questions remain on the best evidence-based practices and strategies for these communication technologies.

Recognizing these gaps, an interprofessional meeting was organized to bring a diverse group of leading experts and stakeholders together to discuss effective ways to design and evaluate communication systems that can enhance clinical communication processes in health care organizations. Funded by the Canadian Institutes of Health Research (CIHR), the key objectives of the meeting were to 1) develop principles of an effective communication system, 2) identify key research priorities and paradigms in General Internal Medicine (GIM)–related communication areas, and 3) propose methods in which technological changes can be both implemented and evaluated successfully.

Methods

Workshop

An Interprofessional Communication and Collaboration Meeting was held in Toronto, Canada, on April 29, 2011, to discuss solutions that address the rapidly changing demands and needs of hospital communication and collaboration. Attendees came from different clinical and academic backgrounds to ensure that heterogeneous viewpoints were represented on a wide range of topics relating to interprofessional collaboration and information and communication technologies (ICT). The meeting involved the following steps: 1) the identification of an array of experts who were invited to participate from the fields of clinical communication and health care technologies. These experts were identified through literature reviews and professional networks and represented a diversity of professions that included medicine, nursing, pharmacy, academic research, health informatics, engineers, and hospital administration; 2) attendees were asked to provide position statements on the design and research models of an interdisciplinary team communication system, which were compiled and circulated among the group for review prior to the meeting; 3) the meeting began with presentations delivered by three of the invited experts who described their current research work and ICT implementations conducted at their respective hospitals and/or institutions; 4)

based on their area(s) of expertise, participants were subsequently assigned to one of the two working groups: A) Design and Implementation or B) Research and Evaluation; 5) the working groups were co-facilitated by the invited experts who worked through previously formulated case studies and key questions (see Appendix 1) to develop recommendations for their respective domains; 6) the recommendations were presented to the larger group, which concluded with roundtable discussions among the attendees on the themes and issues raised during the sessions; 7) written field notes of the discussions were recorded by note-takers and subsequently transcribed into raw Microsoft Word documents after the meeting; 8) using an inductive thematic analysis, recommendations and discussion notes were then analyzed, summarized, and drafted for an initial report by two members of the research team (RW and VL); and 9) the report was sent to all participants for content validation.

Workshop Participants

The key issues related to clinical interprofessional communication and collaboration transcended the traditional disciplinary boundaries and demanded a broad range of interests and areas of expertise. These professional groups included medicine, nursing, pharmacy, academic research, health informatics, engineering, and hospital administration. Experts were identified through published literature as well as recommendations from professional networks. The invitees were selected by reviewing their experience and knowledge in various domains, publication records, and participation in initiatives or projects related to clinical interprofessional communication or ICT. Despite the efforts to have geographic, academic, and clinical diversity among the meeting invitees, many of the identified experts came from a few regions and organizations where active work in the field was being undertaken. To facilitate the process of inviting national and international experts, we made every effort to ensure out-of-town participants were able to attend in person or via videoconferencing (Skype). The majority of the invitees were able to attend the meeting in person with the exception of one attendee who participated via Skype.

Definitions

Interprofessional communication was defined as information exchanges of patient-related issues between different care providers and professions. These included face to face, verbal, and text messages, and both scheduled communications such as interprofessional rounds and “as needed” or “ad-hoc” communications.

Interprofessional collaboration was defined as different professions working together as a team toward a common goal of providing optimal patient care using the skills/expertise of other professions.

Results

Nineteen expert participants from Canada and the United States attended the Interprofessional Communication and Collaboration Meeting. This summary presents the major themes developed by the attendees in each session and concludes with issues and prospects for both the design/implementation and research communities (Table 1).

Table 1. Summary of major themes for design recommendations of an ideal communication system and recommendations for future research.

Theme	Subtheme
Key Principles of the Ideal System	
	Safety – The new communication system should help minimize communication errors and improve patient safety.
	Patient-Centered Focus – The system centers on the patient instead of specific providers, promoting the inclusiveness of all individuals and team integration.
	Cost – The cost of the existing communication inefficiencies outweighs the financial burden of implementing a new communication system.
Design Recommendations	
	Improve Workflow Through Contextual Awareness and Prioritization – Minimize interruptions by allowing message receivers to set their availability and prioritize messages by urgency.
	Promote Stronger Collaborative Relationships – Provide capability to communicate to more than one team members.
	Enhance Patient-Centered Care – Allow patients to be a part of the communications.
	Allow Interoperability and Scalability – Allow communication to clinicians providing care to patients, regardless of the institution with which they are affiliated.
	Support Multiple Technologies – Support different communication technologies such as pagers, cell phones, and different types of smartphones.
Research Recommendations	
	Considerations of the Contexts and Processes in Which Technology are Embedded – A broad approach looking at the processes and contexts in which these technologies are adopted including professional and organizational cultures.
	Need for Extensive Research Frameworks – It is important to examine the processes and how technology interacts from multiple perspectives, including 1) Education, 2) Clinical Practices, 3) Culture, 4) Interprofessional Collaboration and Communication, and 5) Organization of Care.
	Need for Multifaceted Outcome Measurements – Mixed methods consisting of qualitative and quantitative approaches should be used to obtain multiple data sources when evaluating complex interventions.

Group A: Design and Implementation

A total of eight attendees were assigned to the Design and Implementation group. Participants of the group were tasked to identify important design principles for an improved communication and collaboration system that can be adopted in health care settings. The following key themes emerged from the discussions.

Themes

Key Fundamentals/Principles of the Ideal System

During the discussions, the group agreed that there should be key principles to guide in the design of the system such that the new communication system should produce positive or neutral impact at the very least. Specifically, the impact should affect the following areas:

Safety

The new communication system should help to minimize communication errors and improve patient safety. Participants noted that in the growing body of literature, patient safety has often been compromised by factors such as frequent interruptions in clinicians' routines, which are exacerbated by the untimely delivery of messages. Thus, participants agreed that an ideal communication system should both minimize the occurrences of adverse events and improve patient safety.

Patient-Centered Focus

The interprofessional team should consider using one communication system to collaborate on the care of patients, as opposed to separate systems that merely serve the needs of different professions. Participants recognized that communication inefficiencies could create unnecessary patient inconvenience. Thus, there is a necessity for a system that could a) coordinate and manage all parts of the communication process, b) accomplish a defined set of goals that streamline workflow, and c) maximize practice efficiency and productivity amongst clinicians. Consequently, participants believed that the optimal model should encompass solutions that address these gaps of health care communication by promoting a) inclusiveness of all individuals and team integration, b) standardization the language used between team members, and c) timeliness and accuracy of information so that high-quality patient care can be delivered.

Cost

The cost of the existing communication inefficiencies outweighs the financial burden of implementing a new communication system. Participants acknowledged that strategic and operational considerations also need to be included in the long-term decision-making plans and sustainability of the health care systems. Thus, an important consideration for analysis is to understand the potential value, opportunities, and cost associated

with the modifications or changes in the use of clinical communication resources. Subsequently, the incremental improvements in the new clinical communication system should have major positive impacts on overall health care expenditure, which should outweigh the financial cost of the new system.

Design Recommendations

From the discussions, the following key design recommendations emerged as important features to a successful clinical communication system:

Improve Workflow Through Contextual Awareness and Prioritization

Participants observed that existing disruptions in communication and information transfer often create unnecessary inefficiencies in clinical workflow. Thus, one of the key goals of the ideal communication system should aim to improve clinical workflow whereby only essential communication interrupts clinicians. This could be accomplished through appropriate prioritization. However, it was acknowledged that the concept of urgency may vary depending on the perspective of the sender and the receiver. That is, a message deemed to be urgent by the sender may not be considered urgent by the receiver (and vice versa). Furthermore, although different clinicians may agree on the categorization of messaging, their perception of appropriate response time may differ. For example, although medical and nursing staff may agree that a meeting with the patient's family members for an update is "non-urgent", they may disagree as to how long the family should wait to speak to the physician. This discordance can escalate messaging frequency and in some cases lead to conflict. Some attendees proposed adopting practices such as standardizing a set of guidelines that define urgent and non-urgent issues. Others advocated that the system should provide clinicians the flexibility to designate the urgency of the message and the status of receiver. One possibility is to design a system that allows clinicians to indicate their situation and availability (ie, "context awareness"). For example, the system should enable clinicians from a range of professions to indicate their locations (eg, in an isolation room, in a teaching session, in the operating room) as well as their ability to respond to messages (eg, available, performing a procedure—do not interrupt, in a critical family meeting—do not interrupt). In the absence of a response from the original receiver, the system should have an algorithm to help escalate the sender's messages to the next level and provide two-way feedback loops to both the sender and the receiver that the message and information has been escalated and dealt with.

Promote Stronger Collaborative Relationships

The second attribute of the ideal communication system should include features that promote teamwork and create stronger collaborative relationships. One of the biggest barriers to collaboration and cohesion involves misunderstandings and frustrations over discrepancies in the flow of information between clinicians. Often, clinicians face challenges identifying members of the patient's care team resulting in information being transmitted between single individuals that excluded other GIM staff. Thus, another key aspect of the ideal system is to ensure an accurate list of the different clinicians caring for the patient that is integrated with the necessary communication channels. This would make it easy for all clinicians on the team to be included in the conversations by being updated and kept in the loop on the information they need. Moreover, in the area of message distribution, the system should have functionalities that allow mass broadcasting to all members of the clinical teams but also "tiering" for subgroups to receive specific targeted messages. Nonetheless, despite the support for new technological implementations and functionalities, caution should be taken to ensure that the new technology does not replace face-to-face interaction but rather augment it in ways that support team collaboration.

Enhance Patient-Centered Care

Participants agreed that while the system is designed for use mainly among clinicians, it should be one that promotes patient-centered communication focusing on unified communication strategies to connect patients and providers. Drawing from the perspective of the social networking model that looks at how relationships between individuals are connected between one another, attendees advocated that the patient should also be allowed to partake in the communication dialogue. At a minimum, the ideal communication system should keep patients informed on who is caring for them and be updated on the status of their care or treatment plans. Furthermore, attendees believed that patients should have a voice in how they would like to share their health care information. One possible way identified during the discussions was to design a communication system that either allows patients, family members, and/or physicians to determine the level of privacy and security in the flow of the patient health information and records. Thus, a patient-centric approach with the flexibility to designate user views and preferences should be considered. Nevertheless, these technological components should also be grounded in the principles of equality such that the system is highly usable by patients and their family members and is accessible regardless of age, culture, or ethnicity.

Allow Interoperability and Scalability

Attendees agreed that interoperability between institutions and health settings should be another important feature of the ideal system. Since many patients have multiple chronic diseases, they have multiple care providers distributed throughout the institutions and health settings. Specifically, the ideal system should establish common and understandable professional and interprofessional language and platforms that meet both clinicians' and stakeholders' needs. Additionally, the new system should be cost-effective and scalable to different hospital sizes, policies, and requirements. Moreover, the group agreed that user guidance and education should be provided alongside communication system implementation and knowledge transfer phases to encourage everyone to share the same goals and objectives.

Support Multiple Technologies

To date, industry solutions have focused on either expensive proprietary systems or the deployment of business/consumer-level communication devices and platforms in the health care marketplace. The lack of open and common standards is a significant barrier to application and device development for health care mobility at the point of care. Thus,

considering the rapid advancement and evolution of new technology, the ideal system should be "device agnostic". That is to say, the system should have the capabilities to accommodate different mediums of communication technologies that leverage diverse platforms and technologies rather than being constrained to one specific channel or platform. Nonetheless, the group also acknowledged the challenges and complexities of managing different communication mediums. Specifically, having multiple channels of communication (paging, text messaging, calling, or others) but without having a common understanding when it is appropriate to use each channels of communication, considerable confusion may be created among users. This could result in situations of delay and misunderstanding that could be worse than the traditional, simpler communication systems.

Group B: Research and Evaluation

A total of 11 attendees were assigned to the Research and Evaluation group. Participants of the group were tasked with identifying the current research gaps and approaches in evaluating the effects and impact of a new communication system implementation. Specifically, participants were asked to consider the types of frameworks, methodologies, and theories that ought to be adopted when conducting evaluation research in the area of clinical communication and technology. The following themes emerged:

Themes

Research Recommendations

Considerations of the Contexts and Processes in Which Technology are Embedded

Participants recognized that current assessments and perspectives of health care communication are often fragmented. Existing studies often focus on technology and its direct impact alone, which is inadequate and limiting in understanding the complexities of communication in health care settings. Instead, participants advocated a broader approach of looking at the *processes* and *contexts* to which these technologies are adopted. Specifically, research on health care communication domains—such as interprofessional collaboration and ICT—need to consider other factors such as professional and organizational culture(s) and socialization processes, which are often intertwined. Experts argued that new technology interventions and designs should consider challenging the traditional workflow.

Attendees also recognized that when organizations are transitioning between systems, it is important to be aware that certain information or collaborative opportunities may be lost in the adoption of the new technology. Thus, it is important for researchers to consider the roles of technology and its processes at the organizational level by exploring interactions that occur between institutional cultures and technology. In particular, participants believed that the existence of different patterns of communication—brought about by the cultural aspects of professional tensions and hierarchies—may be the critical pieces to improving communication and collaboration.

Need for Extensive Research Frameworks

Participants agreed that existing research frameworks have been successful at tracking metrics such as monitoring patient outcomes that are often found in quality improvement interventions. Although these quality improvement studies may have provided meaningful knowledge on how to enhance the quality and safety of care, these studies often lacked the scholarly conceptualizations, as seen in the dearth of social and organizational level theories in explaining the phenomena. Experts acknowledged that the focus on technology usage is only one piece to understanding the dynamics of how information and communication technologies impact the field of interprofessional collaboration. Additionally, evaluators of health information technology should also assess the impacts of how these interventions affect different levels of the system and organization. Specifically, experts identified the following issues that researchers need to consider. These include examining the processes and how technology interacts with 1) Education, 2) Clinical Practices, 3) Culture, 4) Interprofessional Collaboration and Communication, and 5) Organization of Care.

Need for Multifaceted Outcome Measurements

Experts also recognized the challenges of identifying the appropriate measures and their assessments in the research models. Questions were raised over the types and definitions of outcomes and how to measure them. In the current literature, participants noted that different definitions and benchmarks were often used to measure similar concepts or outcomes, which are problematic for the researchers. For example, the definition of communication failures could be measured as either a) disruptions to the flow of information, b) frequency of communication events, or c) simple references and classifications of communication failures. Thus, suggestions were made that championed evaluation research to apply rigorous scientific approaches and designs.

Yet, the types of methods used to investigate these multifaceted and complex processes can be problematic and difficult for assessments. In particular, it was noted that complex interventions do not only affect one singular outcome but rather, produce multiple consequences on diverse areas including effects on patient safety through the adverse events, interprofessional relationships, and efficiencies of individuals, teams, and organizations. Thus, considering the complexities of the inherent processes, it was recommended that mixed methods consisting of qualitative and quantitative approaches be used to obtain multiple data sources when evaluating complex interventions.

Discussion

Overall, the discussion points raised in our meeting captured the current concerns raised by many implementers and researchers across various fields, disciplines, and professions. Participants acknowledged that there are essential uses for information and communication technologies in the clinical settings. Yet the creation of the ideal clinical communication will require integration of the design and research paradigms.

From the design and implementation perspective, discussions among the meeting's participants revealed that the ideal health care communication system should be one that allows the large volume of patient information to be conveyed and shared at the *appropriate* time to the *right* person. The optimal communication design should be one that allows clinicians to choose *when* and *what* types of information they wish to receive through these communications. Participants also brought up the importance of designing a health care communication system that focuses on patient-centered care. Consequently, a communication system should allow patients to have greater autonomy in managing their health information and create more equality between patients and clinicians. Finally, the system should also be interoperable and scalable to different institutions that allow ease of knowledge transfer.

From the research and evaluation perspective, participants acknowledged that perspectives traditionally used to examine the role of technology in health care communication have been fragmented with researchers working in separate silos. Thus, new forms of integrative thinking and theorizing that span multiple disciplines and fields (eg, social sciences, quality and safety, and information technology) should be considered. Research priorities should consider the impact of communication technology on health care by examining it across different levels and units of analyses ranging from the organizational level (eg, organizational culture and hierarchies), to the team level (eg, workgroups and clinical practices), and finally at the individual level (eg, the clinician's workflow). While the exploration of health care technology from different perspectives is an important step, the tools used to map the paths are equally as important. In order to illuminate insights into the complex and multifaceted processes that exist at multiple levels, participants advocated that research should explore novel methodologies and define theoretical frameworks as ways to further our understanding of clinical communication activities.

Looking from the lenses of interprofessional communication and collaboration, the rising trends in our aging populations and the emergence of chronic diseases also meant that an increasing number of these patients with chronic and complex conditions are being admitted into our hospitals today. There is a call for a communication system that promotes team cohesion and collaboration as hospitals further adopt an interprofessional approach involving different health care practitioners working in different temporal and spatial zones [16,17]. Thus, a strong communication and collaboration structure should be one that not only keeps all clinicians caring for the patient adequately informed but also one that could establish and standardize a common interprofessional language among diverse clinicians [18,19]. Additionally, providing adequate education and gaining buy-in from stakeholders and end users will be critical success factors for implementing this sophisticated technology in the health care system.

To the best of our knowledge, this is the first meeting with this agenda that has been previously organized. Thus, our meeting serves to expand the knowledge base in the fields of information and communication technologies as well as interprofessional collaboration. Through this meeting, we were able to bring different professionals, key stakeholders, and leading experts in the field of health care communication into one room to discuss how technology can serve the needs of today's health care settings and determine ways to evaluate these implementations. We were able to identify the attributes of what the ideal communication system should entail and determine ways in which we can garner interest and partnerships to help develop and implement a sophisticated system successfully—one that not only improves efficiency at the system level but also promotes quality interprofessional collaboration and patient care. Moreover, we have identified the limitations of current knowledge and existing research gaps and proposed recommendations of what health care communication researchers can do to bridge those gaps. Thus, beyond the identification of existing issues, we were able to brainstorm recommendations that health care systems and researchers could apply in their organizations and fields of study alike.

A limitation of the meeting is that physicians who attended the workshop were primarily from the General Internal Medicine specialty who may be accustomed to particular communication cultures and patterns. Another limitation is that the meeting was hosted for only half a day, which may have constrained participants from delving further in their discussions or brainstorm more solutions and recommendations. Nonetheless, all the workshop participants felt that they were given sufficient time to express their opinions and believed the half-day workshop had achieved the intended goals and purposes. Also, given that the aim was exploratory and geared toward hypothesis generation, we believe that the findings provided us with sufficient grounding to make recommendations that will guide future work. While we had excellent interprofessional representation, representation by patients and family members may have provided improved understanding of their roles and on information governance and confidentiality. Finally, as communication systems become more advanced, information contained within these systems may overlap information with an electronic patient record. The meeting did not explore this issue and how this overlap should be managed but should be considered in the future.

Conclusion

While there is a push to adopt technology as the solution to fix specific health care communication problems, collaboration between clinicians is complex, and answers to overcoming health care communication challenges extend beyond the selection of the latest technology. Health care institutions need to consider their strategies and examine how to effectively facilitate and integrate communication technology with different components of the overall system and organization. Caution is required to ensure that technological changes and implementations are adopted carefully to minimize the failures and unintended consequences by considering how interdependencies among multiple parts of the system — people, processes, policies, cultures, and physical infrastructure — influence health care communication and collaboration outcomes.

Acknowledgments

The authors would like to thank the Canadian Institutes of Health Research (CIHR) for their financial support for the Interprofessional Communication and Collaboration Meeting. In addition, the authors would like to thank all the members of the meeting (for a full list of participants please see below) who showed enormous dedication to this initiative through their generous donation of time, thoughtfulness, and energy.

Interprofessional Communication and Collaboration Working Group

Group A: Design and Implementation

Joseph A Cafazzo, Centre for Centre for Global eHealth Innovation, University Health Network

Eric Chow, Centre for Innovation in Complex Care, University Health Network

Sherry Espin, Daphne Cockwell School of Nursing, Ryerson University

Olavo Fernandes, Leslie Dan Faculty of Pharmacy, University of Toronto

Dante Morra, Department of Medicine, University Health Network

Kevin O'Leary, Feinberg School of Medicine, Northwestern University

Peter Rossos, Department of Medicine, University Health Network

Sherman Quan, Shared Information Management Services, University Health Network

Group B: Research and Evaluation

Diane Doran, Lawrence S. Bloomberg Faculty of Nursing, University of Toronto

Joanne Goldman, Keenan Research Centre, Li Ka Shing Knowledge Institute of St Michael's Hospital

Lianne Jeffs, Keenan Research Centre of the Li Ka Shing Knowledge Institute of St. Michael's Hospital

Arun Keepanasseril, eHealth Program, McMaster University

Andre Kushniruk, School of Health Information Science, University of Victoria

Craig Kuziemsky, Telfer School of Management, University of Ottawa

Vivian Lo, Centre for Innovation in Complex Care, University Health Network

Ann McKibbon, eHealth Program, McMaster University

Scott Reeves, Keenan Research Centre of the Li Ka Shing Knowledge Institute of St. Michael's Hospital

Brian Wong, Department of Medicine, Sunnybrook Health Sciences Centre

Robert C. Wu, Department of Medicine, University Health Network

Authors' Contributions

R Wu and D Morra contributed to the conception and design of the study, aided in the analysis and interpretation of the data, and provided critical revision of the article.
V Lo contributed to the collection, analysis, and interpretation of the data and drafted and provided critical revision of the article.
P Rossos, C Kuziemsky, KJ O'Leary, SD Quan, JA Cafazzo, S Reeves and BM Wong aided in the interpretation of the data and provided critical revision of the article.
All authors provided final approval of the version to be published.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Guiding questions and materials used in the Interprofessional Communication and Collaboration Meeting.

[PDF File (Adobe PDF File), 76KB - ijmr_v1i2e9_app1.pdf]

References

1. Coiera E. When conversation is better than computation. J Am Med Inform Assoc 2000;7(3):277-286 [FREE Full text] [Medline: 10833164]
2. Volpp KG, Grande D. Residents' suggestions for reducing errors in teaching hospitals. N Engl J Med 2003 Feb 27;348(9):851-855. [doi: 10.1056/NEJMsb021667] [Medline: 12606742]
3. O'Leary KJ, Liebovitz DM, Baker DW. How hospitalists spend their time: insights on efficiency and safety. J Hosp Med 2006 Mar;1(2):88-93. [doi: 10.1002/jhm.88] [Medline: 17219478]
4. Brennan TA, Leape LL, Laird NM, Hebert L, Localio AR, Lawthers AG, et al. Incidence of adverse events and negligence in hospitalized patients. Results of the Harvard Medical Practice Study I. N Engl J Med 1991 Feb 7;324(6):370-376 [FREE Full text] [doi: 10.1056/NEJM199102073240604] [Medline: 1987460]
5. Harvey R, Jarrett PG, Peltekian KM. Patterns of paging medical interns during night calls at two teaching hospitals. CMAJ 1994 Aug 1;151(3):307-311 [FREE Full text] [Medline: 8039084]
6. Katz MH, Schroeder SA. The sounds of the hospital. Paging patterns in three teaching hospitals. N Engl J Med 1988 Dec 15;319(24):1585-1589. [doi: 10.1056/NEJM198812153192406] [Medline: 3200267]
7. Zinn C. 14,000 preventable deaths in Australian hospitals. BMJ 1995 Jun 10;310(6993):1487. [Medline: 7787586]
8. Joint Commission on Accreditation of Healthcare Organizations. Poor communication is common cause of errors. Healthcare Benchmarks Qual Improv 2002 Aug;1(2):18-19. [Medline: 12224120]
9. Agarwal R, Sands DZ, Schneider JD. Quantifying the economic impact of communication inefficiencies in U.S. hospitals. J Healthc Manag 2010 Jul;55(4):265-281. [Medline: 20812527]
10. Quan S, Wu R, Morra D, Wong BM, Mraz R, Hamill M, et al. Demonstrating the BlackBerry as a clinical communication tool: a pilot study conducted through the Centre for Innovation in Complex Care. Healthc Q 2008;11(4):94-98. [Medline: 19068938]
11. Wu RC, Morra D, Quan S, Lai S, Zanjani S, Abrams H, et al. The use of smartphones for clinical communication on internal medicine wards. J Hosp Med 2010;5(9):553-559. [doi: 10.1002/jhm.775] [Medline: 20690190]
12. Wu R, Rossos P, Quan S, Reeves S, Lo V, Wong B, et al. An evaluation of the use of smartphones to communicate between clinicians: a mixed-methods study. J Med Internet Res 2011;13(3):e59 [FREE Full text] [doi: 10.2196/jmir.1655] [Medline: 21875849]
13. Manhattan Research. 2009. Physician smartphone adoption rate to reach 81% in 2012 URL: http://manhattanresearch.com/News-and-Events/Press-Releases/physician-smartphones-2012 [accessed 2012-09-13] [WebCite Cache ID 6Aegup99F]
14. Wu RC, Tran K, Lo V, O'Leary KJ, Morra D, Quan SD, et al. Effects of clinical communication interventions in hospitals: A systematic review of information and communication technology adoptions for improved communication between clinicians. Int J Med Inform 2012 Jun 22. [doi: 10.1016/j.ijmedinf.2012.05.014] [Medline: 22727613]
15. Lu YC, Xiao Y, Sears A, Jacko JA. A review and a framework of handheld computer adoption in healthcare. Int J Med Inform 2005 Jun;74(5):409-422. [doi: 10.1016/j.ijmedinf.2005.03.001] [Medline: 15893264]
16. Pryor J. A nursing perspective on the relationship between nursing and allied health in inpatient rehabilitation. Disabil Rehabil 2008 Jun;30(4):314-322. [doi: 10.1080/09638280701256900] [Medline: 18210323]
17. Grinspun D. Healthy workplaces: the case for shared clinical decision making and increased full-time employment. Healthc Pap 2007;7 Spec No:85-91; discussion 109 [FREE Full text] [Medline: 17479007]
18. Reeves S, Lewin S, Espin S, Zwarenstein M. Interprofessional Teamwork for Health and Social Care. Oxford: Blackwell-Wiley; 2010.
19. Gair G, Hartery T. Medical dominance in multidisciplinary teamwork: a case study of discharge decision-making in a geriatric assessment unit. J Nurs Manag 2001 Jan;9(1):3-11. [Medline: 11879439]

Edited by G Eysenbach; submitted 12.12.11; peer-reviewed by M Green; comments to author 30.07.12; revised version received 10.08.12; accepted 24.08.12; published 24.09.12.

Please cite as:
Wu RC, Lo V, Rossos P, Kuziemsky C, O'Leary KJ, Cafazzo JA, Reeves S, Wong BM, Morra D
Improving Hospital Care and Collaborative Communications for the 21st Century: Key Recommendations for General Internal Medicine
Interact J Med Res 2012;1(2):e9
URL: http://www.i-jmr.org/2012/2/e9/
doi:10.2196/ijmr.2022
PMID:23612055

Original Paper

Open Access Capture of Patients With Gastroesophageal Reflux Disease Using an Online Patient-Reported Outcomes Instrument

Merel M Tielemans[1,2], MD; Jan BMJ Jansen[3], MD, PhD; Martijn GH van Oijen[2,4,5], PhD

[1]Department of Gastroenterology and Hepatology, Radboud University Medical Center, Nijmegen, Netherlands

[2]Department of Gastroenterology and Hepatology, University Medical Center Utrecht, Utrecht, Netherlands

[3]Department of Gastroenterology and Hepatology, Elkerliek Hospital, Helmond, Netherlands

[4]David Geffen School of Medicine, Division of Digestive Diseases, UCLA, Los Angeles, CA, United States

[5]UCLA/VA Center for Outcomes Research and Education (CORE), Los Angeles, CA, United States

Corresponding Author:
Merel M Tielemans, MD
Department of Gastroenterology and Hepatology
Radboud University Medical Center
PO Box 9101
Internal code 455
Nijmegen, 6500 HB
Netherlands
Phone: 31 24 3616999
Fax: 31 24 3540103
Email: M.Tielemans@MDL.umcn.nl

Abstract

Background: Persons with gastroesophageal reflux disease (GERD) frequently search online for information about causes and treatment options. The GerdQ self-assessment questionnaire can be used for diagnosis of GERD and follow-up of symptoms.

Objectives: To assess whether it is feasible (1) to study the prevalence and impact of GERD in persons visiting a GERD information website, and (2) to identify partial responsiveness to proton pump inhibitor (PPI) therapy using the GerdQ.

Methods: All visitors (aged 18–79 years) to a GERD information website between November 2008 and May 2011 were invited to complete the GerdQ online. The GerdQ questionnaire consists of 6 questions (score per question: 0–3). In respondents who did not use PPIs, we used the questionnaire to identify those with GERD (total score ≥8) and assess the influence of these symptoms on their daily life, divided into low (total score <3 on impact questions) and high impact (total score ≥3 on impact questions). In PPI users, we used the GerdQ to quantify partial responsiveness by any report of heartburn, regurgitation, sleep disturbance, or over-the-counter medication use for more than 1 day in the preceding week. We subsequently asked GerdQ respondents scoring ≥8 to complete the disease-specific Quality of Life in Reflux and Dyspepsia (QOLRAD) questionnaire.

Results: A total of 131,286 visitors completed the GerdQ, of whom 80.23% (n = 105,329) did not use a PPI. Of these, we identified 67,379 respondents (63.97%) to have GERD (n = 32,935; 48.88% high impact). We invited 14,028 non-PPI users to complete the QOLRAD questionnaire, of whom 1231 (8.78%) completed the questionnaire. Mean total QOLRAD scores were 5.14 (SEM 0.04) for those with high-impact GERD and 5.77 (SEM 0.04) for those with low-impact GERD ($P < .001$). In PPI users, 22,826 of 25,957 respondents (87.94%) reported partial responsiveness. We invited 6238 PPI users to complete the QOLRAD questionnaire, of whom 599 (9.60%) completed the disease-specific quality-of-life questionnaire. Mean total QOLRAD scores were 4.62 (SEM 0.05) for partial responders and 5.88 (SEM 0.14) for adequate responders ($P < .001$).

Conclusions: The GerdQ identified GERD in many website respondents and measured partial responsiveness in the majority of PPI users. Both non-PPI users with GERD and PPI users with partial responsiveness were associated with a decreased health-related quality of life. We have shown the feasibility of GERD patient identification online.

(Interact J Med Res 2012;1(2):e7) doi:10.2196/ijmr.2101

KEYWORDS

Gastroesophageal reflux; proton pump inhibitor; Internet; open access questionnaire; partial responsiveness

Introduction

The Internet has gained major influence in the information supply for both physicians and patients in the last decades and has generated new opportunities to study health care and diseases [1-4]. Traditionally, medical literature, treatment guidelines, and patient brochures on gastroesophageal reflux disease (GERD) have been available mainly at the general practitioner's office, and only 5%-30% of patients with GERD consult a general practitioner for their symptoms [5,6]. A recent study found that more than half of online health information seekers searched the Internet without prior medical consultation [4].

GERD is a chronic relapsing and remitting disorder with heartburn and regurgitation as cardinal symptoms. It is associated with a decreased health-related quality of life [7-9]. The prevalence of GERD in Western countries is 10%-20% [5,10] and the disease accounts for 3%-5% of general practitioner visits [11,12]. The main treatment focus is gastric acid suppression, for which proton pump inhibitors (PPIs) are most effective and are proven to be cost effective [13].

The majority of persons with GERD symptoms are underreported in the literature, because prior studies regarding GERD were mainly conducted in primary care [14]. Most persons with GERD symptoms do not visit a primary care physician, which is a potential limitation in the understanding of symptom prevalence and treatment response. A German study assessed gastrointestinal symptoms and quality of life via an Internet questionnaire in 5256 individuals between 2002 and 2005 [15]. This study concluded that the generated data were in general comparable with non-Internet studies, with the exception that the Internet population was younger. Since then, only a few studies have been conducted on the prevalence of a condition in the general population via the Internet. The majority of Internet-based studies invite participants by email, for example, selected by clinicians or Internet panels [16,17], thereby preselecting participants.

The aims of the current study were to assess whether it is feasible to study the prevalence and impact of GERD in persons visiting a GERD information website and to identify partial responsiveness to PPI therapy using the GerdQ self-assessment questionnaire. Symptom scores were compared with a validated health-related quality-of-life instrument. We hypothesized that the prevalence of GERD in our Internet population would be high and that a higher GerdQ score would reflect a lower health-related quality of life.

Methods

Study Population

The website www.maagzuur.nl contains information regarding GERD symptoms, possible causes, lifestyle advice, and treatment and diagnostic options. In May 2008, the Dutch translation of the GerdQ self-assessment questionnaire was launched on this website and could be completed by all website visitors (see Multimedia Appendix 1). After a preparatory period of 6 months, questionnaires completed between November 24, 2008 and May 4, 2011 were included in this study. We excluded respondents younger than 18 and older than 79 years. In the case of duplicate GerdQ questionnaires—defined as having an identical Internet protocol address, birth year, and gender—we included only the first completed GerdQ questionnaire. Respondents who scored ≥8 on the GerdQ were subsequently asked to complete the Quality of Life in Reflux and Dyspepsia (QOLRAD) questionnaire.

The GerdQ Self-Assessment Questionnaire

The GerdQ is a short and validated self-assessment questionnaire that assesses presence of GERD and determines the influence of symptoms on a patient's daily life [18]. The GerdQ comprises six questions reflecting symptoms in the previous 7 days, and has been developed with questions from the Reflux Disease Questionnaire, the Gastrointestinal Symptom Rating Scale, and the Gastrointestinal Symptom Scale, all of which are validated disease-specific questionnaires [19-21]. The GerdQ consists of the following questions referring to the previous week: (1) How often did you have a burning feeling behind your breastbone (heartburn)?, (2) How often did you have stomach contents (liquid or food) moving upward to your throat or mouth (regurgitation)?, (3) How often did you have a pain in the center of the upper stomach?, (4) How often did you have nausea?, (5) How often did you have difficulty getting a good night's sleep because of your heartburn and/or regurgitation?, and (6) How often did you take additional medication for your heartburn and/or regurgitation, other than what the physician told you to take (such as Maalox)?

The first two questions (1 and 2) are positive predictors of GERD, where a higher symptom frequency is indicated by a higher score. Questions 3 and 4 address dyspeptic symptoms that decrease the probability of having GERD—that is, they are negative predictors of GERD. The two final questions (5 and 6) assess the impact of symptoms on a person's daily life and are also positive predictors of GERD. The score on every question ranges from 0 to 3 for the four positive predictors of GERD (0 days is a score of 0; 1 day scores 1; 2–3 days scores 2, and 4–7 days scores 3, or in reversed order for the two negative predictors of GERD). In people who do not use a PPI, a GerdQ score of ≥8 indicates a high probability of having GERD. A cut-off of ≥3 on the GERD-impact questions 5 and 6 indicates a high impact of symptoms on a person's daily life. We defined partial responsiveness in PPI users as more than 1 day of having heartburn (question 1), regurgitation (question 2), sleep disturbance (question 5), or over-the-counter acid suppressive medication use (question 6), all during the preceding week. We also analyzed partial responsiveness using a more stringent definition of persistence of heartburn, regurgitation, sleep disturbances, or over-the-counter medication use for at least 4 days during the preceding week. The questionnaire was shown to respondents together with a figure of a human torso with the breastbone and center of the upper stomach being marked.

QOLRAD Questionnaire

The validated disease-specific QOLRAD questionnaire was developed to monitor health-related quality of life in patients with heartburn and dyspepsia. It contains 25 questions clustered

in five domains: emotional distress, sleep disturbance, food and drink problems, physical and social functioning, and vitality [22,23]. Every question was assessed on a 7-point Likert scale, with a lower score indicating a more severe impact on daily functioning (1 = always, 2 = usually, 3 = frequently, etc, to 7 = never) [24].

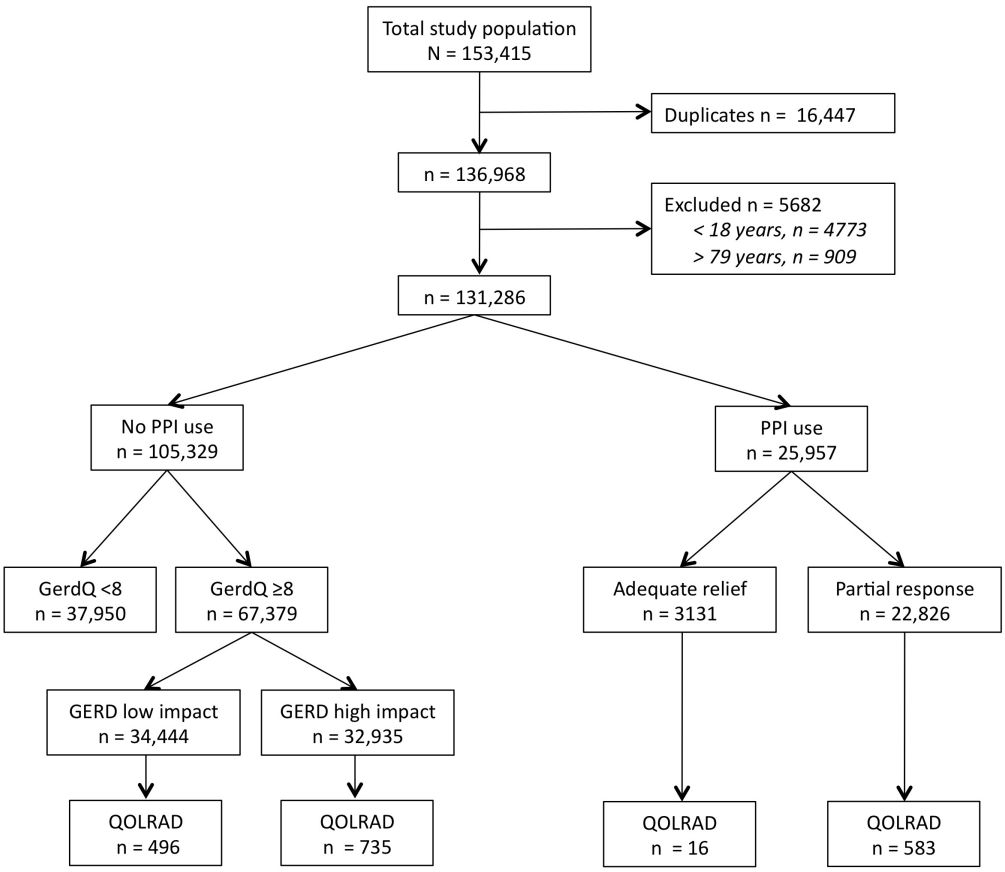
Data Analysis

Questionnaires were stored online in a specially designed website content management system (TripTic bv, Eindhoven, The Netherlands). Data were analyzed using SPSS version 16.0 (IBM Corporation, Somers, NY, USA). We calculated total GerdQ score by summing scores for all of the GerdQ questions. The mean age of respondents with high-impact GERD and low-impact GERD were analyzed using the Student *t* test. The mean age of PPI users with adequate relief and partial responders were also compared by Student *t* test. We compared dichotomous variables, such as gender, by chi-square analysis. Over-the-counter medication use and duration of symptoms were analyzed using descriptive statistics. An overall mean QOLRAD score was calculated by summing scores for all QOLRAD questions, divided by 25 for subgroups of PPI users and non-PPI users. We also calculated a mean score for each domain for respondents with high-impact GERD, low-impact GERD, PPI users with adequate relief, and partial responders to PPI therapy. In respondents with partial responsiveness, we analyzed subgroups of respondents with symptoms persisting at least 4 days per week versus those with less frequent symptoms. Using the Mann-Whitney *U* test, we compared mean scores in each QOLRAD domain between non-PPI users with low impact and those with high impact, and between PPI users with relief and those with partial response. We also compared mean scores in each QOLRAD domain between partial responders with symptoms persisting at least 4 days per week and those with symptoms persisting at most 3 days per week. A *P* value of <.05 was considered statistically significant.

Results

The GerdQ self-assessment questionnaire was completed 153,415 times between November 2008 and May 2011. After removing duplicate entries (n = 16,447) and excluding respondents aged less than 18 years or 80 years and over (n = 5682), we entered 131,286 GerdQ questionnaires into our analysis (Figure 1). A total of 105,329 respondents (80.23%) reported no use of PPIs and 25,957 respondents (19.77%) reported PPI use (Figure 1).

Figure 1. Flow of participants through the study. PPI = proton pump inhibitor. GERD = gastroesophageal reflux disease. QOLRAD = Quality of Life in Reflux and Dyspepsia.



Respondents Without Proton Pump Inhibitor Use

The mean age of the 105,329 respondents who did not use PPIs was 41.6 (SD 14) years, and 49.72% (n = 52,369) were male. A total of 37,950 respondents (36.03%) scored <8 on the GerdQ, indicating a low probability for GERD. The remainder (n = 67,379; 64.0%) scored ≥8, of whom half (n = 32,935; 48.88%) reported GERD with a high impact on the respondent's daily life. Respondents with GERD were older than those without GERD, and the mean age was even higher in respondents with GERD with high impact (Table 1).

Table 1. Baseline characteristics of respondents with and without proton pump inhibitor (PPI) use.

Characteristic	No PPI use			PPI use	
	No GERD[a] (n = 37,950)	Low-impact GERD (n = 34,444)	High-impact GERD (n = 32,935)	Adequate relief (n= 3131)	Partial response[b] (n = 22,826)
Male, n (%)	17,562 (46.28%)	18,035 (52.36%)[c]	16,772 (50.92%)	1,539 (49.15%)[d]	10,132 (44.39%)
Age (years), mean (SD)	39.2 (14)	41.7 (14)[c]	44.3 (14)	49.9 (14)[d]	48.3 (14)
Age categories (years), n (%)					
18–30	12,937 (34.09%)	9346 (27.13%)[c]	6500 (19.74%)	349 (11.15%)[d]	2719 (11.91%)
31–40	7953 (20.96%)	7096 (20.60%)	6721 (20.41%)	437 (13.96%)	3821 (16.74%)
41–50	8157 (21.49%)	8051 (23.37%)	8252 (25.06%)	717 (22.90%)	5787 (25.35%)
51–60	5833 (15.37%)	6237 (18.11%)	7217 (21.91%)	861 (27.50%)	5815 (25.48%)
61–70	2575 (6.79%)	3038 (8.82%)	3527 (10.71%)	603 (19.26%)	3644 (15.96%)
71–79	495 (1.30%)	676 (1.96%)	718 (2.18%)	164 (5.24%)	1040 (4.56%)

[a] Gastroesophageal reflux disease.

[b] Partial response: heartburn, regurgitation, sleep disturbance, or over-the-counter medication use for >1 day during the preceding week.

[c] $P < .001$ comparing low-impact GERD versus high-impact GERD.

[d] $P < .001$ comparing adequate relief versus partial response in PPI users.

Of respondents with low-impact GERD, 61.59% (n = 21,215) took over-the-counter medication less than once per week, compared with 8.64% (n = 2846) of respondents with high-impact GERD (Table 2).

Table 2. Frequency of over-the-counter medication use in respondents with and without proton pump inhibitor (PPI) use.

Frequency (days/week)	No PPI use			PPI use	
	No GERD[a] (n = 37,950)	Low-impact GERD (n = 34,444)	High-impact GERD (n = 32,935)	Adequate relief (n = 3131)	Partial response[b] (n = 22,826)
<1	31,673 (83.46%)	21,215 (61.59%)	2846 (8.64%)	2221 (70.94%)	8352 (36.59%)
1	4086 (10.77%)	9128 (26.50%)	3169 (9.62%)	910 (29.06%)	2195 (9.62%)
2–3	1692 (4.46%)	4101 (11.91%)	13,427 (40.77%)	0 (0%)	4587 (20.10%)
4–7	499 (1.31%)	0 (0%)	13,493 (40.97%)	0 (0%)	7692 (33.70%)

[a] Gastroesophageal reflux disease.

[b] Partial response: heartburn, regurgitation, sleep disturbance, or over-the-counter medication use for >1 day during the preceding week.

In a subset of respondents we inquired about duration of symptoms. Of those with low-impact GERD, 45.6% (n = 554) reported symptom duration of 1 year or less, while 56.3% (n = 930) of those with high-impact GERD reported symptoms for more than 2 years (Table 3).

Table 3. Duration of symptoms in respondents with and without proton pump inhibitor (PPI) use.

Duration (months)	No PPI use		PPI use	
	Low-impact GERD[a] (n = 1215)	High-impact GERD (n = 1652)	Adequate relief (n = 185)	Partial response[b] (n = 1381)
0–6	376 (30.95%)	290 (17.55%)[c]	34 (18.4%)	190 (13.76%)[d]
7–12	178 (14.65%)	213 (12.89%)	14 (7.6%)	123 (8.91%)
13–24	130 (10.70%)	219 (13.26%)	13 (7.0%)	131 (9.49%)
>24	531 (43.70%)	930 (56.30%)	124 (67.0%)	937 (67.85%)

[a] Gastroesophageal reflux disease.

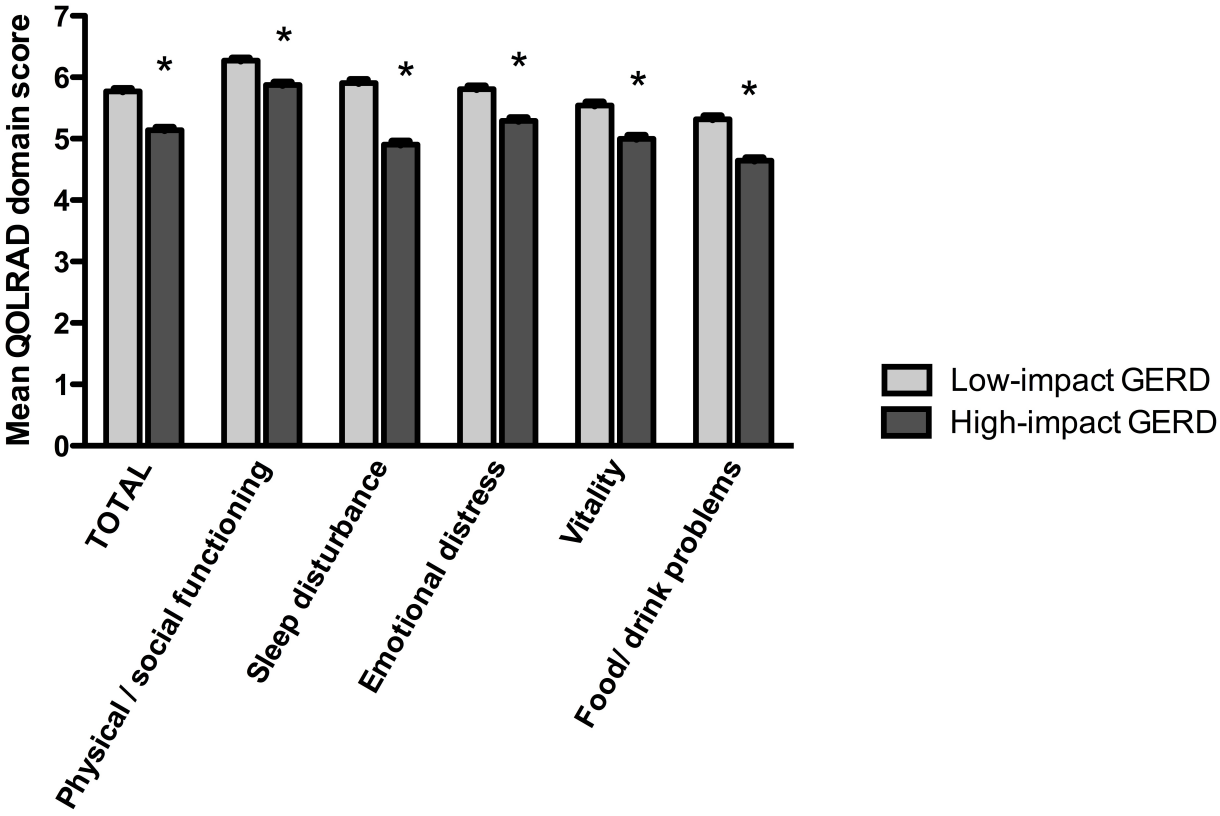
[b] Partial response: heartburn, regurgitation, sleep disturbance, or over-the-counter medication use for >1 day during the preceding week.

[c] $P < .001$ comparing low-impact GERD versus high-impact GERD.

[d] $P = .28$ comparing adequate relief versus partial response.

A total of 14,028 respondents were eligible for (ie, GerdQ score ≥8) and invited to complete the QOLRAD questionnaire, of whom 1231 (8.78%) completed the questionnaire. The total mean QOLRAD score in respondents with GERD with low impact on daily life was 5.77 (SEM 0.04), compared with 5.14 (SEM 0.04) in those with high-impact GERD ($P < .001$; Figure 2). Quality of life was most impaired in the food/drink domain, and the differences in scores between high-impact and low-impact GERD were most pronounced in sleep disturbances and food/drink problems.

Figure 2. Quality of Life in Reflux and Dyspepsia (QOLRAD) scores by domain in respondents with gastroesophageal reflux disease (GERD) who did not use proton pump inhibitors (PPIs). Error bars indicate SEM. *$P < .001$.



Proton Pump Inhibitor Users

The mean age of PPI users was 48.5 (SD 14) years, and 44.96% (n = 11,671) were male. A total of 22,826 PPI users (87.94%) reported having heartburn or regurgitation, sleep disturbances due to GERD symptoms, or intake of over-the-counter acid suppressive medication for more than 1 day per week. We classified these PPI users as partial responders, and this subgroup was younger and had a higher proportion of women (Table 1). Over-the-counter medication use for at least 4 days per week was reported by 33.70% (n = 7692) of PPI users with partial response, whereas the majority of adequate responders (n = 2221, 70.94%) reported over-the-counter acid suppression

medication use of less than once per week (Table 2). After applying a more stringent definition of partial response, of symptoms persisting at least 4 days per week, we obtained a total of 15,975 (61.54%) reporting partial response.

A total of 6238 PPI users were eligible for and invited to complete the QOLRAD questionnaire, of whom 599 (9.60%) completed the disease-specific quality-of-life questionnaire. The total mean QOLRAD score over all domains was 5.88 (SEM 0.14) in PPI users with adequate relief and 4.62 (SEM 0.05) in PPI users with partial response ($P < .001$; Figure 3).

In both groups of PPI users, scores in the vitality and food/drink domains were lowest, with a consistently lower score in those with partial response. The total mean QOLRAD scores in the two subgroups of partial responders were 5.14 (SEM 0.09) for responders with symptoms persisting at most 3 days per week and 4.43 (SEM 0.06) for responders with symptoms persisting at least 4 days per week ($P < .001$ for all domains; Figure 4).

Figure 3. Quality of Life in Reflux and Dyspepsia (QOLRAD) scores by domain in proton pump inhibitor (PPI) users. Error bars indicate SEM. [a] < .001, [b]$P = .003$, [c]$P = .001$, [d]$P = .002$.

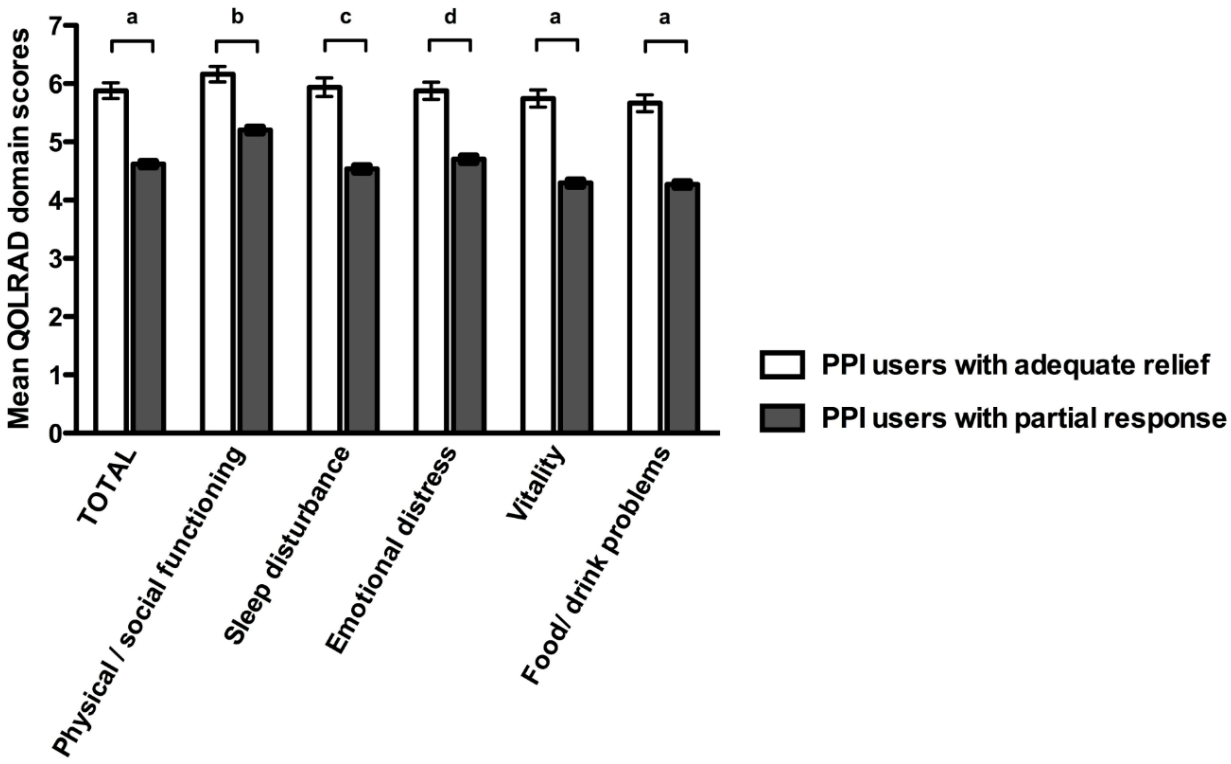
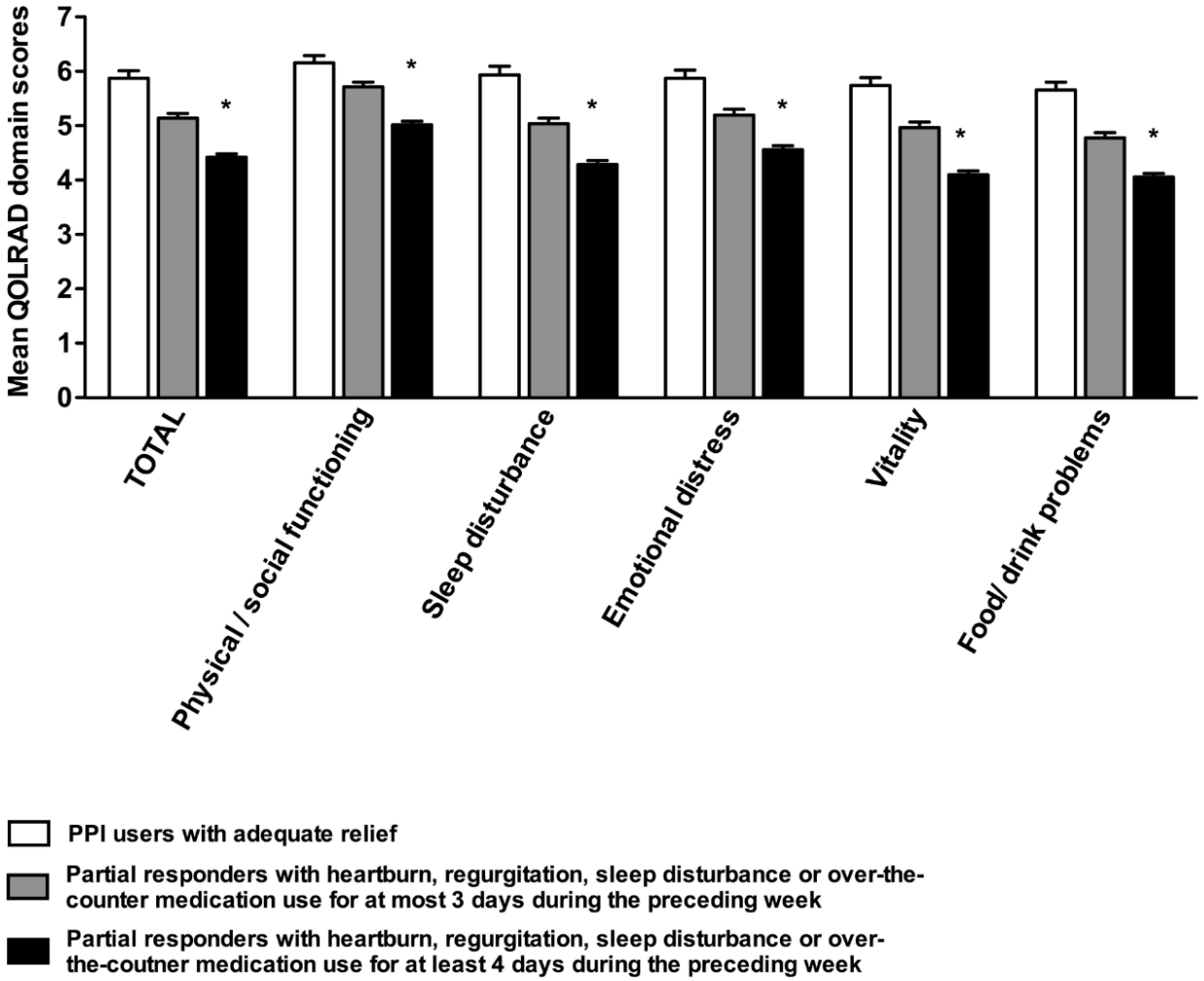


Figure 4. Quality of Life in Reflux and Dyspepsia (QOLRAD) scores by domain in proton pump inhibitor (PPI) users with subdivision of partial responders. *$P < .001$ for comparison between partial responders with symptoms persisting at most 3 days per week and those with symptoms persisting at least 4 days per weeks. Error bars indicate SEM.



Discussion

Principal Results

We found that the prevalence of GERD in website visitors was high, as over 60% of responders without PPI use scored at or above the predefined cut-off on the GerdQ questionnaire. Of the respondents with GERD who did not use a PPI, 49% reported that their symptoms had a great influence on their daily life, in the form of sleep disturbances, and that they needed over-the-counter medications. This was associated with a decreased health-related quality of life. Almost 90% of PPI users reported persistent GERD symptoms for at least 1 day per week. Partial responders taking PPI therapy had a lower health-related quality of life than those who did not use PPIs and those with adequate symptom relief obtained from PPI therapy.

We used the validated self-assessment questionnaire GerdQ to assess the prevalence of GERD among website visitors. Research via the Internet has several advantages and generates new possibilities. As only a minority of patients with GERD visit a health care provider, we can use the Internet to study people who are normally out of the scope of traditional research methods [25]. Another advantage is that missing answers can be directly supplemented during completion of the questionnaire. Data are directly stored electronically, avoiding unreadable handwriting and subsequent mistakes [26]. Data processing via Internet research saves time, especially in studies with many participants. Respondents are able to complete an Internet questionnaire at any time of day, anywhere.

We have shown that it is possible to detect patients with GERD symptoms through a dedicated website. This method can also be used for other conditions. We found that over 150,000 respondents completed the GerdQ questionnaire made accessible online on a health information website, emphasizing the need for disease information on the Internet. However, the skills of the general population to adequately seek health information on the Internet have been shown to be insufficient [27]. These deficiencies varied from problems with opening various common file formats and using hyperlinks embedded in different formats, to problems with appropriately evaluating the information they found [27].

In our study, only 10% of invited respondents completed the QOLRAD questionnaire. We consider the low response rate on completing the QOLRAD questionnaire to be the main drawback of research via an open access questionnaire. Respondents lack face-to-face contact and miss any relationship with the researchers, reducing their willingness to complete a questionnaire without any expected personal gain. A previous study by McCambridge et al assessed the effect of length and relevance of questionnaires on completion rates [28]. They found that only relevance, and not length of the questionnaire, influenced response rate. Another limitation of Internet research is that researchers are unaware of the accuracy of the given information. However, this also applies partly to telephone survey and paper-based questionnaires.

Partial Responsiveness in Proton Pump Inhibitor Users

We used the GerdQ self-assessment questionnaire to identify partial responsiveness in PPI users. This is a novel and very promising feature of the GerdQ. We found that almost 90% of all PPI users had heartburn or regurgitation, sleep problems, or over-the-counter acid suppressive medication use for more than 1 day per week. Of the PPI users, 62% reported persistent symptoms on at least 4 days during the preceding week. Respondents with symptoms persisting at least 4 days per week reported the lowest health-related quality of life in our survey.

A recently published systematic review found that reflux symptoms during PPI therapy persisted in 17%-45% of patients in primary care and the general population [14]. We found a higher proportion of partial responders. This may be due to three independent elements. First, the definitions used in the included articles of the systematic review were not uniform and did not take aspects of quality of life into account. Second, in our study, all website visitors could complete the GerdQ, including those with comorbidity, who are normally excluded from trials. To obtain a maximal treatment effect in clinical trials, respondents with a high risk for decreased efficacy are normally excluded [29]. Third, people with incomplete symptom relief are likelier to search the Internet for more information.

Strengths and Limitations

Our study has several strengths. We included over 130,000 participants in our study, which is the largest population studied for GERD so far [7,8,30]. We used a new, innovative way to collect data. Online data collection can be adequately used in the Netherlands, because more than 85% of Dutch inhabitants already had Internet access in 2008. This is the highest Internet coverage in Europe and would only have increased further during the last 4 years [31]. Using the GerdQ as a promising tool to assess the response of GERD patients to PPI therapy is a novelty. The GerdQ can be used as an easy and quick questionnaire to identify people with an incomplete response. Studies have demonstrated that most physicians presume that PPI therapy is effective in GERD [32]. However, PPIs do not help a significant percentage of patients, which is related to a decreased health-related quality of life [33,34].

Our study also has limitations. First, we have to take selection bias into account. Online health information seekers are probably younger and more educated than are people who search for health information offline [35]. We hypothesize that respondents with more severe symptoms might be overrepresented, as they are likely more motivated to search for information [36]. However, a US survey comparing characteristics of offline and online health information seekers found that online seekers reported a better health status [35]. Another aspect of selection bias in our study is that only a minority of respondents completed the QOLRAD questionnaire. A second limitation is that information regarding comorbidity, medical history, or use of other medications was not available. Third, respondents with suspected GERD symptoms did not undergo endoscopy or pH recording. However, previous research demonstrated that the GerdQ has the same sensitivity and specificity as a gastroenterologist in diagnosing GERD [18].

Implications

The results of our study have some important implications for clinical practice. Many persons searching the Internet for information about reflux have GERD. This generates new opportunities for using the Internet to recognize and treat GERD. It is possible to detect people with GERD and to advise them at first to adjust their lifestyle and take an over-the-counter medication. If these measures are ineffective, these people can be advised to seek medical treatment. People can also regularly complete the GerdQ self-assessment questionnaire via the Internet to assess the effectiveness of their treatment. If they are dissatisfied, they can contact a health care practitioner.

Most PPI users searching the Internet report persistent symptoms or use over-the-counter medication in addition to PPI treatment. General practitioners and gastroenterologists assume that most patients with GERD are adequately treated [32], while our study showed the contrary. Health care providers can now use the GerdQ at every consultation to assess persistent symptoms on PPI therapy and the impact of reflux symptoms on daily life. When necessary, treatment can be adjusted. Further research should investigate the superiority of GerdQ-assisted practice over standard care. The first study to assess incorporation of the GerdQ in daily practice was recently published [37]. It compared the GerdQ with an endoscopy-based approach for diagnosis and initial treatment of GERD, and concluded that using the GerdQ reduced health care costs with comparable efficacy.

We have shown that it is feasible to find patients through a dedicated website for GERD. This concept will also be applicable to other conditions and diseases.

Conclusions

The GerdQ self-assessment questionnaire was completed by over 130,000 website visitors. Two-thirds of respondents who did not use PPIs obtained a score suggestive of GERD. The prevalence of partial responsiveness to PPI therapy was high. Respondents reporting a high impact of GERD had a decreased disease-specific health-related quality of life. Identification of people with GERD through a GERD information website has been shown to be feasible.

Acknowledgments

We thank TripTic bv, Eindhoven, for data storage and provision, and in particular Eefje op den Buysch for her assistance. We thank the editorial board (Jan-Willem Straathof MD PhD, and Hugo Boender MD) and we thank AstraZeneca BV, the Netherlands, for financial support of the website www.maagzuur.nl. The sponsor did not have any influence on the analysis, writing, and conclusions of the article.

Authors' Contributions

MGHvO designed the study. MMT and MGHvO performed statistical analyses. All authors had full access to study data. MMT and MGHvO drafted the first and subsequent versions of the manuscript with input from JBMJ. All authors read and approved the final manuscript.

Conflicts of Interest

MGH van Oijen has received grant support from AstraZeneca and Janssen, and has served as a consultant for AstraZeneca and Pfizer. MM Tielemans and JBMJ Jansen have no conflicts of interest.

Multimedia Appendix 1

The GerdQ self-assessment questionnaire.

[PDF File (Adobe PDF File), 27KB - ijmr_v1i5e7_app1.pdf]

References

1. Fox S. Health information online. Washington, DC: Pew Internet & American Life Project; 2005 May 17. URL: http://www.pewinternet.org/~/media//Files/Reports/2005/PIP_Healthtopics_May05.pdf.pdf [accessed 2012-09-04] [WebCite Cache ID 6AQzcYwCr]
2. Steinbrook R. Searching for the right search--reaching the medical literature. N Engl J Med 2006 Jan 5;354(1):4-7. [doi: 10.1056/NEJMp058128] [Medline: 16394296]
3. Tang H, Ng JH. Googling for a diagnosis--use of Google as a diagnostic aid: internet based study. BMJ 2006 Dec 2;333(7579):1143-1145 [FREE Full text] [doi: 10.1136/bmj.39003.640567.AE] [Medline: 17098763]
4. Powell J, Inglis N, Ronnie J, Large S. The characteristics and motivations of online health information seekers: cross-sectional survey and qualitative interview study. J Med Internet Res 2011;13(1):e20 [FREE Full text] [doi: 10.2196/jmir.1600] [Medline: 21345783]
5. Locke GR, Talley NJ, Fett SL, Zinsmeister AR, Melton LJ. Prevalence and clinical spectrum of gastroesophageal reflux: a population-based study in Olmsted County, Minnesota. Gastroenterology 1997 May;112(5):1448-1456. [Medline: 9136821]
6. Kennedy T, Jones R. The prevalence of gastro-oesophageal reflux symptoms in a UK population and the consultation behaviour of patients with these symptoms. Aliment Pharmacol Ther 2000 Dec;14(12):1589-1594 [FREE Full text] [Medline: 11121906]
7. Kulig M, Leodolter A, Vieth M, Schulte E, Jaspersen D, Labenz J, et al. Quality of life in relation to symptoms in patients with gastro-oesophageal reflux disease-- an analysis based on the ProGERD initiative. Aliment Pharmacol Ther 2003 Oct 15;18(8):767-776 [FREE Full text] [Medline: 14535869]
8. Pace F, Negrini C, Wiklund I, Rossi C, Savarino V, ITALIAN ONE INVESTIGATORS STUDY GROUP. Quality of life in acute and maintenance treatment of non-erosive and mild erosive gastro-oesophageal reflux disease. Aliment Pharmacol Ther 2005 Aug 15;22(4):349-356 [FREE Full text] [doi: 10.1111/j.1365-2036.2005.02558.x] [Medline: 16098002]
9. Wiklund I. Review of the quality of life and burden of illness in gastroesophageal reflux disease. Dig Dis 2004;22(2):108-114. [doi: 10.1159/000080308] [Medline: 15383750]
10. Dent J, El-Serag HB, Wallander MA, Johansson S. Epidemiology of gastro-oesophageal reflux disease: a systematic review. Gut 2005 May;54(5):710-717 [FREE Full text] [doi: 10.1136/gut.2004.051821] [Medline: 15831922]
11. Gisbert JP, Cooper A, Karagiannis D, Hatlebakk J, Agréus L, Jablonowski H, et al. Impact of gastroesophageal reflux disease on patients' daily lives: a European observational study in the primary care setting. Health Qual Life Outcomes 2009;7:60 [FREE Full text] [doi: 10.1186/1477-7525-7-60] [Medline: 19573227]
12. Jones R. Gastro-oesophageal reflux disease in general practice. Scand J Gastroenterol Suppl 1995;211:35-38. [Medline: 8545629]
13. Moayyedi P, Armstrong D, Hunt RH, Lei Y, Bukoski M, White RJ. The gain in quality-adjusted life months by switching to esomeprazole in those with continued reflux symptoms in primary care: EncomPASS--a cluster-randomized trial. Am J Gastroenterol 2010 Nov;105(11):2341-2346. [doi: 10.1038/ajg.2010.368] [Medline: 20842110]

14. El-Serag H, Becher A, Jones R. Systematic review: persistent reflux symptoms on proton pump inhibitor therapy in primary care and community studies. Aliment Pharmacol Ther 2010 Sep;32(6):720-737. [doi: 10.1111/j.1365-2036.2010.04406.x] [Medline: 20662774]
15. Enck P, Kowalski A, Martens U, Klosterhalfen S. Internet-based assessment of bowel symptoms and quality of life. Eur J Gastroenterol Hepatol 2006 Dec;18(12):1263-1269. [doi: 10.1097/01.meg.0000243879.25067.32] [Medline: 17099374]
16. de Wit M, Snoek FJ. Depressive symptoms and unmet psychological needs of Dutch youth with type 1 diabetes: results of a web-survey. Pediatr Diabetes 2011 May;12(3 Pt 1):172-176. [doi: 10.1111/j.1399-5448.2010.00673.x] [Medline: 20561242]
17. Coyne KS, Sexton CC, Thompson CL, Milsom I, Irwin D, Kopp ZS, et al. The prevalence of lower urinary tract symptoms (LUTS) in the USA, the UK and Sweden: results from the Epidemiology of LUTS (EpiLUTS) study. BJU Int 2009 Aug;104(3):352-360. [doi: 10.1111/j.1464-410X.2009.08427.x] [Medline: 19281467]
18. Jones R, Junghard O, Dent J, Vakil N, Halling K, Wernersson B, et al. Development of the GerdQ, a tool for the diagnosis and management of gastro-oesophageal reflux disease in primary care. Aliment Pharmacol Ther 2009 Nov 15;30(10):1030-1038. [doi: 10.1111/j.1365-2036.2009.04142.x] [Medline: 19737151]
19. Jones R, Coyne K, Wiklund I. The gastro-oesophageal reflux disease impact scale: a patient management tool for primary care. Aliment Pharmacol Ther 2007 Jun 15;25(12):1451-1459 [FREE Full text] [doi: 10.1111/j.1365-2036.2007.03343.x] [Medline: 17539985]
20. Revicki DA, Wood M, Wiklund I, Crawley J. Reliability and validity of the Gastrointestinal Symptom Rating Scale in patients with gastroesophageal reflux disease. Qual Life Res 1998 Jan;7(1):75-83. [Medline: 9481153]
21. Shaw M, Dent J, Beebe T, Junghard O, Wiklund I, Lind T, et al. The Reflux Disease Questionnaire: a measure for assessment of treatment response in clinical trials. Health Qual Life Outcomes 2008;6:31 [FREE Full text] [doi: 10.1186/1477-7525-6-31] [Medline: 18447946]
22. Engels LG, Klinkenberg-Knol EC, Carlsson J, Halling K. Psychometric validation of the Dutch translation of the quality of life in reflux and dyspepsia (QOLRAD) questionnaire in patients with gastroesophageal reflux disease. Health Qual Life Outcomes 2010;8:85 [FREE Full text] [doi: 10.1186/1477-7525-8-85] [Medline: 20716328]
23. Wiklund IK, Junghard O, Grace E, Talley NJ, Kamm M, Veldhuyzen van Zanten S, et al. Quality of Life in Reflux and Dyspepsia patients. Psychometric documentation of a new disease-specific questionnaire (QOLRAD). Eur J Surg Suppl 1998(583):41-49. [Medline: 10027672]
24. Likert R. A technique for the measurement of attitudes. Arch Psychol 1932;22(140):1-55.
25. Holmes S. Methodological and ethical considerations in designing an Internet study of quality of life: a discussion paper. Int J Nurs Stud 2009 Mar;46(3):394-405. [doi: 10.1016/j.ijnurstu.2008.08.004] [Medline: 18838135]
26. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. J Med Internet Res 2002 Nov;4(2):E13 [FREE Full text] [doi: 10.2196/jmir.4.2.e13] [Medline: 12554560]
27. van Deursen AJ, van Dijk JA. Internet skills performance tests: are people ready for eHealth? J Med Internet Res 2011;13(2):e35 [FREE Full text] [doi: 10.2196/jmir.1581] [Medline: 21531690]
28. McCambridge J, Kalaitzaki E, White IR, Khadjesari Z, Murray E, Linke S, et al. Impact of length or relevance of questionnaires on attrition in online trials: randomized controlled trial. J Med Internet Res 2011;13(4):e96 [FREE Full text] [doi: 10.2196/jmir.1733] [Medline: 22100793]
29. Brass EP. The gap between clinical trials and clinical practice: the use of pragmatic clinical trials to inform regulatory decision making. Clin Pharmacol Ther 2010 Mar;87(3):351-355. [doi: 10.1038/clpt.2009.218] [Medline: 20010554]
30. Crawley JA, Schmitt CM. How satisfied are chronic heartburn sufferers with their prescription medications? Results of the patient unmet needs survey. J Clin Outcomes Manag 2000;7:29-34.
31. Hoksbergen M. Statistics Netherlands Web Magazine. 2008 Dec 18. No Home Internet Access for 1.2 Million Dutch People URL: http://www.cbs.nl/en-GB/menu/themas/vrije-tijd-cultuur/publicaties/artikelen/archief/2008/2008-2641-wm.htm [accessed 2012-02-12] [WebCite Cache ID 65P6Nw8Ua]
32. Bretagne JF, Honnorat C, Richard-Molard B, Soufflet C, Barthélemy P. Perceptions and practices on the management of gastro-oesophageal reflux disease: results of a national survey comparing primary care physicians and gastroenterologists. Aliment Pharmacol Ther 2007 Apr 1;25(7):823-833 [FREE Full text] [doi: 10.1111/j.1365-2036.2007.03265.x] [Medline: 17373921]
33. Fass R, Sifrim D. Management of heartburn not responding to proton pump inhibitors. Gut 2009 Feb;58(2):295-309. [doi: 10.1136/gut.2007.145581] [Medline: 19136523]
34. Becher A, El-Serag H. Systematic review: the association between symptomatic response to proton pump inhibitors and health-related quality of life in patients with gastro-oesophageal reflux disease. Aliment Pharmacol Ther 2011 Sep;34(6):618-627. [doi: 10.1111/j.1365-2036.2011.04774.x] [Medline: 21770991]
35. Cotten SR, Gupta SS. Characteristics of online and offline health information seekers and factors that discriminate between them. Soc Sci Med 2004 Nov;59(9):1795-1806. [doi: 10.1016/j.socscimed.2004.02.020] [Medline: 15312915]
36. Klovning A, Sandvik H, Hunskaar S. Web-based survey attracted age-biased sample with more severe illness than paper-based survey. J Clin Epidemiol 2009 Oct;62(10):1068-1074. [doi: 10.1016/j.jclinepi.2008.10.015] [Medline: 19246177]

37. Jonasson C, Moum B, Bang C, Andersen KR, Hatlebakk JG. Randomised clinical trial: a comparison between a GerdQ-based algorithm and an endoscopy-based approach for the diagnosis and initial treatment of GERD. Aliment Pharmacol Ther 2012 Jun;35(11):1290-1300. [doi: 10.1111/j.1365-2036.2012.05092.x] [Medline: 22510027]

Abbreviations

GERD: gastroesophageal reflux disease
PPI: proton pump inhibitor
QOLRAD: Quality of Life in Reflux and Dyspepsia

Edited by G Eysenbach; submitted 06.03.12; peer-reviewed by J Gisbert; accepted 30.07.12; published 26.09.12.

Please cite as:
Tielemans MM, Jansen JBMJ, van Oijen MGH
Open Access Capture of Patients With Gastroesophageal Reflux Disease Using an Online Patient-Reported Outcomes Instrument
Interact J Med Res 2012;1(2):e7
URL: http://www.jmir.org/2012/2/e7/
doi:10.2196/ijmr.2101
PMID:23611985

Original Paper

Standardization of Questions in Rare Disease Registries: The PRISM Library Project

Rachel Lynn Richesson[1], PhD, MPH; Denise Shereff[2], MLIS; James Everett Andrews[3], PhD

[1]Duke University School of Nursing, Durham, NC, United States

[2]University of South Florida (USF) College of Medicine, Division of Bioinformatics and Biostatistics, Tampa, FL, United States

[3]School of Information, University of South Florida, Tampa, FL, United States

Corresponding Author:
Rachel Lynn Richesson, PhD, MPH
Duke University School of Nursing
311 Trent Drive
Durham, NC, 27710
United States
Phone: 1 919 681 0825
Fax: 1 919 668 2335
Email: rachel.richesson@dm.duke.edu

Abstract

Background: Patient registries are often a helpful first step in estimating the impact and understanding the etiology of rare diseases - both requisites for the development of new diagnostics and therapeutics. The value and utility of patient registries rely on the use of both well-constructed structured research questions and relevant answer sets accompanying them. There are currently no clear standards or specifications for developing registry questions, and there are no banks of existing questions to support registry developers.

Objective: This paper introduces the [Rare Disease] PRISM (Patient Registry Item Specifications and Metadata for Rare Disease) project, a library of standardized questions covering a broad spectrum of rare diseases that can be used to support the development of new registries, including Internet-based registries.

Methods: A convenience sample of questions was identified from well-established (>5 years) natural history studies in various diseases and from several existing registries. Face validity of the questions was determined by review by many experts (both terminology experts at the College of American Pathologists (CAP) and research and informatics experts at the University of South Florida (USF)) for commonality, clarity, and organization. Questions were re-worded slightly, as needed, to make the full semantics of the question clear and to make the questions generalizable to multiple diseases where possible. Questions were indexed with metadata (structured and descriptive information) using a standard metadata framework to record such information as context, format, question asker and responder, and data standards information.

Results: At present, PRISM contains over 2,200 questions, with content of PRISM relevant to virtually all rare diseases. While the inclusion of disease-specific questions for thousands of rare disease organizations seeking to develop registries would present a challenge for traditional standards development organizations, the PRISM library could serve as a platform to liaison between rare disease communities and existing standardized controlled terminologies, item banks, and coding systems.

Conclusions: If widely used, PRISM will enable the re-use of questions across registries, reduce variation in registry data collection, and facilitate a bottom-up standardization of patient registries. Although it was initially developed to fulfill an urgent need in the rare disease community for shared resources, the PRISM library of patient-directed registry questions can be a valuable resource for registries in any disease – whether common or rare.

Trial Registration: N/A

(Interact J Med Res 2012;1(2):e10) doi:10.2196/ijmr.2107

KEYWORDS

Patient registries; data standards; rare diseases; metadata

Introduction

The value and utility of patient registries are largely contingent upon the use of both high-quality research questions and structured and relevant answer sets accompanying them. The interoperability of registries or registry data—including the feeding of registry data into more rigorous clinical studies and regulatory submissions for new agents—depends upon the use of data standards; yet, there is currently no clear specification for developing registry questions nor are there banks of existing questions to support registry developers. The diverse nature of registries, sponsors, and disease-specific data requirements complicate efforts at standardization in registry applications. This paper introduces the *[Rare Disease] PRISM* (Patient Registry Item Specifications and Metadata for Rare Disease) project, a library of standardized questions covering a broad spectrum of rare diseases that can be used to support the development of new registries in all disease areas. If widely used, PRISM will enable the re-use of questions across registries, hence reducing the variation in registry data collection and facilitating a bottom-up standardization of patient registries. Although it was developed to support the rare disease community's urgent need for shared resources, the PRISM library of patient-directed registry questions can be a valuable resource to registries for all diseases.

Background

Rare Diseases Registries

Often patient registries are a helpful first step in estimating the impact and understanding the etiology of rare diseases—both requisites for the development of new diagnostics and therapeutics. Because of the small numbers of patients affected by rare diseases, these registries present unique challenges related to registry design, enrollment of patients, and data collection [1-5], as well as for analysis and interpretation of registry data [5,6]. In recent years, the U.S. Food and Drug Administration (FDA) and umbrella patient advocacy support organizations (eg, the Genetic Alliance) have publicly encouraged rare disease patient advocacy groups (PAGs) to develop registries as part of a comprehensive strategic research plan. As such, a proliferation of patient registries is underway, and there are currently no dedicated or centralized efforts for developing standards that will reduce the time required to develop new patient registries and facilitate opportunities for shared data.

Patient Registry Variation and Standards

There is tremendous variability in the type of data and specific questions that patient registries collect, due in part to the lack of registry-specific data standards and also to the heterogeneity of registries' purposes and sponsors. Patient registries can be designed for many purposes, including public health surveillance, epidemiologic and longitudinal research, patient education, research recruitment, and population monitoring for the safety of post-marketed drugs and devices. Patient registries can include data reported by patients, researchers, or clinicians. (A characterization of registry types is summarized in Richesson and Vehik [2].) Sponsors and developers of patient registries are varied as well and include governments, academic scientists, and clinical investigators. Often patient registries (and the supporting questions, typically targeted to patients and caregivers rather than physicians) are developed ad hoc by the PAGs themselves, and there is currently no clear specification for standards or banks of existing questions for them to access. Further, the content of these registries (ie, the questions and associated answer sets) may change over time as more becomes known about the disease and its clinical variations, or when new therapeutics and devices become available. There is a tremendous need—especially for the thousands of rare diseases that do not have patient registries—for resources that help registry sponsors and developers to identify well-constructed and meaningful questions.

There is also a clear role for data standards to promote shared efficiencies in registry development and enable opportunities for data sharing. These needs are particularly pronounced for rare diseases, which have sparse resources and significantly fewer—and highly distributed—domain experts and affected patients. An important standards challenge is the fact that there is no central control of patient registries—there is no single funding or regulatory agency that can oversee all the different registry types and implementations. Because registries are developed by many sponsors to address distinct functions, a top-down standards effort would require countless stakeholders and is not feasible. Additionally, there is no central authority to monitor or enforce standards compliance once developed. Given the tremendous need for standards and the scope of data collected across disease-specific registries, and given that there is no incentive or regulatory means to develop standards or enforce compliance, alternatives to complement traditional Standards Development Organizations (SDOs) are needed. Non-traditional strategies for developing and promoting standards can be effective and embraced across various rare diseases if these various research communities perceive them as accessible, useful, helpful, and easy to adopt.

The PRISM project (funded by an American Recovery and Reinvestment Act (ARRA) grant administered through the National Library of Medicine (NIH), NLM Grant Number 1RC1LM010455-01, and supported by the Office of Rare Diseases Research) was developed to provide a useful resource to promote the efficient development of patient registries and standardized quality data collection by supporting the sharing and re-use of existing registry questions and data standards. The fundamental idea behind PRISM is that if registry developers could access questions used by other rare disease registries, they could consider and likely re-use these questions, thereby reducing the variation in questions/data collection across various patient registries, and leading to a *bottom-up* development of standards. In addition, the utility and scalability of PRISM is based on the notion that the PRISM content is accessible and open to any registry sponsor, regardless of prior standards knowledge or experience. The authors are engaged with various standards and informatics organizations and deliberately designed PRISM to facilitate linkage with other standards and research resources as appropriate. The PRISM project, therefore, provides not only a library of questions but a vehicle for the registry developers and rare disease organizations to interact,

learn, and develop consensus requirements, which can in turn be directed to various standards development organizations (eg, CDISC [7], HL7 [8], LOINC [9]), research initiatives (eg, caDSR [10] and CSHARE, PROMIS [11,12], PhenX [13]), and national interoperability initiatives emerging from the U.S. DHHS, Office of the National Coordinator [14,15].

As a demonstration project, PRISM explored foundational issues related to the types of questions included in the bank (relative to other standards and question repositories) and the inclusion of metadata that will facilitate their search and retrieval. We describe our methodological approach to developing the PRISM library in the Methods section and present the resulting library structure, features, and composition afterward in the Results section.

Methods

The first questions identified for PRISM included a convenience sample of questions from well-established (>5 years) natural history studies in various diseases (metabolic [16], vascular [17], developmental disorders [18]) and several existing registries. The questions were obtained largely from the NIH-funded Rare Diseases Clinical Research Network as the authors worked in the data center for the network and had familiarity with the various studies and investigators. These were examined by many experts (both terminology experts at the College of American Pathologists (CAP) and research and informatics experts at the University of South Florida (USF)) for face validity, including commonality, clarity, and organization. Questions were re-worded slightly, as needed, to make the full semantics of the question clear and to make the question generalizable to multiple diseases as possible. For example, a registry data entry item “Genetic test?” would be entered into PRISM as “Have you had a genetic test to confirm your diagnosis?”, representing the intended semantics of the question in a generalizable way, rather than “Have you had a genetic test to confirm your Rett Syndrome diagnosis?”, which was the intent of the question from the source document. In addition, some registry questions were taken from a variety of established rare disease patient registries that authors were acquainted with. To ensure that questions were “stable” and field-tested, authors used well-established content that had been considered final by a PAG after a multi-disciplinary review process and pilot testing. Generic content relative to many or all patient registries was also incorporated, including standard elements from the Rare Diseases Clinical Research Network (RDCRN) Contact Registry (supporting over 200 rare diseases) [19], the OMB and NIH demographics [20], selected questions from the NINDS Common Data Elements [21], and all of the data elements recommended for the Global Registry for Rare Diseases [22], an initiative led by the U.S. Office of Rare Diseases Research. A typology was developed to organize the questions in PRISM and to support internal curation of the library. Questions were indexed by one or more keywords that describe the general content category (eg, demographic, medications, medical history, special histories, etc.). The specific representations of the metadata keywords include common forms (eg, Demographics, Medical History) and form headings (eg, Medication History, Special Education Services) that group related data (by type or by source) for registries and observational studies.

We describe our selected strategy and design features for PRISM in the next section in terms of content, search and retrieval requirements, indexing model and metadata, and strategy for growing the content of the PRISM library. The authors thoroughly explored other standards throughout the design of PRISM and present the relationships and definitions between PRISM and other efforts related to standardized questions and patient reported data. Finally, in the discussion section, we describe immediate future directions for PRISM, including requirements for sharing the library, interface design, and future plans for maintenance and governance of PRISM.

Results

Content

At present, PRISM contains over 2,200 questions. A sample of 224 questions and selected metadata is presented in Appendix 1. Many questions (especially the most general, such as “List current medications you are taking” and “List any other major diseases you have had”) are relevant to virtually all rare diseases, and others are relevant to a great many rare diseases (eg, “Do you require an assistive device for walking?”, or “Was your child born full-term?”). However, the majority of PRISM questions—such as “Does your child hoard food?”, “How many times per week does child pick at own skin until it bleeds?”, “Does your child need TV to fall asleep?”, and “Approximately how many bone fractures have you had in your lifetime?”—are relevant only to particular diseases, and often the valid answer sets associated with each question are disease-specific. For example, valid answers to the question “What age was your child’s first bone fracture?” would include prenatal ages in a patient registry targeted to Osteogenesis Imperfecta, thereby requiring different units and ranges of expected values. The obvious challenge that thousands of such disease-specific questions (relevant to the thousands of rare diseases seeking to develop registries) will present for traditional standards development organizations is what drove the design decision for the scope of PRISM. The scope of PRISM, therefore, deliberately includes a range of disease-specific content, along with narrative definitions and metadata for indexing and source preservation. The metadata was selected to facilitate linkages between rare disease communities and existing SDOs and controlled terminologies, as described in the following sections.

Related Efforts

To prevent overlapping with other standards efforts related to the collection of clinical data, a deliberate search for relevant standardization efforts was undertaken by the authors. This search of existing standards and informatics and library resources revealed several related and potentially relevant efforts, and informed the design of PRISM to leverage related efforts. As described in the introduction, the focus of PRISM is on *patient-reported* questions that are not a part of standardized, validated patient assessment instruments. Given that focus, several organizations are engaged in various attempts to inventory and codify standardized assessment instruments and items, specifically Clinical LOINC [23], PROMIS, the

caDSR and NINDS Common Data Elements [24], and to some extent SNOMED CT. Because PRISM was developed with the intent to leverage and coordinate existing standards as much as possible, the PRISM scope was clearly defined at the onset to prevent any overlap. However, some areas of potential overlap with other initiatives exist. Table 1 and Table 2 describe the potential relationship of PRISM to other initiatives that are also developing repositories of questions or data elements.

Table 1. Relationship of PRISM to Related Standards Efforts and Resources (LOINC, caDSR, and PhenX).

Initiative[a]	Primary Sponsor	Objective	Scope of standard	Proposed relationship with PRISM
Clinical LOINC http://www.webcitation.org/6BJJ3JZIm	NLM	Messaging and interoperability of clinical information. Specifically, the LOINC database provides a set of universal names and ID codes for identifying laboratory and clinical test results.	Health care (primarily) and research	Patient assessment scales are not generally included in PRISM. PRISM documentation directs users to LOINC for this content. [Note: Clinical LOINC does not contain every assessment scale ever published. Registry developers need to search Clinical LOINC or RELMA.] LOINC is interested in supporting electronic health record (EHR) data standards. The rare disease community should leverage LOINC to support transfer of EHR data to registries. PRISM could coordinate this. The majority of PRISM content is variable and might not be messaged or collected in EHRs. But, some PRISM items can be submitted to LOINC if appropriate. PRISM will filter and act as a feeder of selected content into LOINC; PRISM definitions and metadata will aid in this.
caDSR (CSHARE) http://www.webcitation.org/6BJJ5fd1Y	NCI	Research data elements. "caDSR is a database and a set of APIs and tools to create, edit, control, deploy, and find common data elements (CDEs) …for use in software development." (CSHARE is expected to emerge as an expanded pan-disease version of the caDSR but is not yet available.)	Data elements for collection in clinical research studies. NCI's vocab services and metadata repositories do support diseases other than cancer, and have standard data elements from FDA and NIH institutes and centers.	caDSR content and tools are targeted to clinical researchers. Much of caDSR content could be relevant to PRISM and rare disease registry developers, but is not complete nor easily searchable by rare disease users. PRISM includes a focus on registries and rare diseases and a community forum for rare disease registry standards. PRISM maintains a link to source for all content. Users can see explicit links to caDSR if that is the source. PRISM questions could be imported into caDSR for use in clinical research projects. (The PRISM metadata has enough detail to support their transfer if caDSR curators want.)
PhenX http://www.webcitation.org/6BJJ8xNzU	National Human Genome Research Institute (NHGRI)	To provide investigators with high-quality, relatively low-burden measures for inclusion in genome-wide association studies (GWAS) and other large-scale research efforts.	Data elements used in new research data collection, or used/queried from various electronic health records. Content focuses on common diseases but is growing.	Much of PRISM content is very disease specific, often idiosyncharic, and not included in PhenX. PhenX is consulted as a resource for generic questions, which were incorporated into PRISM when authors thought appropriate (underlying PRISM database includes PhenX code in these cases). Where content overlap exists, PRISM points users to PhenX measures. PhenX is aimed at researchers, but user interface is intuitive and easily accessible to PRISM users.

[a] These initiatives are not specific to patient registries or rare diseases.

Table 2. Relationship of PRISM to Related Standards Efforts and Resources (PROMIS, SNOMED CT, RxNorm).

Initiative[a]	Primary Sponsor	Objective	Scope of Standard	Proposed relationship with PRISM
PROMIS (Patient Reported Outcomes Measurement Information System) http://www.webcitation.org/6BJIvffJ8	NIH	A system of highly reliable, precise measures of patient–reported health status for physical, mental, and social well-being.	Functional and quality-of-life assessment questions. Validated measures only; focus on psychosocial constructs across domains, not only specific diseases.	PRISM explicitly avoids content that is validated and intended to be used for measurement. Where there are common questions, PROMIS and PRISM will cross-reference each other. PRISM directs users to PROMIS and describes its potential application in patient registries.
SNOMED CT (SCT) http://www.webcitation.org/6BJIyeJ5C	Int'l. standards development organisation (IHTSDO) [Supported by dues from member nations] US residents may use SNOMED CT free of charge, supported by NLM.	Provides a consistent way to index, store, retrieve, and aggregate clinical data across specialties and sites of care.	Comprehen-sive clinical terminology covering nursing and medical diagnoses, signs and symptoms, functional status, interventions, procedures, and outcomes.	SCT is used for indexing in PRISM. Each PRISM question associated with one or more codes that best represent the important content of the PRISM QAS. The most specific SCT code is used, with the understanding that only some PRISM questions get very precise representation in SCT. Similarly, multiple SCT codes can be used to index the clinical content of multi-concept questions. [In 2006, the SCT nursing working group (prior to the IHTSDO) developed a model for coding assessment scales to SCT, which was approved by the SCT International Editorial Board. So, SCT should be added to the "specifically" list. The Observable Entity-Answer approach to coding PRISM follows the SCT model for coding assessments.]
RxNorm	NLM	RxNorm contains the names of prescription and many non-prescription formulations in the US; aims to support electronic exchange of medication information and clinical decision support related to CPOE in health care contexts.	Standardized nomenclature for clinical drugs and drug delivery devices (mostly in US); gives normalized names for clinical drugs and links its names to many drug vocabularies used in pharmacy management and drug interaction software.	RxNorm does not represent data elements but is a nomenclature for clinical drugs. Many registries ask questions about specific medications. RxNorm is used for indexing questions about medications in PRISM. Each medication-related question in PRISM is associated with RxNorm codes that best represent the named drug – either by clinical drug name, (generic) ingredient, or packaged products.

[a] These initiatives are not specific to patient registries or rare diseases.

Some data elements, like height and weight are in clinical data element repositories such as Clinical LOINC and research data element registries like caDSR. To enable "one-stop shopping" for registry developers, these items are also in PRISM, with the idea that future PRISM interfaces can identify linkages to these other standards where and when appropriate. These linkages can inform PRISM users that they indeed are using items from another designated standard and can also inform PRISM curators to ensure that they do not create or support future variations in that item.

In the interest of rapidly assembling content relevant to patient registries for any and all rare diseases, the initial PRISM strategy has been to accept virtually all questions, with the idea that either patient communities or curators might later filter, rate, or rank them for PRISM users. Despite limiting the inclusion of questions to those from well-established registries, some questions used in rare disease registries were poorly constructed or not as clear as they could be. Because PRISM is motivated to address the registry question needs for a spectrum of registry designs and diseases, it does contain registry data elements (in actual use) that might not be ideally constructed or might actually conflict with another value set. Regardless, the liberal acceptance policy of PRISM increases the breadth and volume of questions and ultimately increases the value of PRISM as a central resource for questions (which should be supplemented by advice on selection and use). We hope that others use PRISM as a resource for standards development or build applications that can facilitate the ranking or endorsement of certain PRISM questions over others in specific diseases or contexts.

Search and Retrieval Requirements

As mentioned in the previous section, the growth of the library brought with it challenges for curation and use. Internally, a process was developed to add new content without duplicating questions and ensure that related questions or variants could be indexed for effective retrieval and comparison. We operated under the assumption that PRISM needed to be useful to provide value. We explored user roles and searching techniques to determine the best method for indexing question and answer sets (QAS). Our indexing scheme (including the use of

controlled terminology) is described later. A key strategy for identifying the search and retrieval requirements was employing use cases.

The PRISM team developed narrative *use cases* (for question searching, re-use, and new registry development) that have informed our indexing approaches. For our indexing strategy, PRISM also developed use cases of retrieval requirements as a guide of needed functionality and detail, and leveraged existing metadata and controlled vocabularies. The use cases describe the development of new registries using PRISM as a resource for the questions. Specifically, the use cases describe the development of a prospective vasculitis and pregnancy registry—including the re-use of PRISM questions and the submission of new questions to PRISM—and the subsequent development of another registry using these same pregnancy-related questions. Finally, the use cases articulate a demonstration and initial proof of the interoperability of data collected from different registries and data sources. These functional use cases ultimately will also support the evaluation of PRISM. The use cases were developed by the authors and not widely vetted in the rare disease community. They can be found on PRISM website [25].

Indexing Model and Metadata

A critical and largely unaddressed problem for registries (and clinical research data collection in general) is the need for patient registry questions and answers to be indexed in such a way that they can be retrieved for re-use, for example, to support rapid development of another related rare disease registry. In essence, indexing is the practice of applying metadata (structured and descriptive information) to items in a database for efficient and accurate retrieval [26]. An important measure of indexing is specificity, which refers to the detail or precision of the indexing process and its depth [26(p161)]. Achieving the appropriate level of specificity in encoding the question and answer sets was a continuous challenge. Our approach was guided by the ideal situation where the pairs would be encoded to convey the semantic meaning of the clinical concepts and at the same time allow efficient collocation. The specificity and overall indexing approach informed the detailed use cases discussed earlier. The indexing scheme also addresses semantic ambiguities inherent in research questions and answer sets, and includes metadata for the following constructs:

1. Context (type of study, disease or treatment of interest, etc.)

2. Format of questions and location of semantics

3. Who is asking the question (patient, relative, doctor)?

4. Audience or person being asked the question (patient, family member or caregiver)

5. Relevant data standards for specific answer sets

Metadata and Controlled Vocabularies

Metadata are used to describe information resource-type features of questions, such as terms and attributes, and controlled vocabularies represent the actual content. Both are important for retrieval and, ultimately, interoperability. PRISM uses Dublin Core (DC) as a metadata framework for indexing PRISM QAS, and within DC metadata uses controlled terminology to reflect the semantics for each question. Our approach is described in detail in [27]. Terminology control, when implemented correctly and consistently, can dramatically improve the quality of search results in most contexts. SNOMED CT is an ONC recommended standard for many aspects of the electronic medical record, and previous research has indicated that SNOMED CT is also well suited for clinical concepts in research [28].

The use of Dublin Core metadata to annotate various QAS in PRISM offers a way to employ the most appropriate controlled vocabulary(s) for the content, while preserving retrievability. In addition to selected Dublin Core metadata elements and controlled terminologies, other decisions were made to ensure that these elements and vocabularies were used appropriately and consistently. Specifically, assuring each QAS is usable, reproducible, and understandable on its own merit. For example, a form that addresses gynecological issues may include a QAS addressing menstrual symptoms, such as “Do you experience cramps?” From the context of the form, this can refer only to menstrual cramps and may be coded with or without the “menstrual semantics”. However, when a later user searches the registry for library questions about abdominal, leg or other body site cramps, this question may be inappropriately selected. The QAS metadata (including the embedded SNOMED CT codes and the narrative definition of the question that PRISM includes) can be used to easily disambiguate this term.

Under the leadership of terminology experts from the College of American Pathologists, guidelines for using SNOMED CT were developed related to post-coordination, selection of hierarchies, and level of specificity. Guidelines for the use of SCT within the PRISM Library data model were developed collaboratively by the PRISM team and included the following decisions:

1. Assigning codes for entire question and answer set groups vs. discrete codes for questions and answers

2. An approach to take the (semantically) closest available SNOMED CT concept rather than creating a new one

3. Consequently, we considered but rejected the idea to create a SNOMED extension (“Ref Set”) mechanism

4. Versioning and change protocols were developed between USF and CAP partners

Procedures for Adding New Items / Growing PRISM content

As described in the previous section, PRISM has developed a useful, consistent, and standards-compliant solution for the encoding of questions in PRISM. Ultimately, the metadata model and indexing strategy will be tested as the content of PRISM grows. Implementation plans should ensure that as the size of the PRISM library grows, duplicate questions are not inadvertently added. (Anecdotally, this is an issue with other question and metadata repositories, owing to the fact that a complete search of existing content must be undertaken before new content is added, and this search is both time-consuming and generally not incentivized.) For PRISM to remain a useful

resource to rare disease registry developers, and for PRISM to support the goal to reduce question variation across rare disease registries, the library should not contain obvious duplicates and the library indexing model should be sufficient for users to search *and retrieve* relevant registry questions from PRISM before they re-create their own variation. As new content is introduced to the PRISM outreach team at USF, they first determine if the content is in scope of PRISM (ie, is not purview of other standards initiatives shown in Table 1 and Table 2). Then, as questions are identified for possible addition into the library, library professionals perform multiple and semantically enhanced queries to compare to QAS already in the database. Through collaborations with rare disease stakeholders who are part of the PRISM scientific review committee, they are then edited for structure and clarity, and definitions are provided. After the questions have been edited, semantic and administrative metadata are applied (including controlled terminology) before integrating them into the library's database.

Discussion

Limited Scope

One of the biggest challenges of the PRISM project has been to keep the scope reasonable and practical. Since the targeted audience for PRISM is researchers and registry developers (including PAGs with non-research backgrounds), we took a minimalist approach to coding with the goal of easy retrieval. It is our expectation that easy retrieval will drive increased usage, which will cultivate de facto standards, and those standards will ultimately support interoperability (see Figure 1).

We recognize that interoperability between registry questions and other data (eg, EHR data) would require more sophisticated coding with SNOMED CT and other data standards. Given the short duration of this project and the desire for maximum retrievability, we determined that this level of coding would be out of the scope for PRISM at this time.

It is not clear at this point how large a corpus of sharable items that there is among disparate and different rare diseases. Likewise, it is still unclear whether items for rare diseases are likely to be similar to items for more common diseases, and if so, whether there would be value in finding a way to include and reuse those items via PRISM as well. The future use and evaluation of PRISM content by multiple disease representatives will yield information on the reusability of the questions in PRISM within and across rare and common diseases, as well as provide practical examples of cross-disease standards and determination of standards gaps. To understand the reusability and generalizability of PRISM questions, additional work needs to explore the validity and reproducibility of the categorization and indexing of questions.

Figure 1. Theory of PRISM design to interoperability.

The Future

PRISM fills a void for the rare diseases research and registry development communities. The PRISM library resource can support standardized data collection in patient registries by reducing unnecessary variation. PRISM is free and available to search through the project website [25]. Authors are also in the process of making the PRISM library metadata model and content accessible through the National Center for Biomedical Ontology [29] Bioportal. While the library content is available, we encourage innovators and developers to build tools that integrate PRISM content into registry development efforts. It is possible that the PRISM library resource could be leveraged in large research data collection tools such as the NIH-funded REDCap project.

Authors are hopeful for future funding that will allow PRISM content to grow to meet the needs of the thousands of rare diseases registry applications and to allow computer mediated methods for adding and presenting content. The notion of using a distributed community of registry developers to curate this resource by commenting on and ranking items—as with the demonstration of caDSR content that is described in [30]—is very appealing. Such projects would require extensive marketing and publicity of PRISM to a comprehensive group of rare disease registry stakeholders and researchers in order to bolster the extensive use of PRISM that would be required to effectively demonstrate a bottom-up community driven development of standards. An increased usage and future growth of PRISM will in turn require formal governance for PRISM and official policies related to content and a submissions and update processes.

The PRISM developers, with the cooperation and support of the National Organization for Rare Disorders (NORD), are working to make the PRISM resource available and useful to registry developers representing all rare diseases and all countries. Currently, several rare disease patient advocacy organizations are participating in focus groups and expert interviews to inform the development of best interfaces and retrieval strategies to ensure that PRISM is a useful and accessible community-driven resource. In addition, we are developing international collaborations to explore the translation of items to support global rare disease research. Our overarching goal is that—given the sheer number of rare diseases, the variety of registry designs, and the number of languages that might need to be addressed—the PRISM leadership seizes and implements standards opportunities without burdening resource-strapped rare disease communities.

Summary

The lack of a clear set of standards and specifications for data collection using patient registries represents a significant data standards gap in an explosively growing application area—important to both drug development and patient-directed health communities. Standardization of patient registries can enable the interoperability of health and research data, as registries should be able to receive data from health care system or transmit data into various clinical research or pharmacovigilance applications. PRISM can be used to facilitate interoperability of existing and newly developed registries and to ensure that moving forward, registries use standard sets of questions. Without the use of such a resource, the proliferation of patient registries and variation of data collection questions will be inevitable. This central resource, the PRISM library, will support a bottom-up and incremental standards promulgation. By using a standard set of metadata elements and SNOMED CT to facilitate the retrieval and re-use of existing questions and standards, PRISM will reduce variation in the rare diseases registry community and assist registry implementers to produce high quality registries much more efficiently than ever before. Once variation in patient registries is reduced (ie, “standards” emerge), then issues related to harmonizing, mapping, and relating to the different standards communities for health care (eg, HL7) and research (eg, CDISC) can be addressed in an efficient manner. In this approach, the standardization of patient registry questions can serve to improve efficiencies, collaboration, and resource sharing across the entire drug development process.

Information regarding the development of PRISM and access can be found on the PRISM website [24] .

Acknowledgments

The [RD] PRISM Library project was funded by the American Reinvestment and Recovery Act. Support for the [RD] PRISM Library project is administered through the National Library of Medicine (NLM Grant Number 1RC1LM010455-01), a component of the NIH, and supported by the Office of Rare Diseases Research. The contents are solely the responsibility of the authors and do not necessarily represent the official views of NLM or ORDR or NIH.

The authors wish to thank Alice Graves, Heather Guillette, Jamie Malloy, Shu Liu, and Sarah Austin from the University of South Florida for their contributions to the PRISM project, and science and program officers from the NIH Office of Rare Diseases Research for their support.

The Cancer Checklists and the tools described in this article were developed by our colleagues at the College of American Pathologists, who are our partners in this work. Specifically we want to thank Narciso Albarracin, Christine Spisla, Rich Moldwin, Debra Konicek, Jaleh Mirza, and Debbie Klieman for their substantial efforts in the [RD] PRISM project.

The authors also thank the National Organization for Rare Disorders (NORD) and their member organizations for their support and advice, and for participation in ongoing evaluation and quality improvement activities.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sample questions and selected metadata.

[PDF File (Adobe PDF File), 318KB - ijmr_v1i2e10_app1.pdf]

References

1. Watson MS, Epstein C, Howell RR, Jones MC, Korf BR, McCabe ER, et al. Developing a national collaborative study system for rare genetic diseases. Genet Med 2008 May;10(5):325-329. [doi: 10.1097/GIM.0b013e31817b80fd] [Medline: 18496030]
2. Richesson R, Vehik K. Patient registries: utility, validity and inference. Adv Exp Med Biol 2010;686:87-104. [doi: 10.1007/978-90-481-9485-8_6] [Medline: 20824441]
3. Luisetti M, Campo I, Scabini R, Zorzetto M, Kadija Z, Mariani F, et al. The problems of clinical trials and registries in rare diseases. Respir Med 2010 Jul;104 Suppl 1:S42-S44. [doi: 10.1016/j.rmed.2010.03.016] [Medline: 20413285]
4. Hollak CEM. OrphaNews Europe: the newsletter of the European Union Committee of Experts on Rare Diseases Internet. 2011. Disease-based registries for rare disorders URL: http://www.orpha.net/actor/cgi-bin/OAhome.php?Ltr=EuropaNews [accessed 2012-09-17] [WebCite Cache ID 6AktOboSY]
5. Jones S, James E, Prasad S. Disease registries and outcomes research in children: focus on lysosomal storage disorders. Paediatr Drugs 2011 Feb 1;13(1):33-47. [doi: 10.2165/11586860-000000000-00000] [Medline: 21162599]
6. Green SB, Byar DP. Using observational data from registries to compare treatments: the fallacy of omnimetrics. Stat Med 1984;3(4):361-373. [Medline: 6396792]
7. CDISC. CDISC. 2005. Clinical Data Interchange Standards Consortium URL: http://www.cdisc.org/ [accessed 2012-09-17] [WebCite Cache ID 6AktSnLMO]
8. HL7. Health Level Seven, Inc. Health Level Seven URL: http://www.hl7.org/ [accessed 2012-09-17] [WebCite Cache ID 6AktURs3q]
9. McDonald CJ, Huff SM, Suico JG, Hill G, Leavelle D, Aller R, et al. LOINC, a universal standard for identifying laboratory observations: a 5-year update. Clin Chem 2003 Apr;49(4):624-633 [FREE Full text] [Medline: 12651816]
10. Warzel DB, Andonaydis C, McCurry B, Chilukuri R, Ishmukhamedov S, Covitz P. Common data element (CDE) management and deployment in clinical trials. 2003 Presented at: AMIA Annu Symp Proc. :1048; 2003; Washington, DC.
11. Lawson DM. PROMIS: a new tool for the clinician scientist. J Can Chiropr Assoc 2011 Mar;55(1):16-19. [Medline: 21403776]
12. Gershon RC, Rothrock N, Hanrahan R, Bass M, Cella D. The use of PROMIS and assessment center to deliver patient-reported outcome measures in clinical research. J Appl Meas 2010;11(3):304-314. [Medline: 20847477]
13. Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, et al. The PhenX Toolkit: get the most from your measures. Am J Epidemiol 2011 Aug 1;174(3):253-260 [FREE Full text] [doi: 10.1093/aje/kwr193] [Medline: 21749974]
14. DHHS. U.S. Department of Health and Human Services. 2007 Apr 3. Office of the National Coordinator for Health Information Technology (ONC) URL: http://www.hhs.gov/healthit/chiinitiative.html [accessed 2012-09-17] [WebCite Cache ID 6AkuoYqsU]
15. ONC. Comments related to registries and management of special patient populations: Hearing before the Meaningful Use Workshop. In: US Office of the National Coordinator (ONC) for Health Information Technology. 2007 Presented at: Meaningful Use Workshop; 2007; Washington.
16. Seminara J, Tuchman M, Krivitzky L, Krischer J, Lee HS, Lemons C, et al. Establishing a consortium for the study of rare diseases: The Urea Cycle Disorders Consortium. Mol Genet Metab 2010;100 Suppl 1:S97-105. [doi: 10.1016/j.ymgme.2010.01.014] [Medline: 20188616]
17. Merkel PA, Cuthbertson DD, Hellmich B, Hoffman GS, Jayne DR, Kallenberg CG, Vasculitis Clinical Research Consortium. Comparison of disease activity measures for anti-neutrophil cytoplasmic autoantibody (ANCA)-associated vasculitis. Ann Rheum Dis 2009 Jan;68(1):103-106. [doi: 10.1136/ard.2008.097758] [Medline: 18664546]
18. Percy AK, Neul JL, Glaze DG, Motil KJ, Skinner SA, Khwaja O, et al. Rett syndrome diagnostic criteria: lessons from the Natural History Study. Ann Neurol 2010 Dec;68(6):951-955. [doi: 10.1002/ana.22154] [Medline: 21104896]
19. Richesson RL, Lee HS, Cuthbertson D, Lloyd J, Young K, Krischer JP. An automated communication system in a contact registry for persons with rare diseases: scalable tools for identifying and recruiting clinical research participants. Contemp Clin Trials 2009 Jan;30(1):55-62. [doi: 10.1016/j.cct.2008.09.002] [Medline: 18804556]
20. NIH. NIH Policy on Reporting Race and Ethnicity Data. NIH Policy on Reporting Race and Ethnicity Data: Subjects in Clinical Research; Notice NOT-OD-01-053; Release Date: August 8, 2001. National Institutes of Health; 2001. URL: http://grants.nih.gov/grants/guide/notice-files/NOT-OD-01-053.html [accessed 2012-09-17] [WebCite Cache ID 6AkuJjFGl]

21. NINDS. NINDS.: National Institute of Neurological Disorders and Stroke; 2010. NINDS Common Data Elements. harmonizing information. Streamlining research. Project overview URL: http://www.commondataelements.ninds.nih.gov/ProjReview.aspx [accessed 2012-09-17] [WebCite Cache ID 6AkuScHEk]
22. Rubinstein YR, Groft SC, Bartek R, Brown K, Christensen RA, Collier E, et al. Creating a global rare disease patient registry linked to a rare diseases biorepository database: Rare Disease-HUB (RD-HUB). Contemp Clin Trials 2010 Sep;31(5):394-404. [doi: 10.1016/j.cct.2010.06.007] [Medline: 20609392]
23. Vreeman DJ, McDonald CJ, Huff SM. LOINC® - A Universal Catalog of Individual Clinical Observations and Uniform Representation of Enumerated Collections. Int J Funct Inform Personal Med 2010;3(4):273-291. [doi: 10.1504/IJFIPM.2010.040211] [Medline: 22899966]
24. Loring DW, Lowenstein DH, Barbaro NM, Fureman BE, Odenkirchen J, Jacobs MP, et al. Common data elements in epilepsy research: development and implementation of the NINDS epilepsy CDE project. Epilepsia 2011 Jun;52(6):1186-1191. [doi: 10.1111/j.1528-1167.2011.03018.x] [Medline: 21426327]
25. PRISM. 2012. What is PRISM? Homepage URL: http://prism.epi.usf.edu/ [accessed 2012-09-17] [WebCite Cache ID 6AkuZ6YdO]
26. Hersh W. Information retrieval: a health and biomedical perspective. New York: Springer; 2003.
27. Richesson R, Shereff D, Andrews J. [RD] PRISM Library: Patient Registry Item Specifications and Metadata for Rare Diseases. J Libr Metadata 2010 Apr 1;10(2-3):119-135. [doi: 10.1080/19386389.2010.506385] [Medline: 21057650]
28. Richesson RL, Andrews JE, Krischer JP. Use of SNOMED CT to represent clinical research data: a semantic characterization of data items on case report forms in vasculitis research. J Am Med Inform Assoc 2006;13(5):536-546 [FREE Full text] [doi: 10.1197/jamia.M2093] [Medline: 16799121]
29. Jonquet C, Lependu P, Falconer S, Coulet A, Noy NF, Musen MA, et al. NCBO Resource Index: Ontology-Based Search and Mining of Biomedical Resources. Web Semant 2011 Sep 1;9(3):316-324. [doi: 10.1016/j.websem.2011.06.005] [Medline: 21918645]
30. Moser RP, Hesse BW, Shaikh AR, Courtney P, Morgan G, Augustson E, et al. Grid-enabled measures: using Science 2.0 to standardize measures and share data. Am J Prev Med 2011 May;40(5 Suppl 2):S134-S143. [doi: 10.1016/j.amepre.2011.01.004] [Medline: 21521586]

Abbreviations

caDSR: Cancer Data Standards Registry and Repository
CAP: College of American Pathologists SNOMED Terminology Solutions
CDISC: Clinical Data Interchange Standards Consortium
CSHARE: CDISC Shared Health And Clinical Research Electronic Library (http://www.cdisc.org/cdisc-share)
DC: Dublin Core Metadata Initiative
HL7: Health Level Seven
LOINC: Logical Observation Identifiers Names and Codes
NINDS: National Institute of Neuromuscular Disorders and Stroke
NLM: National Library of Medicine
NORD: National Organization for Rare Disorders
OMB: The White House Office of Management and Budget
ONC: Office of National Coordinator for Health Information Technology
ORDR: Office of Rare Diseases Research
PAG: Patient Advocacy Group
PhenX: Consensus measures for Phenotypes and eXposures
PRISM: Patient Registry Item Specification and Metadata
PROMIS: Patient Reported Outcomes Measurement Information System
QAS: Question and Answer Set
RDCRN: Rare Disease Clinical Research Network (http://www.rarediseasesnetwork.org)
SCT: SNOMED CT
SDO: Standards Development Organization
SNOMED CT: Systematized Nomenclature of Medicine–Clinical Terms
USF: University of South Florida

Edited by K El Emam, G Eysenbach; submitted 15.03.12; peer-reviewed by T Rosenbloom, A Sprague; comments to author 14.05.12; revised version received 02.07.12; accepted 30.08.12; published 10.10.12.

Please cite as:
Richesson RL, Shereff D, Andrews JE
Standardization of Questions in Rare Disease Registries: The PRISM Library Project
Interact J Med Res 2012;1(2):e10
URL: http://www.i-jmr.org/2012/2/e10/
doi:10.2196/ijmr.2107
PMID:23611924

Original Paper

Physician Satisfaction Following Electronic Health Record Adoption in Three Massachusetts Communities

Leonie Heyworth[1,2], MD, MPH; Fang Zhang[3], PhD; Chelsea A Jenter[3], MPH; Rachel Kell[4], MPH; Lynn A Volk[5], MHS; Micky Tripathi[4], PhD, MPP; David W Bates[2,5], MD, MSc; Steven R Simon[2,6], MD, MPH

[1]VA Boston Healthcare System, Section of General Internal Medicine, Jamaica Plain, MA, United States

[2]Brigham and Women's Hospital, Division of General Internal Medicine, Boston, MA, United States

[3]Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, MA, United States

[4]Massachusetts eHealth Collaborative, Waltham, MA, United States

[5]Partners Health Care System, Wellesley, MA, United States

[6]VA Boston Health Care, General Internal Medicine, Boston, MA, United States

Corresponding Author:
Leonie Heyworth, MD, MPH
VA Boston Healthcare System
Section of General Internal Medicine
150 S. Huntington Ave, Suite 425
Jamaica Plain, MA, 02130
United States
Phone: 1 857 364 5704
Fax: 1 857 364 6990
Email: lheyworth@gmail.com

Abstract

Background: Despite mandates and incentives for electronic health record (EHR) adoption, little is known about factors predicting physicians' satisfaction following EHR implementation.

Objective: To measure predictors of physician satisfaction following EHR adoption.

Methods: A total of 163 physicians completed a mailed survey before and after EHR implementation through a statewide pilot project in Massachusetts. Multivariable logistic regression identified predictors of physician satisfaction with their current practice situation in 2009 and generalized estimating equations accounted for clustering.

Results: The response rate was 77% in 2005 and 68% in 2009. In 2005, prior to EHR adoption, 28% of physicians were very satisfied with their current practice situation compared to 25% in 2009, following EHR adoption ($P < .001$). In multivariate analysis, physician satisfaction following EHR adoption was correlated with self-reported ease of EHR implementation (adjusted odds ratio [OR] = 5.7, 95% CI 2.1 - 16), resources for practice improvement (adjusted OR = 2.6, 95% CI 1.2 - 6.1), pre-intervention satisfaction (adjusted OR = 4.8, 95% CI 1.5 - 15), and stress (adjusted OR = 5.3, 95% CI 1.1 - 25). Male physicians reported lower satisfaction following EHR adoption (adjusted OR = 0.3, 95% CI 0.2 - 0.6).

Conclusions: Interventions to expand EHR use should consider additional support for practices with fewer resources for improvement and ensure ease of EHR implementation. EHR adoption may be a factor in alleviating physicians' stress. Addressing physicians' satisfaction prior to practice transformation and anticipating greater dissatisfaction among male physicians will be essential to retaining the physician workforce and ensuring the quality of care they deliver.

(Interact J Med Res 2012;1(2):e12) doi:10.2196/ijmr.2064

KEYWORDS

electronic health record; physician satisfaction; implementation; Massachusetts eHealth collaborative

Introduction

Electronic health records (EHRs) have the potential to transform health care by improving the quality and safety of care delivered, particularly if adopted on a wide-scale basis. EHRs may also reduce the cost of providing ambulatory care [1,2]. Despite emerging evidence about these potential benefits, increasing public policy attention toward universal EHR adoption by

legislative leaders, and financial incentives for EHR use, little is known about physician experience following EHR adoption.

Successful EHR implementation generally requires users to be satisfied; therefore, defining predictors of satisfaction may increase the likelihood of a system's success. Recent studies examining existing physician EHR users found that larger practices and younger physicians were more likely to adopt EHRs [3,4]. In addition, EHR users were found to be generally satisfied with their system [3]. Another cross-sectional study of physicians' EHR-related satisfaction found that it was time-dependent and associated with use of a more robust EHR [5]. Prior studies have demonstrated physician satisfaction with new EHR adoption to be associated with the success of electronic implementation [6]; however, few prospective data are available regarding physicians' overall satisfaction with their practice following EHR adoption.

It is estimated that 27% of practices in the United States were using a fully functional EHR in 2010 [7]. The mandate for universal EHR implementation and robust clinical data exchange by 2014 [8], and incentives in place for EHR adoption [7] mean that very large numbers of practices in the United States are, or likely soon will be, in the process of EHR implementation. EHR adoption may not only impact how a physician writes prescriptions and notes, but may also influence the overall practice environment by changing workflow and physician-patient interaction. In addition, the emerging patient-centered medical home as a model for health care delivery rewards practices for EHR adoption [9], even though the EHRs of today are variable at meeting the needs of medical homes [10,11]. It is likely that EHR implementation is occurring simultaneously with practice changes to fit the medical home model, further underscoring the need to assess physicians' satisfaction with an EHR, but also with their practice overall.

Using Konrad and Linzer's framework for measuring physicians' work-life satisfaction [12,13], we identified several key facets that would be relevant to physicians' satisfaction with their practice post-EHR adoption. We speculated that intrinsic factors (eg, years in practice, gender, race, self-reported stress, computer literacy, and satisfaction prior to EHR implementation), pay (in the form of bonuses for EHR use), and availability of practice resources would predict physician satisfaction following EHR adoption. Therefore, we undertook the present study to identify correlates of physician satisfaction with their practice following EHR adoption, in the context of a pilot community-based collaborative model in Massachusetts.

Methods

The methods of survey development and administration have been described elsewhere and are summarized subsequently [14]. The survey and research protocol were approved by the Partners HealthCare Human Research Committee.

Setting: The Massachusetts eHealth Collaborative

The Massachusetts eHealth Collaborative (MAeHC) is a not-for-profit consortium established in 2004 as a pilot demonstration project to facilitate the adoption and implementation of EHRs. Following a competitive application process involving 35 eligible communities, three communities were selected as demonstration sites in March 2005: North Adams/Williamstown, Brockton, and Newburyport. Physicians and practices within these communities were found to be generally representative of Massachusetts [14]. A total of 548 physicians (97.7%) in 159 practices (95.2%) accepted the offer to participate. A select group of EHR vendors were identified through a comprehensive process and agreed to competitive EHR pricing. The EHR vendors Allscripts, eClinicalWorks, NextGen Healthcare Information Systems, and GE Healthcare were selected (all of whom were subsequently certified by the Certification Commission for Health Information Technology). As of the pilot's completion date [15], overall participation was 84.0% for physicians in 134 practices in the pilot's three communities. EHRs were implemented in 97.9% of practices between March 2006 and March 2008.

The Massachusetts eHealth Collaborative Survey

The MAeHC physicians were surveyed using the same survey instrument developed for a Massachusetts statewide survey of EHR use, described elsewhere [4] and available from the corresponding author upon request. The survey included demographic information about the physician, including age, gender, and self-reported race (physicians had the option to select one or more of the following categories: Asian, American Indian or Alaskan Native, black or African American, Native Hawaiian or other Pacific Islander, white, or "other"). Race data were collected in order to compare survey data with census figures about population distributions. The baseline MAeHC survey was delivered to participating physicians between September and October 2005. Between January and February 2009, a similar follow-up survey was delivered after physicians had been using a new EHR for at least one year. Surveys were mailed to all physicians who had signed agreements to participate in the MAeHC. Physicians were instructed to complete the survey and return the survey by mail or to an MAeHC staff member visiting their practice site. No cash incentive was offered for survey completion because participation in the evaluations was a condition for MAeHC involvement. Between 2005 and 2009, some practices dissolved, some physicians departed, and new physicians entered the communities. A total of 163 physicians from 134 practices completed both the 2005 and 2009 survey questionnaires [16].

Statistical Analysis

For all statistical analyses, SAS version 9.2 software (SAS Institute Inc, Cary, NC, USA) was used. Logistic regression was used to analyze predictors of self-reported physician satisfaction following EHR implementation, and Chi-square analysis was used to examine the association between physician satisfaction and perceived EHR vendor characteristics. The dependent variable in these analyses was the response to the 2009 survey item: "Overall, how satisfied are you with your current practice situation?" Physicians were asked to check 1 of 4 possible responses: very satisfied, generally satisfied, somewhat dissatisfied, or very dissatisfied. These responses were dichotomized in the analyses to "very satisfied" versus "generally satisfied, somewhat dissatisfied, or very dissatisfied." In this logistic regression model, physician satisfaction was

examined as a function of exogenous and perceptual predictor variables modeled after Konrad and Linzer's 10 facets of measuring physician satisfaction [12,13]. Exogenous predictor variables included were whether the practice had financial incentives for EHR adoption, the level of financial resources available for practice improvement or expansion, size of practice (1-4 vs 5+ physicians or nurse practitioners and physician assistants), practice specialty (primary care vs specialty practice), and whether the practice experienced difficulty with EHR implementation (self-reported). Perceptual variables included physicians' views of whether personal or professional stress was a problem in their practice and how comfortable they felt with the use of computers prior to EHR implementation and their daily patient volume in 2009. Physician characteristics, including age, gender, self-reported satisfaction in 2005, and practice ownership, were also included *a priori* in the final logistic regression model. Generalized estimating equations were used for the logistic regression to account for the correlations within each practice [17,18].

Results

Respondent Characteristics

A total of 163 surveys were returned from the same physicians in 2005 and 2009. The response rate in 2005 was 76.5% and was 68.2% in 2009. Respondents and non-respondents were similar with respect to specialty and practice size [14]. Of the 163 respondents to both surveys, 147 responded to the question, "Overall, how satisfied are you with your current practice situation?" in both the 2005 and 2009 surveys. Of the respondents to both surveys, 44 had missing responses to questions used as explanatory variables, resulting in 119 respondents' surveys eligible for subsequent analysis.

Physician Satisfaction

Physicians responding to both the 2005 and the 2009 MAeHC surveys had adopted an EHR for at least one year, and were primarily white (79.1%) men (73.6%) in practice for more than 15 years (89.0%). Overall, 26.4% reported feeling very satisfied with their current practice situation in 2005, compared to 23.9% in 2009 ($P < .001$). Approximately half (46.6%) of responding physicians worked in practices with > 5 physicians or nurse practitioners (NPs) or physician assistants (PAs), and 60.7% worked in specialty (non-primary care) practices. A majority (61.4%) reported incentives for health information technology (HIT) use. A total of 20.2% reported that their practices had financial resources available for practice improvement or expansion. A minority of physicians (16.6%) post-EHR adoption reported that they would prefer to give up their EHR and return to paper records (Table 1).

Predictors of Physician Satisfaction with Current Practice

In logistic regression analyses, ease of EHR implementation and male gender were the strongest independent predictors of physician satisfaction among physicians adopting a new EHR (Table 2). Physicians who found that their EHR implementation process was not very difficult were much more likely to report feeling satisfied in their current practice in 2009 (adjusted odds ratio [OR] = 5.7, 95% CI 2.1 - 16). Male physicians were much less likely to be satisfied following EHR adoption (adjusted OR = 0.3, 95% CI 0.2 - 0.6), whereas physicians who reported higher personal or professional stress before EHR adoption were more satisfied following implementation (adjusted OR = 5.3, 95% CI 1.1 - 25). Physicians' report of feeling very satisfied in their current practice at the time of their baseline survey in 2005 was also strongly correlated with satisfaction in 2009 (adjusted OR = 4.8, 95% CI 1.5 - 15). Physicians in practices with financial resources available for practice improvement or expansion also were more likely to report feeling satisfied in their current practice in 2009 (adjusted OR = 2.6, 95% CI 1.2 - 6.1).

In the Chi-square analysis, we found that baseline levels of satisfaction were not related to post-intervention (2009) reports of easy EHR implementation ($P = .62$). This analysis provided evidence that the observed relationship between ease of EHR implementation and 2009 satisfaction was not likely to be confounded by baseline levels of satisfaction. We also confirmed that satisfaction post-EHR adoption was strongly associated with the physician-reported training level of the EHR vendor ($P = .002$) and the quality of assistance provided by MAeHC senior leadership ($P < .001$) and staff ($P = .002$), but found that prompt support from the information technology (IT) support team for problems related to the EHR application, or for hardware or network connectivity, did not impact practice satisfaction (Table 3).

Table 1. Sample characteristics and demographics of physician respondents to both the 2005 and the 2009 Massachusetts eHealth Collaborative (MAeHC) surveys (n = 163).[a]

Characteristic	n (%)
Gender	
Male	120 (73.6%)
Female	38 (23.3%)
Missing response	5 (3.1%)
Years in practice	
15 or greater	145 (89.0%)
Less than 15	18 (11.0%)
Race	
White	129 (79.1%)
Asian	15 (9.2%)
American Indian or Alaskan Native	1 (0.6%)
Black or African American	4 (2.5%)
Native Hawaiian or other Pacific Islander	0
Other	4 (2.5%)
Missing response	10 (6.1%)
Practice size	
Solo or partner practice	47 (28.8%)
3-5 physicians or NPs/PAs	24 (14.7%)
> 5 physicians or NPs/PAs	76 (46.6%)
Missing response	16 (9.8%)
Practice type [b]	
Primary care	63 (38.7%)
Specialty	99 (60.7%)
Missing response	1 (0.6%)
Moderate or extensive financial resources available for practice improvement or expansion	33 (20.2%)
Incentives for HIT	100 (61.4%)
Satisfaction with current practice	
Very satisfied in 2005	43 (26.4%)
Very satisfied in 2009	39 (23.9%)
Personal/professional stress	
A moderate or serious problem in 2005	57 (35.0%)
A moderate or serious problem in 2009	61 (37.4%)
Comfort with computers	
Very comfortable in 2005	74 (45.4%)
Very comfortable in 2009	87 (53.4%)
EHR implementation [c]	

Characteristic	n (%)
Very difficult	54 (33.1%)
Somewhat or not difficult	102 (62.6%)
Missing response	7 (4.3%)
Preference to return to paper records in 2009 [d]	
Yes	27 (16.6%)
No	136 (83.4%)

[a] Not all respondents answered all questions. Responses are from 2009 survey unless otherwise indicated.

[b] Primary care practices in 2005 include solo/group/partnership practices; specialty practices include single or multispecialty solo/group/partnership practices.

[c] There were no responses to the answer “not sure.”

[d] There were no responses to the answer “not applicable.”

Table 2. Multivariate relationship between practice characteristics and physician perceptions of practices and physician satisfaction (n = 119).[a]

Characteristic	Adjusted OR	95% CI	*P* value
Gender			
Male	0.3	0.2 - 0.6	< .001
Years in practice			
15 or greater	2.3	0.7 - 7.2	.17
Race			
White	2.1	0.8 - 5.5	.14
Practice ownership	3.0	0.3 - 28	.33
Practice size			
Solo or partner practice (reference)	--	--	
3-5 physicians or NPs/PAs			
> 5 physicians or NPs/PAs	2.5	0.8 - 7.9	.11
Practice type [b]			
Primary care	1.3	0.4 - 0.6	.73
Daily patient volume			
≤ 15 (reference)	--	--	
> 16	0.5	0.1 - 2.5	.40
Moderate or extensive financial resources available for practice improvement or expansion	2.6	1.2 - 6.1	.02
Incentives for HIT [c]	1.7	0.2 - 10	.50
Physician perceptions			
Very satisfied with current practice (2005)	4.8	1.5 - 15	.007
Personal/professional stress a moderate or serious problem (2005)	5.3	1.1 - 25	.04
Very comfortable with computers (2005)	1.6	0.3 - 10	.60
No difficulty with EHR implementation (2009)	5.7	2.1 - 16	< .001

[a] Logistic regression analysis with generalized estimating equations, modeling the outcome, physician satisfaction with current practice situation, as a function of all listed characteristics. The model included all respondents (n = 163) with non-missing values for all variables included in the model.

[b] Primary care practices in 2005 include solo/group/partnership practices.

[c] Practice income or personal earnings for the use of electronic information systems.

Table 3. Bivariate relationship between physicians feeling very satisfied with their practice following EHR adoption and their perceptions of EHR vendor characteristics (n = 159).

Characteristic	*P* value
Excellent or very good EHR vendor quality	.002
Excellent or very good help provided by MAeHC	
Senior leadership	< .001
Senior staff	< .001
Prompt support from information technology team	
For hardware/ network connectivity	.82
For EHR application issues	.74

Discussion

We examined physicians' satisfaction with their practice in the context of new EHR adoption through a pilot community-based collaborative model in Massachusetts (MAeHC) and found that physician satisfaction declined modestly overall after EHR adoption. We found that one of the strongest predictors of physicians reportingsatisfaction post-EHR adoption was the ease with which the EHR implementation occurred. In subsequent analyses, we found that satisfaction was related to physician-perceived EHR vendor training and the quality of assistance provided by MAeHC senior leadership and staff. There was no association with availability of prompt support from the IT team for hardware or connectivity issues, or for problems related to the EHR application. This finding emphasizes the importance of site-specific experience and potentially explains why physicians within the MAeHC may have reported varying experiences with ease of EHR implementation.

In the context of Konrad and Linzer's framework of the necessary facets for physician job satisfaction [13], our results support the collective presence of a practice's resources (financial resources for improvement or expansion), community (MAeHC, a community-wide pilot project), and administrative features (ease of EHR implementation) in nurturing physician satisfaction during a period of EHR adoption. Other predictors of satisfaction included physician's satisfaction with their practice prior to EHR implementation. We also found that increased self-reported physician stress before EHR adoption was a predictor of satisfaction following EHR implementation. Although little research exists on the issue of physician stress and workflow, our results raise the intriguing possibility that physicians with higher stress levels may perceive a greater benefit from EHR implementation, perhaps by improving workflow, efficiency, and care coordination, resulting in higher levels of satisfaction. We did not find any associations between satisfaction following EHR adoption and physician age, practice size or ownership, physicians' patient volume, or number of years in practice. The presence of financial incentives for EHR adoption was not associated with post-adoption satisfaction in this study.

Our finding that male physicians were much less likely to report being satisfied following EHR implementation adds to the existing mixed literature on gender differences in physician career satisfaction. The largest study of female physicians, the Women Physicians' Health Study [19], found that 84% of respondents reported being "usually," "almost always," or "always" satisfied. Another study demonstrated that male physicians report greater job dissatisfaction, work-life stress, and isolation in their work environment [20]. Existing social science literature has proposed several potential explanations for the higher reported satisfaction rates among professional women, including lower job expectations, socialization not to express discontent, and appreciation of different career characteristics compared with men [21-23]. The Physician Work Life Study, a large, nationally representative, randomly stratified sample, showed similar global satisfaction scores for men and women, although researchers found that women were more satisfied with their specialty and patient relationships than their levels of autonomy, relationships with the community, pay, and resources [24]. In the context of this literature, greater male dissatisfaction, particularly following practice transformation, is not an unexpected finding.

Our finding that overall physician satisfaction decreased between 2005 and 2009 may be explained by a possible ceiling effect within our study sample. We observed that 92% of physicians reported feeling very satisfied or satisfied in 2005, and 75% reported feeling very satisfied or satisfied in 2009, The physicians in our study may have been more satisfied at baseline than average physicians in Massachusetts, where the proportion of physicians reporting that they felt either very satisfied or satisfied remained approximately the same (39% in 2006 and 38% in 2009) [25]. The Massachusetts health care insurance reform law in October 2006, pressure from pay-for-performance measures, or the economic climate may have contributed toward the observed trends in our sample.

Few studies have examined physicians' satisfaction with their practices in the context of community-based EHR adoption. One strength of this study was that EHR adoption was facilitated through the same pilot program among all surveyed physicians in three communities. This pilot was specifically designed to streamline robust EHR implementation, which has been found previously to be correlated with physician satisfaction [5]. Since participation in evaluation was a condition for MAeHC involvement, we were fortunate to have robust survey response rates for both the 2005 and 2009 surveys [15]. Surveys conducted before and after EHR implementation also allowed

us to account for baseline satisfaction—our outcome of interest—in our analysis.

This study has several limitations. Despite high response rates, the sample size (n = 119) restricted the number of independent variables that could be examined as correlates of physician satisfaction. In addition, as in any observational study, unmeasured confounders may exist, although our models included a number of covariates to control for confounding to the extent possible. A 2010 study found that satisfaction with EHR use seems to increase over time, but this study was limited by a low response rate (28%) [5]. It is possible that the physicians in our study, with relatively recent EHR implementation (at least 1 year, but many less than 2 years) would have reported higher levels of overall satisfaction with their practice if surveyed at a later date [26]. Since meaningful use criteria were established after these surveys were conducted, clinician experience in the context of meaningful use and EHR implementation was not captured in this study. The study's setting and participants, generally comparable to physicians practicing within Massachusetts [4], may have limited generalizability to other states or communities, and these results may not generalize outside the United States. In addition, because all physicians were members of practices assisted by MAeHC with EHR implementation, physicians' experiences with EHR implementation in other communities by alternate EHR vendors may differ.

EHR adoption and meaningful use of health information exchange represent a national health care priority in the United States. Before implementation of the American Recovery and Reinvestment Act of 2009, EHR adoption rates had been sluggish in the United States. This has led policy makers to implement broad national incentives for providers, and policy makers have also actively sought successful programs as models for EHR design and implementation in order to meet the national goal of universal EHR implementation and robust clinical data exchange by 2014 [8]. Demonstration projects, such as the MAeHC, have great potential to play a transformative role in the quest for widespread EHR adoption when partnered with on-site effective leadership and staff support. Importantly, our finding of ease of EHR implementation being strongly associated with physicians' satisfaction is particularly noteworthy because sustained usage of EHR and physician promotion of EHR use within professional networks is fundamentally dependent on physicians' satisfaction. Despite relatively widespread use of incentives for EHR use and physicians' reports of comfort with computer use, neither of these was found to be associated with satisfaction in practice following EHR implementation. Our data, in-line with prior literature, underscore the need to adequately support practices with fewer financial resources for improvement and ensure physician's practice satisfaction even before embarking on EHR implementation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

MAeHC survey 2005.

[PDF File (Adobe PDF File), 186KB - ijmr_v1i2e12_app1.pdf]

Multimedia Appendix 2

MAeHC survey 2009.

[PDF File (Adobe PDF File), 318KB - ijmr_v1i2e12_app2.pdf]

References

1. Hillestad R, Bigelow J, Bower A, Girosi F, Meili R, Scoville R, et al. Can electronic medical record systems transform health care? Potential health benefits, savings, and costs. Health Aff (Millwood) 2005 Oct;24(5):1103-1117. [doi: 10.1377/hlthaff.24.5.1103] [Medline: 16162551]
2. Miller RH, West C, Brown TM, Sim I, Ganchoff C. The value of electronic health records in solo or small group practices. Health Aff (Millwood) 2005 Oct;24(5):1127-1137. [doi: 10.1377/hlthaff.24.5.1127] [Medline: 16162555]
3. DesRoches CM, Campbell EG, Rao SR, Donelan K, Ferris TG, Jha A, et al. Electronic health records in ambulatory care--a national survey of physicians. N Engl J Med 2008 Jul 3;359(1):50-60. [doi: 10.1056/NEJMsa0802005] [Medline: 18565855]
4. Simon SR, Kaushal R, Cleary PD, Jenter CA, Volk LA, Poon EG, et al. Correlates of electronic health record adoption in office practices: a statewide survey. J Am Med Inform Assoc 2007 Feb;14(1):110-117 [FREE Full text] [doi: 10.1197/jamia.M2187] [Medline: 17068351]
5. Menachemi N, Powers T, Au DW, Brooks RG. Predictors of physician satisfaction among electronic health record system users. J Healthc Qual 2010;32(1):35-41. [Medline: 20151590]
6. Mazzoleni MC, Baiardi P, Giorgi I, Franchi G, Marconi R, Cortesi M. Assessing users' satisfaction through perception of usefulness and ease of use in the daily interaction with a hospital information system. In: Proc AMIA Annual Fall Symp. 1996 Presented at: Proc AMIA Annual Fall Symp; 1996; Washington, DC p. 752-756.

7. Blumenthal D, Tavenner M. The "meaningful use" regulation for electronic health records. N Engl J Med 2010 Aug 5;363(6):501-504. [doi: 10.1056/NEJMp1006114] [Medline: 20647183]
8. The White House. President George W. Bush. Transforming Health Care: The President's Health Information Technology Plan; 2006.
9. Fisher ES. Building a medical neighborhood for the medical home. N Engl J Med 2008 Sep 18;359(12):1202-1205. [doi: 10.1056/NEJMp0806233] [Medline: 18799556]
10. Bates DW, Bitton A. The future of health information technology in the patient-centered medical home. Health Aff (Millwood) 2010 Apr;29(4):614-621. [doi: 10.1377/hlthaff.2010.0007] [Medline: 20368590]
11. Fernandopulle R, Patel N. How the electronic health record did not measure up to the demands of our medical home practice. Health Aff (Millwood) 2010 Apr;29(4):622-628. [doi: 10.1377/hlthaff.2010.0065] [Medline: 20368591]
12. Linzer M, Konrad TR, Douglas J, McMurray JE, Pathman DE, Williams ES, et al. Managed care, time pressure, and physician job satisfaction: results from the physician worklife study. J Gen Intern Med 2000 Jul;15(7):441-450. [Medline: 10940129]
13. Konrad TR, Williams ES, Linzer M, McMurray J, Pathman DE, Gerrity M, et al. Measuring physician job satisfaction in a changing workplace and a challenging environment. SGIM Career Satisfaction Study Group. Society of General Internal Medicine. Med Care 1999 Nov;37(11):1174-1182. [Medline: 10549620]
14. Simon SR, Kaushal R, Jenter CA, Volk LA, Burdick E, Poon EG, et al. Readiness for electronic health records: comparison of characteristics of practices in a collaborative with the remainder of Massachusetts. Inform Prim Care 2008;16(2):129-137. [Medline: 18713529]
15. Goroll AH, Simon SR, Tripathi M, Ascenzo C, Bates DW. Community-wide implementation of health information technology: the Massachusetts eHealth Collaborative experience. J Am Med Inform Assoc 2009 Feb;16(1):132-139 [FREE Full text] [doi: 10.1197/jamia.M2899] [Medline: 18952937]
16. Fleurant M, Kell R, Love J, Jenter C, Volk LA, Zhang F, et al. Massachusetts e-Health Project increased physicians' ability to use registries, and signals progress toward better care. Health Aff (Millwood) 2011 Jul;30(7):1256-1264. [doi: 10.1377/hlthaff.2010.1020] [Medline: 21734198]
17. Zeger SL, Liang KY. Longitudinal data analysis for discrete and continuous outcomes. Biometrics 1986 Mar;42(1):121-130. [Medline: 3719049]
18. Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. Biometrics 1988 Dec;44(4):1049-1060. [Medline: 3233245]
19. Frank E, McMurray JE, Linzer M, Elon L. Career satisfaction of US women physicians: results from the Women Physicians' Health Study. Society of General Internal Medicine Career Satisfaction Study Group. Arch Intern Med 1999 Jul 12;159(13):1417-1426. [Medline: 10399893]
20. Quinn MA, Wilcox A, Orav EJ, Bates DW, Simon SR. The relationship between perceived practice quality and quality improvement activities and physician practice dissatisfaction, professional isolation, and work-life stress. Med Care 2009 Aug;47(8):924-928. [doi: 10.1097/MLR.0b013e3181a393e4] [Medline: 19543122]
21. Phelan J. The paradox of the contented female worker: an assessment of alternative explanations. Soc Psychol Q 1994;57:95-107.
22. Robinson GE. STUDENTJAMA. Career satisfaction in female physicians. JAMA 2004 Feb 4;291(5):635. [doi: 10.1001/jama.291.5.635] [Medline: 14762047]
23. Chiu C. Do professional women have lower job satisfaction than professional men? Lawyers as a case study. Sex Roles 1998;38:521-536.
24. McMurray JE, Linzer M, Konrad TR, Douglas J, Shugerman R, Nelson K. The work lives of women physicians results from the physician work life study. The SGIM Career Satisfaction Study Group. J Gen Intern Med 2000 Jun;15(6):372-380. [Medline: 10886471]
25. Massachusetts Medical Society. Section 5: Physician Satisfaction, Attitudes Toward the Profession, And Future Career Plans. Physician Workforce Study 2011:34-50 [FREE Full text]
26. Zhou L, Soran CS, Jenter CA, Volk LA, Orav EJ, Bates DW, et al. The relationship between electronic health record use and quality of care over time. J Am Med Inform Assoc 2009 Aug;16(4):457-464 [FREE Full text] [doi: 10.1197/jamia.M3128] [Medline: 19390094]
27. Ford EW, Menachemi N, Peterson LT, Huerta TR. Resistance is futile: but it is slowing the pace of EHR adoption nonetheless. J Am Med Inform Assoc 2009 Jun;16(3):274-281 [FREE Full text] [doi: 10.1197/jamia.M3042] [Medline: 19261931]

Abbreviations

EHR: electronic health record
HIT: health information technology
MAeHC: Massachusetts eHealth Collaborative
NP: nurse practitioner
OR: odds ratio

PA: physician assistant

Edited by G Eysenbach; submitted 26.02.12; peer-reviewed by D Burke, C Or, PY Yen, K Zheng, C Kuziemsky; comments to author 29.07.12; revised version received 04.09.12; accepted 21.09.12; published 08.11.12.

Please cite as:
Heyworth L, Zhang F, Jenter CA, Kell R, Volk LA, Tripathi M, Bates DW, Simon SR
Physician Satisfaction Following Electronic Health Record Adoption in Three Massachusetts Communities
Interact J Med Res 2012;1(2):e12
URL: http://www.i-jmr.org/2012/2/e12/
doi:10.2196/ijmr.2064
PMID:23611987

Original Paper

DB4US: A Decision Support System for Laboratory Information Management

José M Carmona-Cejudo[1]; Maria Luisa Hortas[2]; Manuel Baena-García[3]; Jorge Lana-Linati[4]; Carlos González[2]; Maximino Redondo[5]; Rafael Morales-Bueno[1]

[1]University of Malaga, Computer Sciences and Languages Department, Málaga, Spain

[2]Hospital Costa del Sol, Clinical Laboratories Department, Marbella, Spain

[3]Clínicas Rincón, Rincón de la Victoria, Spain

[4]Siemens Healthcare Diagnostics, Barcelona, Spain

[5]Laboratorios Clínicos, REDISSEC, Marbella, Spain

Corresponding Author:
José M Carmona-Cejudo
University of Malaga
Computer Sciences and Languages Department
Departamento de Lenguajes y Ciencias de la Computación, Complejo Tecnológico, Campus de Teatinos.
Málaga, 29071
Spain
Phone: 34 952132863
Fax: 34 952131397
Email: jmcarmona@lcc.uma.es

Abstract

Background: Until recently, laboratory automation has focused primarily on improving hardware. Future advances are concentrated on intelligent software since laboratories performing clinical diagnostic testing require improved information systems to address their data processing needs. In this paper, we propose DB4US, an application that automates information related to laboratory quality indicators information. Currently, there is a lack of ready-to-use management quality measures. This application addresses this deficiency through the extraction, consolidation, statistical analysis, and visualization of data related to the use of demographics, reagents, and turn-around times. The design and implementation issues, as well as the technologies used for the implementation of this system, are discussed in this paper.

Objective: To develop a general methodology that integrates the computation of ready-to-use management quality measures and a dashboard to easily analyze the overall performance of a laboratory, as well as automatically detect anomalies or errors. The novelty of our approach lies in the application of integrated web-based dashboards as an information management system in hospital laboratories.

Methods: We propose a new methodology for laboratory information management based on the extraction, consolidation, statistical analysis, and visualization of data related to demographics, reagents, and turn-around times, offering a dashboard-like user web interface to the laboratory manager. The methodology comprises a unified data warehouse that stores and consolidates multidimensional data from different data sources. The methodology is illustrated through the implementation and validation of DB4US, a novel web application based on this methodology that constructs an interface to obtain ready-to-use indicators, and offers the possibility to drill down from high-level metrics to more detailed summaries. The offered indicators are calculated beforehand so that they are ready to use when the user needs them. The design is based on a set of different parallel processes to precalculate indicators. The application displays information related to tests, requests, samples, and turn-around times. The dashboard is designed to show the set of indicators on a single screen.

Results: DB4US was deployed for the first time in the Hospital Costa del Sol in 2008. In our evaluation we show the positive impact of this methodology for laboratory professionals, since the use of our application has reduced the time needed for the elaboration of the different statistical indicators and has also provided information that has been used to optimize the usage of laboratory resources by the discovery of anomalies in the indicators. DB4US users benefit from Internet-based communication of results, since this information is available from any computer without having to install any additional software.

Conclusions: The proposed methodology and the accompanying web application, DB4US, automates the processing of information related to laboratory quality indicators and offers a novel approach for managing laboratory-related information, benefiting from an Internet-based communication mechanism. The application of this methodology has been shown to improve the usage of time, as well as other laboratory resources.

(Interact J Med Res 2012;1(2):e16) doi:10.2196/ijmr.2126

KEYWORDS

Automation, laboratory; Medical Informatics Applications; Data Mining; Quality Indicators, Health Care

Introduction

These days, laboratories tend towards automation. The data they manage are structured around information systems that store not only purely analytical data, but also management data, such as samples, timestamps, and demographics. Nevertheless, there is a lack of ready-to-use management quality measures or indicators, which makes it very difficult to easily analyze the overall performance of the laboratory and thus to use the available information to make evidence-supported decisions [1].

The reagents used for laboratory analysis have a profound impact on the global economic balance of the laboratory. Clinical laboratory costs increase rapidly due to increased laboratory utilization and to inflationary trends [2]. In particular, the Hospital Costa del Sol laboratory expenditures increase about 10%-15% annually. The costs of this laboratory have an impact of 3-4% on the total hospital expenditures.

It has been pointed out that having accurate information on laboratory test costs effectively leads to reduction in hospital expenditure [3]. For that reason, an accurate estimation of the money spent on reagents in a given period is of great importance from an organizational point of view. In order for it to be reasonably accurate, a large amount of data on number of determinations, quality controls, or calibrations for different tests and analytical instruments has to be available beforehand. These data are already stored in the instruments and information systems of the laboratory, but not in an easily accessible fashion. Each data source stores a large amount of partially redundant data using a different format. The collection and analysis of all the necessary data represents a cumbersome, error-prone, and time-intensive process, and the estimation of the desired values has to make use of extrapolation of past trends and other indirect means. Such calculations involve several hours, or even days, of manual work, combining data from different sources.

Furthermore, interesting information such as patterns and trends can be hidden among the massive amounts of data distributed between different systems. For this reason, it would be of great use to have a system able to gain access to all the necessary data in the information systems, in order to automatically unify, process, and provide ready-to-use summaries to the laboratory manager. Indicators defined by the user would be readily available, allowing relevant performance trends to be seen at a glance and possibly revealing some aspects of the laboratory performance that could be optimized [4].

Decision support systems capable of providing such performance measures constitute a solution for this task, as explained by Power's article [5], which offers a review of the most important milestones in the history of decision support systems. In that work, decision support systems are presented as the result of the convergence of various technology threads, which started in the early 1960s. The focus shifted towards web-based analytical applications in the late 1990s, with the emergence of ad-hoc corporate intranets and web-based applications. The adoption of such information systems has been a key factor for quality improvement in medical centers such as the Legacy Good Samaritan Hospital in Portland, Oregon, the Ranking Medical Center in Brandon, Mississippi, and St. Mary's Health Care System in Athens, Georgia [6]. Some specific examples of successful decision support systems for clinical laboratories are QCIS [7], designed to identify relevant clinical information, and COAT [8], a system able to assess outcomes and measure performance by gathering, formatting, and abstracting data.

A suitable tool for providing quality measures is the management dashboard, an interface where complex and heterogeneous data from various sources are consolidated and displayed, in order to provide easy-to-read summaries of previously defined performance metrics. Dashboards are the result of the evolution and convergence of classical Decision Support Systems, Executive Information Systems, Data Warehouses, and Business Intelligence [9]. They represent a key tool to improve efficiency, accelerate decisions, and reduce oversights and errors in clinical practice [10]. We refer to the work of Kroch et al [11], where several dashboards across different hospitals are studied, showing the tool's effectiveness in quality management.

In this paper, we present a methodology based on a management dashboard providing performance measures for laboratories and a web application that illustrates this methodology, DB4US. This tool is the product of a project created as part of an ongoing collaboration between the University of Málaga, Hospital Costa del Sol of Marbella, and Siemens Healthcare Diagnostics, S.L.

Our goal was to design a methodology that would enable physicians to obtain information about the overall performance of the laboratory in the form of requests, tests, and turn-around times and to validate this methodology implementing a web application. As an easy-to-read summary of the data, we propose the use of a dashboard screen, which offers a high-level view of the most important indicators.

The main objective is thus the design of this methodology and the implementation and deployment of a dashboard-like application capable of summarizing metrics related to the number of tests, samples, and requests handled by the laboratory of the Hospital, as well as turn-around times. The

implementation of the project comprises the construction of a unified data warehouse that stores and consolidates multidimensional data from various information sources, offering an interface to obtain ready-to-use indicators, and a user-oriented front-end application that offers the possibility to drill down from high-level metrics to more detailed summaries.

Methods

The laboratory of the Hospital Costa del Sol is organized around the ADVIA LabCell system, with several analytic instruments: two ADVIA 2400, three ADVIA Centaur XP systems, and one IMMULITE 2000 Xpi system. The core of the information system of the laboratory is the ADVIA CentraLink data management system, which provides an integrated foundation for automating and consolidating data from the laboratory. Another important information source in the laboratory is Linemaster, which offers more detailed information about the samples' life-cycle and their related timestamps.

The Lab Cell System processes around 1200 patients daily and has a portfolio of 97 different tests (eg, biochemistry, immunochemistry, serology, allergy tests, etc.). The impact on the general budget of the laboratory is 50%, approximately. Note that our application uses only data from the automation chain. Data that are not provided by the aforementioned systems are out of the scope of the application described in this article.

It is necessary to access the different sources of data to achieve the objectives, specifically the ADVIA CentraLink system and Linemaster, to consolidate the obtained data and to calculate several statistical indicators to be concise, informative, and easy to read and interpret [12-14].

The methodology followed by DB4US works with up-to-date data, which means that processes have to be designed to regularly extract data from the sources and store that data in a data warehouse. As the application works with large amounts of data, these extraction processes have been carefully designed to be as efficient as possible with respect to execution time, as well as to memory and storage space.

The users of this application cannot retrieve all the detailed information stored in the database, but only summarized views or indicators, such that any given indicator can be immediately retrieved from the data warehouse. For that reason, the offered indicators have been calculated beforehand, so that they are ready to use when the user needs them. This necessity leads to the design of a set of different parallel processes to precalculate indicators, employing the consolidated data from the data warehouse.

A graphical user interface (GUI) offers easy access to the indicators. Its structure makes it clear how to move around the different sections of the statistics. Those are presented in an informative and interpretable way, using information representation techniques such as data charts and tables. For any given indicator, the GUI allows us to compare its value for different months and years. It has been made flexible and customizable, offered in several languages, and the association between instruments, groups of instruments, and parameters, as well as the user-defined thresholds, is configurable.

To provide a more detailed methodological and architectural vision, the description of DB4US is divided into two sections. The first section is devoted to describing detailed information about the specific indicators we are working with, and the other one to describing the architecture and some technical details of the implementation.

Indicators Description

We have implemented an application that displays three main kinds of information. The first one corresponds to tests, requests, and samples. The user can look up general information about requests and their associated samples. This allows the user to discover the ratio between these two parameters, showing how many requests have resulted in more than one sample. Parameter-based information can be also looked up, organized by instrument groups and specific parameters, and broken down into requested and uploaded tests, repetitions, quality controls, and calibrations. The second kind of information is related to turn-around times [15], showing global and detailed turn-around time curves for the samples, as well as percentiles. The third kind of information is captured in the dashboard which summarizes laboratory status and gives a global view of activity.

Tests, Requests, and Samples

The first batch of indicators is related to activity. The number of requests, tests, and samples are shown on a monthly basis, and the total quantity for the last years, as well as the accumulated quantity for the current year, is shown for comparison. The number of requested and reported tests is also shown. This number will be lower than the number of total tests performed because repetitions, quality control, and calibrations are not included. The ratio between requested and total performed tests is another interesting measure, since it allows the user to assess how much extra work has been carried out.

The number of samples can be grouped by sample origin, while the number of tests can be grouped by group of instruments and specific parameters. Tests can be subsequently broken down into uploaded, repeated, quality controls, and calibrations. The ratio between the real and requested tests can also be calculated for each parameter.

From the number of total tests performed, it is possible to obtain information about test-related costs. Test costs can be configured by the user and can be used in combination with the number of determinations in order to obtain the total costs, both overall and by instruments and parameters.

Turn-Around Times

The second batch of indicators is related to the turn-around times. One of the key indicators is the turn-around-time for the 50th and 90th percentiles, or for any other pair of percentiles that the user configures. Since turn-around times do not follow a Gaussian distribution, these percentiles are more informative than averages and variances. The percentiles can be grouped by origin since it is illustrative to compare how samples from different origins behave. Turn-around-times can be broken down into pre-analytical and analytical times.

Dashboard

The dashboard is the core of our information system. It is designed to show the following indicators on a single screen:

1. Requests progress: Represents the ratio between the requests for a given month and the requests for the previous month.
2. Tests progress: Represents the ratio between the tests for a given month and the requests for the previous month.
3. Tests/requests ratio: Represents the ratio between the real and the requested tests for each month.
4. Not uploaded tests: Represents the number of tests that have been requested, but not uploaded (eg, because they have not been validated yet).
5. Repetitions/uploaded ratio: Ratio of repetitions for each uploaded result. It is thus a measure of extra work.
6. Quality factor: The ratio of quality controls and calibrations is an indicator of how much reagent has been employed to ensure the quality of the analytical system.
7. Quality factor for different areas ADVIA 2400 system, ADVIA Centaur system, and IMMULITE system: Same as above, but grouped by specific instrument groups.
8. Samples/requests ratio. Represents how many samples are associated with each patient request on average.
9. Samples with turn-around time longer than 10 days: These samples are considered to be anomalous and have to be individually checked by the medical staff.
10. Percentiles 50 and 90 of turn-around time and average values.
11. Percentage of requests with a turn-around time longer than 12 hours: These requests are left for the next day; therefore, it can be considered as the ratio of requests that cannot be completed in the same day. A high number of requests finished after more than 12 hours shows that a problem has delayed the workflow (eg, a system failure).

The web application contains a configuration section where the aforementioned parameters can be modified. For example, the user can opt to monitor samples with turn-around times longer than 12 instead of 10 days, percentiles 95 instead of 90, or percentage of requests with a turn-around time longer than 8 instead of 12 hours.

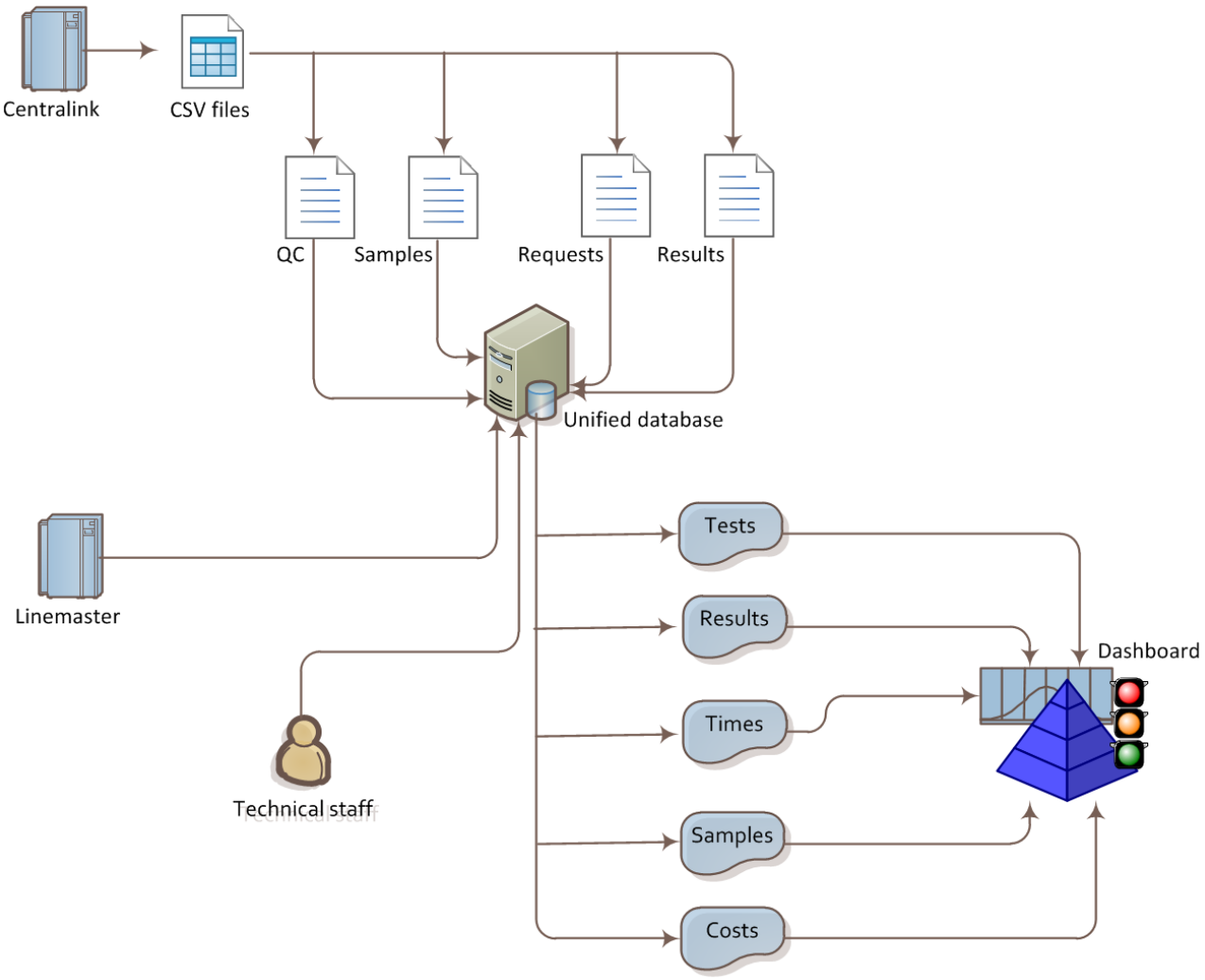
For each indicator, a normal range is defined. It is also configurable by the web application user, who has previously defined a minimum value v_{min} and a maximum value v_{max}. The values are shown, month-by-month, using a three-color code: Green, if the value lies between the defined validity range limit for that indicator; Orange, if the value is outside the limits but the distance to the nearest limit is less than $(v_{max} - v_{min})/2$; and Red, if the value is outside the limits and the distance to the nearest limit is greater or equal to $(v_{max} - v_{min})/2$.

System Description: Technologies and Architecture

The implemented application is based on Java technologies and uses Java EE-based web architecture, with several layers for data access, transformation, and visualization. An outline of the application structure is shown in Figure 1.

We have chosen to implement an information system from scratch instead of using or adapting an already existing business analytics system such as MicroStrategy Reporting Suite for a number of reasons. Already existing systems are typically based on OLAP technologies, which impose a given structure onto the underlying database. Our application is very intensive in terms of space, so the authors have deemed the extra effort of implementing a database from scratch worth the effort because a balance between space and time needed for queries can be more fine-tuned and adapted to the specific laboratory needs. Furthermore, our business layer provides a data representation that makes the indicators totally independent of the final representation, which helps to reuse data structures between different representation strategies, such as HTML tables, Flash graphics, XML data, or Excel files.

Figure 1. Application structure.



Application Architecture

We have implemented our project as a Java EE-based web application [16] that makes use of an external PostgreSQL database [17]. That means that the application has to be installed only once on a server and will be accessible for the final users on their web browsers, without the necessity of installing anything. Java technologies have been chosen because Java is a powerful, platform-independent language with a large number of third party libraries that offer a wide range of functionalities. We have implemented a user interface using the Java Server Faces [18] web framework and InfoSoft Global FusionCharts for graphical charting [19]. The web application relies on Glassfish v.2, an open source application server developed by Sun Microsystems. For the database management system, we have chosen PostgreSQL, an advanced open-source database system. The reasons for the technological decisions are further described in the Discussion section.

The core of our project is the central database, in which clean data from ADVIA CentraLink system and Linemaster is stored, together with several preprocessed summaries, ready to be served to the Java application. Two main database levels have been implemented: the first one stores raw data, which would be used to calculate indicators; the second level is the indicators level, in which the data is summarized.

Our application consists of two main components: a back-end, where calculations take place, and a user-oriented front-end. These two components are further explained in the next subsections.

Back-End

The back-end of our application contains modules that directly interact with the central database. This back-end is divided in two main modules. The first one is responsible for consolidating the data originating from the ADVIA CentraLink system and Linemaster databases, as well as from the manually entered data, such as calibrations. The second one is responsible for summarizing the data and for offering ready-to-use indicators to the web application. While the first module is dependent on the specific infrastructure of this laboratory, and thus has to be adapted for each laboratory, the second one is independent and can be used as is in other laboratories.

In order to adapt the first module to the specific infrastructure of the laboratory, the application needs information about which proprietary systems it is connected to and how the database of each of such systems has to be queried in order to obtain information to be analyzed. This is due to the fact that the

database structure of each proprietary system is different. This is accomplished by implementing a table that stores certain information (see Table 1).

Table 1. Fields in the data table.

Field	Meaning
Query identifier	A unique string identifier for each specific query (eg, number of determinations during the last month)
Proprietary system	The proprietary system to be queried (eg, Advia 2400)
SQL sentence	The SQL query that allows to extract the query defined by the query identifier from the database of the corresponding proprietary system

Additionally, for each proprietary system we need to specify connection parameters (IP address, port, user, and password). These parameters are configurable by the user.

Although the database of each proprietary system is different, it has to be able to provide at least the following parameters in order to be compatible with DB4US: requests; samples associated with each request; demographics associated with each request; tests associated with each sample; parameter, status and result of each test; and timestamps associated with the life-cycle of each test.

An important cross-component that affects these two modules is the scheduler, a piece of software that automatically triggers temporized data importations and summaries according to a previously defined temporal planning. We have defined a specific scheduling for this laboratory, but the application lets the user configure new scheduling plans. In our case, the plan runs weekly. Every weekend, data from the ADVIA CentraLink system is exported to CSV files. Once these files are available at the ADVIA CentraLink system main server, our back-end application parses them and loads the data into the central database, previously normalizing the relational data. Simultaneously, time-related data from Linemaster are imported into a different section of the central database. When these two tasks have been successfully completed, several timestamps, which comprise data from both the ADVIA CentraLink system and Linemaster, are computed. The CSV files are then compressed and backed-up. At this point, we have imported all the raw data that we need for building the indicators. For this calculation, several threads are launched simultaneously, each of them specializing in one specific indicator. All of these tasks, as well as the data import, are automatically stored; if the server goes down, the application automatically resumes the pending tasks. See Figure 2 for an outline of the main activities involved.

The back-end application provides another distinct module that serves the indicators. This module waits until an external request for a given indicator is received from another application, such as our web front-end. The necessary chunks of data required for calculating the indicator are retrieved from the database, and the indicator is serialized and returned to the calling application, which usually represents this data as a table or graphic.

Front-End

For the front-end, we have implemented another Java EE web-based application that acts as an interface between the back-end and the user. The web application provides such features as customization, security via authentication and authorization modules, internationalization, tabular, and graphical representations of the data, and exportation to other formats, such as Microsoft Excel.

To grant access to the users, a login and password are provided, which give them the permissions associated with a given role. For example, to configure the application, the user has to belong to an administrative role; this is not needed to simply see the data. Once the user has logged in, they can navigate through the application with the help of a menu and links.

The front-end acts interactively. The user's actions, such as clicks and selections, are transformed into requests to the back-end application, which, in turn, retrieves data from the database. The web application renders the received data into tables and graphics. For the tables, we have used JSF components from the Woodstock library along with an object-oriented representation of their rows. For the graphics, we have implemented a component that transforms the data returned by the back-end application into the XML format. A Servlet (a Java class that responds to http requests with XML code) is associated with each graphic. Along with these main components, we have implemented a number of helper classes for date-related data. We have used Java internationalization components in order to provide a multilingual interface. See Figure 3 for a screenshot of the dashboard.

Figure 2. Data loading and summarizing outline.

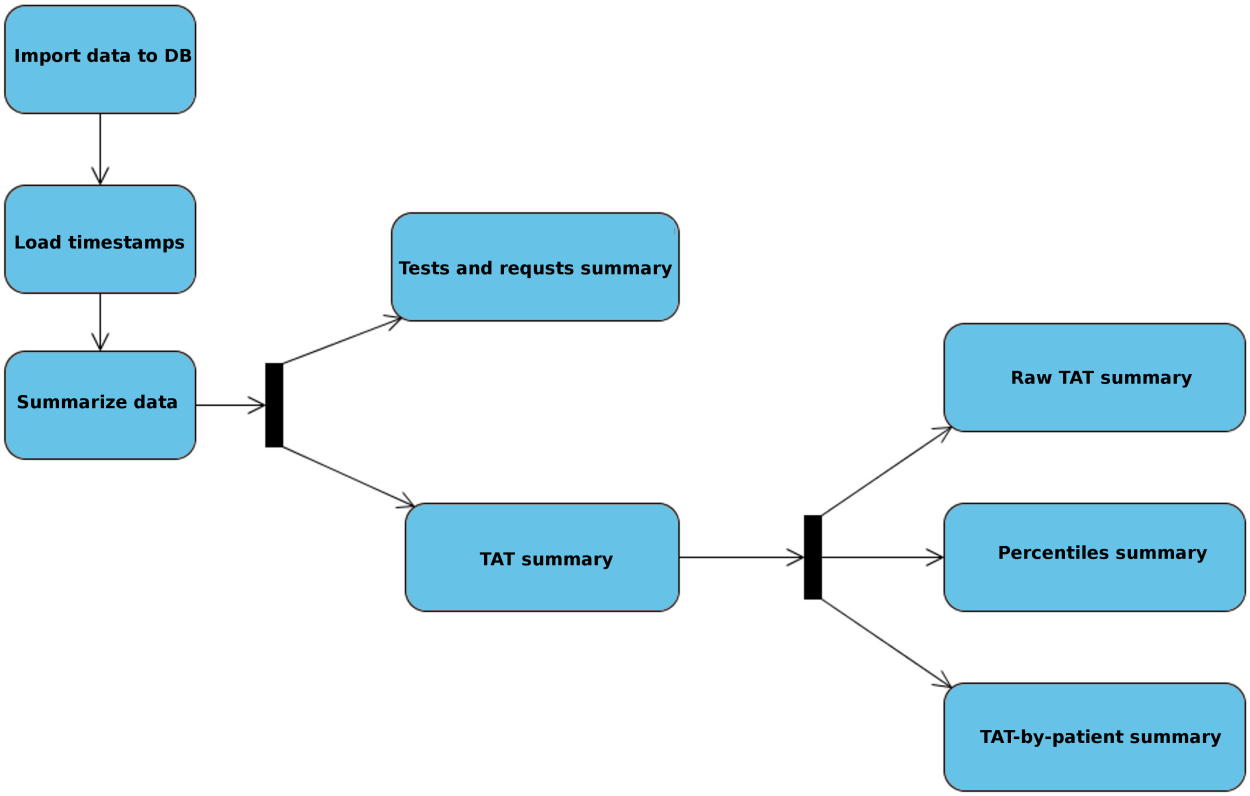


Figure 3. Dashboard screenshot.

Dashboard

Excel

Subscribe to indicator	Indicator	Objective	Apr 2009	May 2009	Jun 2009	Jul 2009	Aug 2009	Sep 2009	Oct 2009	Nov 2009	Dec 2009	Jan 2010	Feb 2010	Mar 2010
☐	Requests progress	0.0 - 4.0 %	-5.21	-2.05	-4.92	-3.33	-14.38	22.19	-1.22	3.22	-9.30	-0.99	14.32	6.57
☐	Tests progress	0.0 - 5.0 %	-6.86	0.87	-1.16	-0.94	-15.32	23.40	-4.23	7.97	-20.51	10.14	13.95	11.49
☐	Tests/requests	0.0 - 15.0	15.04	15.49	16.10	16.50	16.32	16.48	15.98	16.71	14.65	16.29	16.24	16.99
☐	Not uploaded	0.0 - 0.0	3	0	0	0	0	0	1	0	1	0	0	0
☐	Quality factor (all)	0.0 - 2.0 %	4.22	3.75	2.89	4.20	4.53	3.63	3.59	3.32	3.47	3.57	3.50	3.80
☐	Quality factor (Bioch.)	0.0 - 2.0 %	3.33	2.85	2.14	3.07	3.46	2.66	2.54	2.40	2.46	2.62	2.74	3.02
☐	Quality factor (Centaur 1, 2)	0.0 - 10.0 %	12.10	11.65	9.61	13.23	13.30	11.78	12.56	11.18	11.77	11.89	10.23	11.29
☐	Samples/requests	0.0 - 1.2	1.18	1.21	1.21	1.22	1.21	1.21	1.20	1.21	1.19	1.22	1.21	1.23
☐	Average TAT	0.0 - 300.0 min	230.30	254.96	271.06	225.55	324.11	234.22	541.01	327.71	774.04	354.14	333.82	315.01
☐	Percentile 50 (TAT)	0.0 - 180.0 min	155.0	176.0	178.0	142.0	189.0	164.0	210.0	194.0	217.0	170.0	194.0	225.0
☐	Percentile 90 (TAT)	0.0 - 360.0 min	343.0	364.0	389.0	335.0	468.0	352.0	1556.0	425.0	1480.0	350.0	449.0	429.0
☐	Percentage of requests with TAT > 12 h.	0.0 - 1.0 %	3.17	3.87	3.66	2.52	8.63	2.43	16.54	5.92	0.0	0.0	5.19	0.0
☐	Requests with TAT > 10 days	0.0 - 0.0	0	0	2	0	0	0	0	0	0	0	0	0

Results

DB4US was first deployed in the Hospital Costa del Sol in 2008 and has been in continuous use ever since. The system has been evolving over time, in order to cope with new necessities and ideas. The main benefits of the deployment of the system are, on the one hand, the reduction of the time needed for the calculation of indicators, and on the other, access to a global view of the performance of the laboratory that allows identification of aspects that can be optimized.

The automation of the indicators' retrieval means that the time delay is heavily reduced. For example, before the deployment of our application, an Excel spreadsheet for calculating the money spent on reagents was manually compiled, with costs and usage data for each type of reagent. This calculation involved several hours of repetitive and error-prone manual work. Thanks to our application, these summaries can be automatically calculated over the weekend, giving the user access to a ready-to-use summary of the data, saving hours of manual work and ensuring error-free data.

Another important advantage of this application is the ability to provide the final user with access to a global perspective of

the laboratory performance in terms of reagent usage and turn-around times. This means that the user can easily detect trends at a glance. This new possibility has an additional benefit: if there is something anomalous in the indicators, the user can take action immediately, which gives them a chance to improve some aspects of the process. This was shown in the discovery of the anomalous behavior of the direct bilirubin (BD) test requests. In early 2009, an unlikely high value of the total number of BD requests was spotted in the corresponding indicator provided by our application. The reason for this value turned out to be an incorrectly configured trigger rule in the Centralink system, which triggered an unwanted BD request for any new registered sample. This discovery allowed the laboratory staff to disable this anomalous rule. Therefore, an important fact that was somehow hidden among a vast amount of information was revealed by the automated processing of information carried out by our system. Following this discovery, the number of requests for this rule diminished by an average of 2297.1 requests (taking the previous and succeeding 6 months); ie, a 77.89% less reagent usage for this test type. See Table 2.

Table 2. Evolution of direct bilirubin test requests.

Month	Number of requests
August ’08	2255
September ’08	3244
October ’08	3094
November ’08	2729
December ’08	3354
January ’09	2129
February’ 09	621
March ’09	664
April ’09	594
May ’09	634
June ’09	643
July ’09	692

Discussion

Nowadays, clinical laboratories face numerous challenges such as health care reform, cost pressures, tight laboratory regulations, a growing complexity of diagnostic tests, increased workload, and continuous shortening of turn-around times. At the same time, the users have very high expectations of laboratory services. In order to overcome the above-mentioned obstacles, a clinical laboratory needs to incorporate creative solutions and adapt to change. That is why the introduction of artificial intelligence by means of expert systems has gained an important place in the automation process [20]. To convince our administrators, we need new comprehensive tools for providing indicators of activity and quality. In this work, we aimed to develop a general methodology illustrated by the implementation of an application that integrates the computation of quality measures to easily analyze the overall performance of a laboratory, as well as automatically detect anomalies.

Principal Results

During the analysis phase of this project, it was decided to deploy the software as a web rather than a desktop application. The use of Internet and web applications has several advantages: no disk space is required by the client, cross-platform compatibility is provided, and when installing new features, only the server side has to be upgraded. Furthermore, the final users do not have to install anything on their computers. The main drawback of this solution is that the application cannot be accessed in case of a network failure.

As for implementing a web application, we had a large number of available technologies to choose from. Two of the most widely used technologies for implementing complex web applications are Microsoft .NET and Java EE. There are several similarities between these two technologies: both run over a virtual-machine environment and both provide a large library of solutions for GUI construction, database access, persistence, caching, or remote method invocation, among others. Java has the advantage of providing cross-platform compatibility and of being open-sourced software. Furthermore, there exist a large number of external Java libraries for data analysis, such as Weka [21]. These reasons have prompted us to adopt Java EE as our main technology.

There are quite a few web application frameworks that work under Java, many of them following the model-view-controller pattern to separate the data model and business rules from user interface. Examples of these frameworks are Apache Struts, Spring and Sun Java Server Faces (JSF), all of which provide features for security, URL mapping, templating, and Ajax support. We have decided to use Sun JSF, which provides Facelets, a simple, efficient, and powerful view description language. Java EE-based web applications run on an application server. For our purposes, we use Oracle Glassfish, which is based on Apache Tomcat.

As one of the key goals of our application is to provide an easily comprehensible view of data, the use of adequate data visualization components is of utmost importance. InfoSoft Global FusionCharts is a library to automatically generate dynamic Flash charts to be embedded in web applications. It is noteworthy that FusionCharts pioneered the use of Adobe Flash for statistics charting [22].

In our application, the management of data is of vital importance, since we need to efficiently access a very large quantity of data. Thus we need to use a database management system (DBMS), such as Oracle MySQL, Microsoft SQL Server, or PostgreSQL, which has been our selected DBMS. PostgreSQL is open-source software, which is not controlled by any single company. PostgreSQL is a fairly powerful DBMS, offering unlimited database size, several types of indexes, and rich database capabilities.

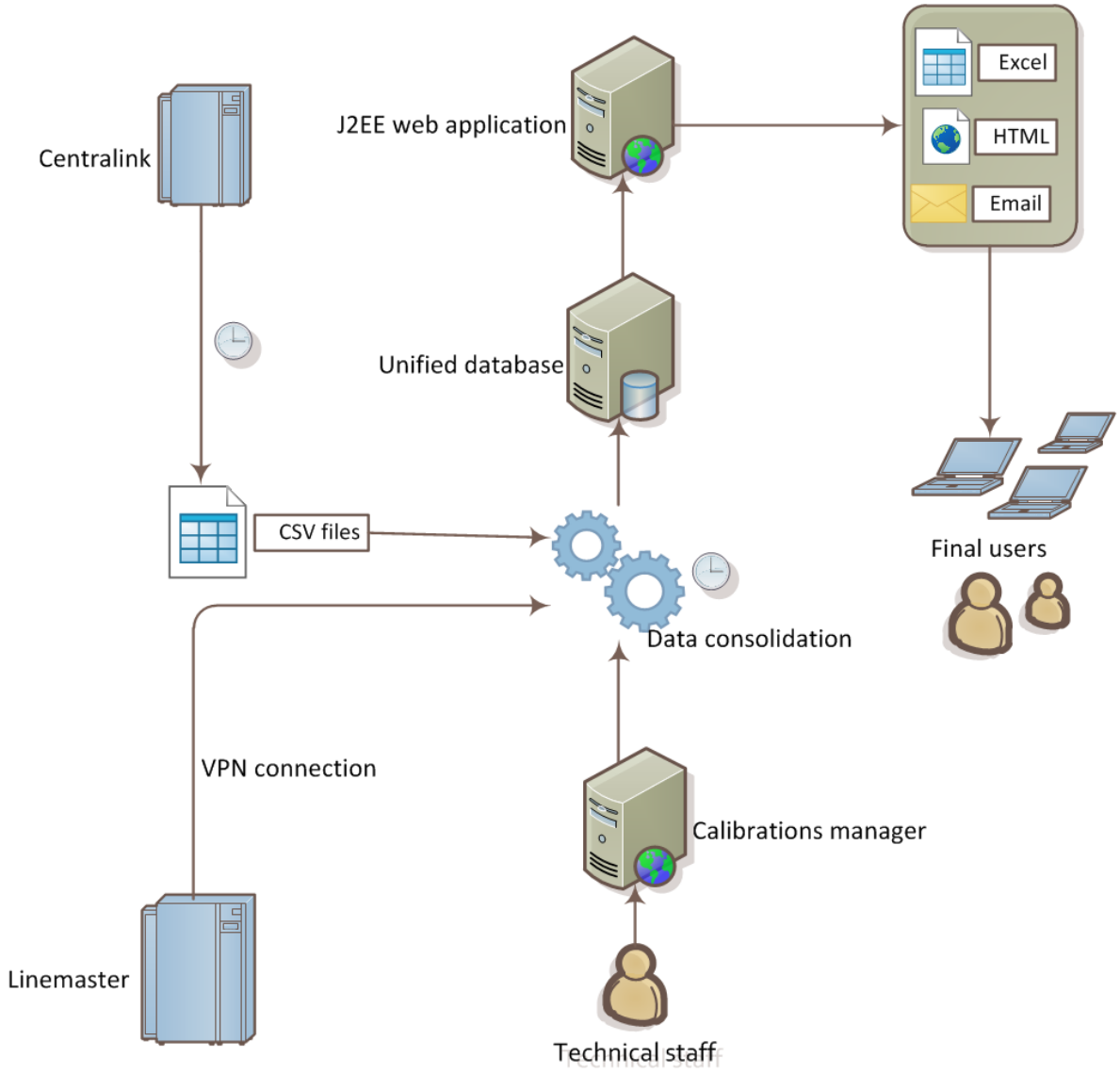
Difficulties

During the deployment of the application, we faced several difficulties respecting raw data access. Specifically, we did not have real-time access to the ADVIA CentraLink database or to the Linemaster database, since the server was in a different subnet. The Linemaster problem was resolved by deploying a virtual private network to gain access to the Linemaster database. To do this, we installed and configured a VNP server on an intermediate machine, with access to both Linemaster and the server where our application is deployed, and a client in our server. We used OpenVPN for the VPN implementation [23].

The ADVIA CentraLink system problem was more difficult to solve. This system offers the possibility of exporting its database into CSV files, which in turn can be parsed by our application to retrieve the data, albeit indirectly. Until recently, the only way to trigger this data exportation was by hand, which meant that the exportation was not schedulable. Nevertheless, an AutoHotkey script was developed to simulate the movements of the mouse and keystrokes of the keyboard. AutoHotkey is a free and open-source automation software frequently used to simulate the actions of the user with the keyboard and mouse. A recent upgrade of the ADVIA CentraLink system client enables the creation of scheduled data exports; as a consequence, the AutoHotkey scripts are no longer needed.

A further difficulty was the synchronization of the clocks of the components (the server where our application runs, the ADVIA CentraLink system and Linemaster server), in order to calculate meaningful turn-around times that involve timestamps from different sources. Finally, there are data that cannot be exported by the instruments, and thus cannot be provided by the ADVIA CentraLink system, such as the calibrations, which are needed to calculate the correct number of tests. In this case, we have implemented a subsidiary web application, with which the technical staff can manually insert the data. Figure 4 shows how the application is deployed.

Figure 4. System deployment.



Conclusions

DB4US is an application using a methodological approach that automates the processing of information related to laboratory quality indicators presented and implemented using a central data warehouse and computing ready-to-use measures. The use of this methodology for laboratory information management has a positive impact on the laboratory, improving the usage of time as well as other laboratory resources.

The introduction of this methodology has allowed access to information that otherwise would be hidden among the vast quantity of data stored in the instruments. This gives, in some cases, the key to optimize some aspects of the laboratory performance and, in all cases, access to arbitrarily complex ad-hoc indicators in a reduced amount of time, benefiting from the advantages of Internet technologies and web-based interfaces. Furthermore, the system can be used to optimize processes and reduce costs by discovering anomalous behaviors.

Acknowledgments

We would like to thank the staff of the Hospital Costa del Sol laboratory for their collaboration.

This work has been partially supported by the SESAAME project (code TIN2008-06582-C03-03) of the MEC, Spain, and by the contract 8.06/5.47.3058 between Universidad de Málaga and E.P. Hospital Costal del Sol. José María Carmona-Cejudo is supported by grant AP2009-1457 from the Spanish government.

Conflicts of Interest

None declared.

References

1. Price CP. Benchmarking in laboratory medicine: are we measuring the right outcome? Benchmarking: An International Journal 2005;12(5):449-466. [doi: 10.1108/14635770510619375]
2. Robinson A. Rationale for cost-effective laboratory medicine. Clin Microbiol Rev 1994 Apr;7(2):185-199 [FREE Full text] [Medline: 8055467]
3. Ellemdin S, Rheeder P, Soma P. Providing clinicians with information on laboratory test costs leads to reduction in hospital expenditure. S Afr Med J 2011 Oct;101(10):746-748. [Medline: 22272854]
4. Tierney WM. Improving clinical decisions and outcomes with information: a review. Int J Med Inform 2001 Jun;62(1):1-9. [Medline: 11340002]
5. Power DJ. A brief history of decision support systems. 2007. URL: http://dssresources.com/history/dsshistory.html [accessed 2011-03-10] [WebCite Cache ID 5x59qCpHX]
6. Silow-Carroll S, Alteras T, Meyer JA. Hospital Quality Improvement. Strategies and lessons from US hospitals. 2007. URL: http://dssresources.com/history/dsshistory.html [accessed 2011-03-10] [WebCite Cache ID 5x59qCpHX]
7. Zeng Q, Cimino JJ, Zou KH. Providing concept-oriented views for clinical data using a knowledge-based system: an evaluation. J Am Med Inform Assoc 2002;9(3):294-305 [FREE Full text] [Medline: 11971890]
8. D'Avolio LW, Bui AA. The Clinical Outcomes Assessment Toolkit: a framework to support automated clinical records-based outcomes assessment and performance measurement research. J Am Med Inform Assoc 2008;15(3):333-340 [FREE Full text] [doi: 10.1197/jamia.M2550] [Medline: 18308990]
9. Ezzelle J, Rodriguez-Chavez IR, Darden JM, Stirewalt M, Kunwar N, Hitchcock R, et al. Guidelines on good clinical laboratory practice: bridging operations between research and clinical research laboratories. J Pharm Biomed Anal 2008 Jan 7;46(1):18-29 [FREE Full text] [doi: 10.1016/j.jpba.2007.10.010] [Medline: 18037599]
10. Eckerson W. Performance dashboards: measuring, monitoring, and managing your business. Hoboken, N.J: John Wiley; 2006.
11. Egan M. Clinical dashboards: impact on workflow, care quality, and patient safety. Crit Care Nurs Q 2006;29(4):354-361. [Medline: 17063102]
12. Kroch E, Vaughn T, Koepke M, Roman S, Foster D. Sinha S: Hospital Boards and Quality Dashboards. J Patient Saf 2006;2(1):10-19.
13. Ogrinc G, Mooney SE, Estrada C, Foster T, Goldmann D, Hall LW, et al. The SQUIRE (Standards for QUality Improvement Reporting Excellence) guidelines for quality improvement reporting: explanation and elaboration. Qual Saf Health Care 2008 Oct;17 Suppl 1:i13-i32 [FREE Full text] [doi: 10.1136/qshc.2008.029058] [Medline: 18836062]
14. Niland JC, Rouse L, Stahl DC. An informatics blueprint for healthcare quality information systems. J Am Med Inform Assoc 2006;13(4):402-417 [FREE Full text] [doi: 10.1197/jamia.M2050] [Medline: 16622161]
15. Zarbo RJ, Jones BA, Friedberg RC, Valenstein PN, Renner SW, Schifman RB, et al. Q-tracks: a College of American Pathologists program of continuous laboratory monitoring and longitudinal tracking. Arch Pathol Lab Med 2002 Sep;126(9):1036-1044. [doi: 10.1043/0003-9985(2002)126<1036:QT>2.0.CO;2] [Medline: 12204052]
16. Steindel SJ, Howanitz PJ. Physician satisfaction and emergency department laboratory test turnaround time. Arch Pathol Lab Med 2001 Jul;125(7):863-871. [doi: 10.1043/0003-9985(2001)125<0863:PSAEDL>2.0.CO;2] [Medline: 11419969]
17. Jendrock E, Ball J, Carson D, Evans I, Fordin S, Haase K. For Sun Java Application Server 9. The Java EE 5 Tutorial URL: http://download.oracle.com/javaee/5/tutorial/doc/ [accessed 2011-03-10] [WebCite Cache ID 5x59vQMu1]
18. The PostgreSQL Global Development Group. 3. PostgreSQL 8 URL: http://www.postgresql.org/docs/8.3/static/index.html [accessed 2011-03-10] [WebCite Cache ID 5x5A9sfTS]
19. Sun Developer Network. Available at. JavaserverFaces Techology URL: http://www.oracle.com/technetwork/java/index.html [accessed 2012-11-08] [WebCite Cache ID 6C1rPBTcD]
20. InfoSoft Global. Available at. FusionCharts v3 Online Documentation URL: http://www.oracle.com/technetwork/java/javaee/javaserverfaces-139869.html [accessed 2011-03-10] [WebCite Cache ID 5x5ADVlMF]
21. Van Hoof V, Wormek A, Schleutermann S, Schumacher T, Lothaire O, Trendelenburg C. Medical expert systems developed in j.MD, a Java based expert system shell: application in clinical laboratories. Stud Health Technol Inform 2004;107(Pt 1):89-93. [Medline: 15360781]
22. Hall M, Frank E, Holmes B, Pfahringer B, Reutemann P, Witten IH. The WEKA Data Mining Software: An Update. ACM SIGKDD Explorations Newsletter 2009;11(1):10. [doi: 10.1145/1656274.1656278]
23. OpenVPN Technologies. Available at. OpenVPN documentation URL: http://openvpn.net/index.php/open-source/documentation.html [accessed 2011-03-10] [WebCite Cache ID 5x5AH4tIW]

Edited by G Eysenbach; submitted 30.03.12; peer-reviewed by S Prieto Menchero; comments to author 30.07.12; revised version received 10.08.12; accepted 21.09.12; published 14.11.12.

Please cite as:
Carmona-Cejudo JM, Hortas ML, Baena-García M, Lana-Linati J, González C, Redondo M, Morales-Bueno R
DB4US: A Decision Support System for Laboratory Information Management
Interact J Med Res 2012;1(2):e16
URL: http://www.i-jmr.org/2012/2/e16/
doi:10.2196/ijmr.2126
PMID:23608745

Original Paper

Development, Implementation, and Evaluation of a Telemedicine Service for the Treatment of Acute Stroke Patients: TeleStroke

Carlos Parra[1*], BEc; Francisco Jódar-Sánchez[1*], MSc; M. Dolores Jiménez-Hernández[2*], MD; Eduardo Vigil[3*], BMed; Alfredo Palomino-García[2*], BMed; Francisco Moniche-Álvarez[2*], BMed; Francisco Javier De la Torre-Laviana[2*], BMed; Patricia Bonachela[4*], BEng; Francisco José Fernández[5*], BEng; Aurelio Cayuela-Domínguez[6*], MD; Sandra Leal[4*], BEng

[1]Virgen del Rocío University Hospital, Technological Innovation Group, Seville, Spain
[2]Virgen del Rocío University Hospital, Neurological Department, Seville, Spain
[3]Virgen del Rocío University Hospital, Clinical Documentation Service, Seville, Spain
[4]Virgen del Rocío University Hospital, Research, Development and Innovation Department, Seville, Spain
[5]Everis, Research and Technological Development Group, Seville, Spain
[6]Seville District, Seville, Spain
[*]all authors contributed equally

Corresponding Author:
Carlos Parra, BEc
Virgen del Rocío University Hospital
Technological Innovation Group
Avda. Manuel Siurot, s/n
Centro de Documentación Clínica Avanzada
Seville, 41013
Spain
Phone: 34 955013662
Fax: 34 955013310
Email: carlos.parra.sspa@juntadeandalucia.es

Abstract

Background: Health care service based on telemedicine can reduce both physical and time barriers in stroke treatments. Moreover, this service connects centers specializing in stroke treatment with other centers and practitioners, thereby increasing accessibility to neurological specialist care and fibrinolytic treatment.

Objective: Development, implementation, and evaluation of a care service for the treatment of acute stroke patients based on telemedicine (TeleStroke) at Virgen del Rocío University Hospital.

Methods: The evaluation phase, conducted from October 2008 to January 2011, involved patients who presented acute stroke symptoms confirmed by the emergency physician; they were examined using TeleStroke in two hospitals, at a distance of 16 and 110 kilometers from Virgen del Rocío University Hospital. We analyzed the number of interconsultation sheets, the percentage of patients treated with fibrinolysis, and the number of times they were treated. To evaluate medical professionals' acceptance of the TeleStroke system, we developed a web-based questionnaire using a Technology Acceptance Model.

Results: A total of 28 patients were evaluated through the interconsultation sheet. Out of 28 patients, 19 (68%) received fibrinolytic treatment. The most common reasons for not treating with fibrinolysis included: clinical criteria in six out of nine patients (66%) and beyond the time window in three out of nine patients (33%). The mean "onset-to-hospital" time was 69 minutes, the mean time from admission to CT image was 33 minutes, the mean "door-to-needle" time was 82 minutes, and the mean "onset-to-needle" time was 150 minutes. Out of 61 medical professionals, 34 (56%) completed a questionnaire to evaluate the acceptability of the TeleStroke system. The mean values for each item were over 6.50, indicating that respondents positively evaluated each item. This survey was assessed using the Cronbach alpha test to determine the reliability of the questionnaire and the results obtained, giving a value of 0.97.

Conclusions: The implementation of TeleStroke has made it possible for patients in the acute phase of stroke to receive effective treatment, something that was previously impossible because of the time required to transfer them to referral hospitals.

(Interact J Med Res 2012;1(2):e15) doi:10.2196/ijmr.2163

KEYWORDS

Telemedicine; Standardization; Stroke; Fibrinolysis

Introduction

The World Health Organization predicts that by 2030 cerebrovascular disease will be the second leading cause of death in the world; the percentage of deaths due to this disease will increase from 9.7% in 2004 to 12.1% in 2030 [1]. In Andalusia in 2007, cerebrovascular diseases presented a crude mortality rate of 86 per 100,000 inhabitants, 12% higher than that for Spain as a whole [2]. It has been demonstrated that both mortality and disability are reduced when these patients are treated by teams specially trained in administering treatment for this type of disease [3]. Therefore, specialized teams are required at referral hospitals (RH) that can respond to consultations from other hospitals without specialist neurologists. Interhospital networks based on urgent transfer have been shown to only partially ensure geographical equity in access to quality medical interventions at the cost of many unnecessary transfers [4].

Health care service based on telemedicine can reduce both physical and time barriers in stroke treatments. Moreover, this service connects centers specializing in stroke treatment with other centers and practitioners, thereby increasing accessibility to neurological specialist care and fibrinolytic treatment.

The principal competitive advantage of the telemedicine system over other similar solutions is its use of communications standards, providing the system with greater flexibility and interoperability [5]. Another advantage of the system is that its development has been directed towards the care process, and expert health personnel have been involved in the design and definition of requirements.

Previous studies used real-time video conference systems for acute stroke care to evaluate clinical neurological status with the National Institute of Health Stroke Scale (NIHSS) [6-8] and computerized tomography (CT) images [9-16]. The evolution of this technology has led telemedicine to be considered as an essential method for the diagnosis and treatment of patients suspected of having had a stroke. Communications via Internet and mobile phone may be the best means of access for all doctors treating these patients without requiring the patients' transfer [17]. This method of work is clearly beneficial to patients, their environment, and health institutions [18]. It has been shown that acute phase treatments can be applied without having to travel long distances and in the absence of resources and high-cost complex structures, which are probably not available at all points of care for these patients [19]. Consequently, it is possible for a team of stroke experts to be available 24 hours a day, at any hospital, and at relatively low cost [20].

Previous Experience in Spain

There are two studies on previous telemedicine applications for the treatment of stroke patients in Spain at the Vall d'Hebron University Hospital [11] in Barcelona and Son Espases University Hospital [16] in Palma de Mallorca. However, these experiences did not include interoperability capabilities with the Electronic Health Record, which is a central issue for the continuity of care [21].

Current Study

This paper describes the development, implementation, and evaluation of a care service based on telemedicine for the emergency treatment of patients with acute stroke at Virgen del Rocío University Hospital in Seville, a fully equipped stroke center.

Methods

Development

The development of the project involves a health care process modeling and development system.

TeleStroke Process Model

Business process modeling (BPM) is a methodology that has been applied before in this area [22]. In fact, it has been used in the health care sphere for years, particularly in the field of information system design. Use of this methodology in health care processes continues to become increasingly common [23-27].

In order to model the treatment process for TeleStroke, we chose business process modeling notation (BPMN), which provides a standard language permitting business process modeling in a workflow format. Its main goal is to provide a standard notation that will be easily read and understood by everyone involved in and affected by the business (ie, the stakeholders). These stakeholders include the users themselves who carry out the process, business analysts who define and redefine processes, technical developers responsible for implementing processes, and business managers and administrators who monitor and direct processes [28].

Modeling began with an exhaustive study of the documentation related to this condition and the contributions from the group of stroke specialists. Subsequently, this information was processed and made into a BPM model using BPMN.

TeleStroke System

The system has been designed to meet the specifications of the World Health Organisation program, eHealth for Health Care Delivery (eHCD). The eHCD states that eHealth technologies must be adapted and applicable to the different health systems [29] in accordance with their possibilities and resources, directed towards specific objectives, complying with international standards and in an established manner. In this way, the international standard Health Level Seven (HL7) [30] is used for the format of data and exchange of information and the protocol, digital imaging, and communication in medicine (DICOM) [31] as the internationally recognized standard for

the exchange of medical images. Following the recommendations of the eHCD, we have drawn on the experience of specialist personnel at the RH to decide which functions must be supported by the system to enable the diagnosis and treatment of stroke patients.

The interoperability scenario is supposed to enable the automation of communication, making the system responsible for collecting patient data and information, medical records, admission details, laboratory data, and images from the picture archiving and communication systems (PACS), etc., to be sent to the RH. One essential and critical requirement in the interoperability of the systems is the traceability of information.

The stages of the interoperability scenarios involved in TeleStroke are as follows:

a) Stage 1: Requesting hospital. The patient arrives at the requesting hospital, normally through the emergency department. The patient is examined at the hospital, suspected as a stroke case, and as a result, the doctor calls the TeleStroke Medical Station.

b) Stage 2: Interconsultation request. The emergency doctor submits an interconsultation request from the TeleStroke Medical Station to the RH. This, in turn, becomes a notification of incoming interconsultation to the mobile phone of the on-call neurologist, who goes to the medical station to collect the patient's data and accept the interconsultation.

c) Stage 3: TeleStroke interconsultation. Once the interconsultation has started, the neurologist may ask for images via videoconference, request further details about the patient, or send instructions to be carried out in the examination of the patient.

d) Stage 4: Closure and record. On conclusion of the examination, the specialist closes the interconsultation, providing details of diagnosis and treatment, so that the doctor at the requesting hospital can complete the information required in the Stroke Record System. The system architecture consists of two different elements: the TeleStroke Medical Station and communication between medical stations.

Implementation of Platform

The RH is the Virgen del Rocío University Hospital, which has an Electronic Health Record, called SIDCA, enabling the management of information and knowledge. Two hospitals were selected as requesting hospitals: San Juan de Dios (SJDH) and José María Díaz Domínguez (JMDDH). Work on the development of the platform began with a study of the infrastructure existing in the participating centers responsible for consultation with the RH. Table 1 shows the features of the information systems installed in both hospitals.

The system is based fundamentally on communication between the requesting hospital where the patient is located and the RH where the specialist group is based.

Table 1. Information systems at the two requesting hospitals.

Features	San Juan de Dios Hospital	José María Díaz Domínguez Hospital
HIS[a]	HIS ACTICX Telvent Interactive 4.1. HL7 communication. HL7 communication.	Aurora from SIEMENS, version 3.0.78. HL7 communication. HL7 communication.
LIS[b]	OpenLAB from Icon Media Lab. HL7 communication. HL7 communication.	Omega 2000 from ROCHE version 2.02.00.b. HL7 communication. HL7 communication.
RIS[c]	PHILIPS. HL7 communication. HL7 communication.	Telvent GESIR 3.0. HL7 communication.
PACS[d]	PACS from PHILLIPS. DICOM communication. DICOM communication.	PACS from Telvent. DICOM communication. DICOM communication.
Security and authentication	Firewall. They currently have an Active Directory	Firewall. Rt domain.
Communications	Symmetric communication with Corporate Network through MACROLAN, 2 Mbps	Symmetric communication with Corporate Network via 100Mb Optical Fibre MetroLAN and 2MB Reinforced Copper MetroLan

[a] Hospital Information System.

[b] Laboratory Information System.

[c] Radiology Information System.

[d] Picture Archiving and Communication Systems.

Evaluation

Study Design

This prospective observational study was conducted from October 2008 to January 2011 with patients who presented acute stroke symptoms. Patient inclusion criteria were: less than 4.5 hours after onset of stroke, age between 18 and 80 years (individualized decision in patients over 80 years according to baseline), and NIHSS scale between 5 and 24. They were examined using TeleStroke in SJDH and JMDDH, which are 16 and 110 kilometers respectively from Virgen del Rocío University Hospital. These hospitals did not have a fully equipped stroke center and patients with acute stroke symptoms could not be treated with fibrinolysis before the introduction of the TeleStroke system.

The Ethics Committee at Virgen del Rocío University Hospital approved this study.

Outcome Measures

We analyzed the number of interconsultation sheets, the percentage of patients treated with fibrinolysis, the time from onset of stroke symptoms to admission ("onset-to-hospital"), time from admission to fibrinolysis ("door-to needle"), time from admission to CT image, and time from onset of stroke symptoms to fibrinolysis ("onset-to-needle"). We also analyzed the reason why patients were not treated with fibrinolysis.

To evaluate medical professionals' acceptance of the TeleStroke system, we developed a web-based questionnaire using a Technology Acceptance Model (TAM) [32]. The TAM is a model based on the intended use of new technologies; it was developed to explain and predict the acceptance of information technology by potential users. The methodology has been used in other telemedicine studies [33].

The original TAM model included two dimensions to determine intended use: perceived usefulness and ease of use. For our study, we extended the original model by including two more dimensions: social norms and enabling conditions. Thus, in the end we evaluated five dimensions. Items used in the questionnaire were chosen from the literature [34]. The validity of the questionnaire's content was evaluated by a panel of experts consisting of doctors and experts in health technology assessment. Some of the questionnaire items were modified after the review of the experts. Under these conditions, we developed the questionnaire in a web form with 27 questions whose items were sorted randomly to avoid bias. In addition to the scores for questions, we used a ten-point Likert scale: one point for strongly disagree and ten points for totally agree. Once we had created the questionnaire, it was emailed to neurologists at Virgen del Rocío University Hospital, and family physicians, internists, and intensivists at SJDH and JMDDH.

Data Analysis

Descriptive analysis of data used absolute (n) and relative (%) frequencies for qualitative variables and mean ± standard deviation (SD) for quantitative variables. We measured the reliability of the instrument by calculating Cronbach alpha.

Results

Development

TeleStroke Process Model

The process model is universally considered to be the first step in analyzing and improving processes. It defines a model that contains all relevant characteristics of the actual process and makes analyzing that process possible. The process modeling was carried out based on acute-phase stroke treatment guides and protocols and the experience of specialists working with this condition (see Multimedia Appendix 1).

TeleStroke System

The TeleStroke Medical Station is responsible for collecting the necessary information from the departmental applications (Hospital Information System, Radiology Information System, Laboratory Information System and PACS) and sending it to the RH. In accordance with the HL7 standard, the information is collected using batch and online procedures. The user interface in the TeleStroke Medical Station has the emergency episode (Figure 1) as a central element.

The user interface has the following information items:

1) The interconsultation sheet. Through this sheet the doctor determines when the health personnel at the requesting hospital may consult the RH. It includes the information required by the specialist group to confirm diagnoses. It includes the following elements:

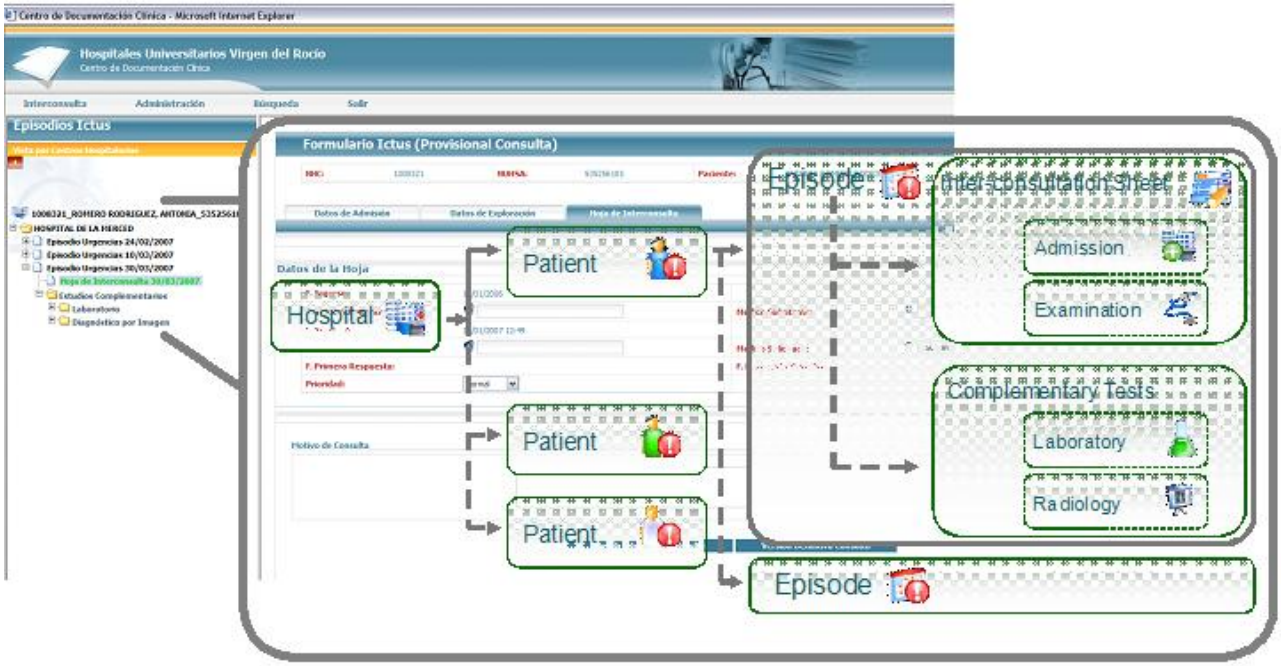
a. Examination details: information about the "in situ" examination of the patient. This includes vital signs (heart rate, body temperature, etc.), personal history (stroke, heart disease, hypertension, etc.), inclusion and exclusion criteria, scales (NIHSS), complementary tests (hemogram, electrocardiogram, etc.), and complementary details.
b. Admission details: details of patient's admission to hospital.
c. Consultation details: record of the consultation with the RH, including replies, assessments, and diagnoses of the group of specialists at the RH.

2) Complementary data consist of the following:

a. Laboratory data: study details (request, date, etc.) and test results, organized by section, chapter, and test.
b. Radiology data: study details (request, date, etc.) and examination (date, results, etc.) together with CT images.

3) Navigation tree is based on the data hierarchy from the medical station.

Figure 1. Data hierarchy from TeleStroke medical station.



Implementation

Communication between requesting hospitals and the RH is via Andalusia regional government's Corporate Network, which has been certified as a safe network. The client equipment that serves as TeleStroke Medical Station is located in a reserved area in the requesting hospital. At this point, we need to highlight the three different elements in communications between medical stations:

1) Transfer of episode and interconsultation sheet. The interconsultation sheet shared between both hospitals has 4 possible statuses:

a. Provisional consultation: the user of the TeleStroke Medical Station saves details of the patient's examination but has not yet consulted the specialist neurologist group.
b. Definitive consultation: details and consultation are sent to RH (SIDCA).
c. Provisional reply: the user of the medical station at SIDCA has saved a reply to the consultation but has not yet sent it to the requesting hospital.
d. Definitive reply: the reply is sent to the requesting hospital.

During the lifecycle of the interconsultation sheet and its complementary data, data modifications may arise that must be shared with the requesting hospital and the RH.

2) Processing, sending and displaying images. Images are collected from DICOM servers. Given the basic requirement of speed in image transmission in order to obtain an urgent diagnosis, images are compressed in accordance with the general-purpose image compression standard. JPEG [35] is also used to compress images automatically. The DICOM images enable a series of display functions that are used by health personnel and have also been developed for JPEG images. Figure 2 shows the CT image display including both the TeleStroke Medical Station and the Medical Station at the RH for displaying PACS images compressed using the JPEG standard. All image processing is performed in the web application itself without requiring installation of additional components in the client equipment.

3) Videoconference and chat. This application allows compartmentalization of the browser window, remote desktop, or blackboard application. It also permits sending files and recordings of videos or sessions.

Figure 2. Computerized tomography image.

Evaluation

A total of 28 patients, 19 from SJDH and 9 from JMDDH, were evaluated through interconsultation sheets. Table 2 summarizes the main characteristics of patients.

Out of 28 patients, 19 (68%) received fibrinolytic treatment. These 19 patients were those who met the inclusion criteria. We analyzed four process indicators for fibrinolysis delivery (Table 3). The most common reasons for not treating with fibrinolysis were: clinical criteria in six out of nine patients (66%) and beyond the time window in three out of nine patients (33%).

Table 4 summarizes the main characteristics of the medical professionals who completed the questionnaire on acceptance of Telestroke. In total, we sent out 61 questionnaires, of which we received 34 (56%) completed. Table 5 shows the descriptive statistics analysis of the items questionnaire. As shown, the mean values for each item were over 6.50, indicating that respondents positively evaluated each item. To assess the reliability of the questionnaire we calculated the Cronbach alpha value, obtaining a value of 0.97.

Table 2. Characteristics of patients; data are presented as n (%).

		No Fibrinolysis	Fibrinolysis
Age	40 – 49	2 (22.2)	3 (15.8)
	50 – 59	2 (22.2)	1 (5.0)
	60 – 69	1 (11.1)	4 (21.0)
	70 – 80	4 (44.4)	11 (58.2)
Sex	Females	3 (33.3)	10 (52.6)
	Males	6 (66.7)	9 (47.4)
Origin[a]	Own decision	3 (33.3)	6 (35.3)
	Health center	3 (33.3)	4 (23.5)
	Emergency department	0 (0)	1 (5.9)
	MICU[b]	3 (33.3)	7 (35.3)
Displacement of the patient	Own vehicle	4 (44.4)	7 (36.8)
	Ambulance	5 (55.6)	12 (63.2)

[a] There are values missing for these variables.

[b] MICU: Mobile Intensive-Care Unit.

Table 3. Indicators in process of fibrinolysis delivery.

	Min-Max	Mean ± SD
Time		
Onset-to-hospital	15-155	68.68 ± 41.76
Door-to-needle	33-128	81.89 ± 28.29
Admission-to-CT scan	6-85	32.87 ± 19.11
Onset-to-needle	70-210	150.58 ± 38.61

Table 4. Characteristics of medical professionals who have used TeleStroke; data are presented as n (%).

Categories		SJDH	JMDDH	VRUH	Total
Sex	Females	6 (17.6)	1 (2.9)	6 (17.6)	13 (38.2)
	Males	13 (38.2)	3 (8.8)	5 (14.7)	21 (61.8)
Age	<40	9 (26.4)	1 (2.9)	8 (23.5)	18 (52.9)
	40-49	9 (26.4)	2 (5.9)	2 (5.9)	13 (38.2)
	≥50	1 (2.9)	1 (2.9)	1 (2.9)	3 (8.8)
Clinical practice years	<10	2 (5.9)	-	5 (14.7)	7 (20.6)
	10-19	17 (50.0)	3 (8.8)	4 (11.7)	24 (70.6)
	≥20	-	1 (2.9)	2 (5.9)	3 (8.8)
Medical speciality	Family doctors	7 (20.6)	-	-	7 (20.6)
	Internists	9 (26.5)	-	-	9 (26.5)
	Intensivists	3 (8.8)	4 (11.7)	-	7 (20.6)
	Neurologists	-	-	11 (32.3)	11 (32.3)
Experience with TeleStroke	Knowledge/training courses	6 (17.6)	-	2 (5.9)	8 (23.5)
	Clinical case	13 (38.2)	4 (11.7)	9 (26.5)	26 (76.5)

Table 5. Descriptive statistics of the TAM questionnaire item.

Items		Min-Max	Mean ± SD
Perceived usefulness			
	TS could enhance my effectiveness of job	1-10	7.35 ± 2.06
	TS would allow greater control over DTP[a]	1-10	7.68 ± 1.98
	TS could support critical aspects in DTP	3-10	8.18 ± 1.80
	If I use TS[b], I will increase my chances to develop my career	1-10	7.06 ± 2.36
	Using TS would make my job easier	1-10	6.62 ± 2.62
	Using TS would improve my job performance	1-10	7.32 ± 2.03
	Using TS would help me to accomplish DTP more quickly	1-10	6.97 ± 2.89
	TS could improve the quality of DTP that I deliver	3-10	7.82 ± 1.87
	Overall, TS could be useful to improve DTP	1-10	8.06 ± 1.98
Perceived ease of use			
	Learning to use TS would be easy for me	4-10	7.56 ± 1.71
	My interaction with TS would be clear and understandable	3-10	7.09 ± 1.91
	I think that DTP made through TS would be clear	3-10	7.82 ± 1.80
	It would be easy for me to become skillful at using TS	4-10	7.21 ± 1.74
	Overall, I believe that TS will be easy to use	2-10	6.65 ± 2.14
Subjective norm			
	Colleagues whose opinions I value think I should use TS	3-10	7.91 ± 1.91
	My superiors think that I should use TS	8-10	9.47 ± 0.75
	Other health professionals whose opinions I value think I should use TS	3-10	7.91 ± 1.90
	The management of the hospital supports me to use TS	2-10	9.03 ± 1.66
	Overall, I believe that the hospital supports the use of TS	5-10	8.94 ± 1.32
Facilitating conditions			
	A specific person will be available to solve problems regarding to TS	1-10	6.65 ± 2.00
	I will have the resources necessary to use TS	3-10	7.94 ± 1.84
	I will receive training to use TS	2-10	7.12 ± 2.21
	TS will be compatible with other systems I use	4-10	8.18 ± 1.58
	The hospital has the infrastructure necessary to I use TS	4-10	7.91 ± 1.83
Intention to use			
	I intend to use TS as it is available in the hospital	4-10	9.03 ± 1.55
	I intend to use TS for DTP as often as needed	4-10	8.97 ± 1.42
	Whenever possible, I intend to use TS	3-10	8.82 ± 1.73

[a] TS: TeleStroke.

[b] DTP: diagnosis and treatment of patients.

Discussion

The implementation of TeleStroke has made it possible for patients in the acute phase of stroke to receive effective treatment, something that was previously impossible because of the time required to transfer them to the RH. In 2006, before implementing the TeleStroke service, 57% of acute stroke patients transferred to the Virgen del Rocío University Hospital emergency department from another hospital presented more than 4.5 hours after the onset of symptoms [36]. However, it is important to note that patients who arrived at the Virgen del Rocío University Hospital emergency department between 3 and 4.5 hours after the onset of symptoms may not be treated with fibrinolysis because of the time required for performing diagnostic tests. Given that this situation does not satisfy the recommended maximum time for treatment, it justifies the deployment of the telemedicine service.

System reliability is vital for the successful implementation of a telemedicine system with a high impact on care. In this sense, formal and rigorous methods of risk management are required, as is good capacity management of the technology to guarantee the service level. We demonstrated that reliability was also essential, and we subjected the system to an exhaustive testing process to ensure this.

Several technological upgrades were carried out during the period of the study in order to increase the system usability and reliability, improving the medical staff access and therefore decreasing the "door-to-needle" time. In turn, the application of BPM modeling techniques to model the care protocol proved to be the right choice because it led to explicit agreement between health professionals on the best way to work. This will make it easier to export it to the rest of the Andalusian Public Health System.

The need to accelerate the communication time of the CT images to the RH has been made clear. Despite conversion to JPG format greatly reducing file size, this remains the most important time period in the whole process. To this end, we have checked the diagnostic validity of JPEG format CT images with regards to DICOM format CT images in a complementary study that is awaiting publication. We found a far higher percentage of patients who had received fibrinolysis compared to other studies that examined stroke patients through videoconference [9,12-14,16], while the number of patients surveyed was higher in most of these studies. The application of BPM modeling techniques may contribute to a higher fibrinolytic treatment rate facilitating agreement and training among professionals involved in the care process.

The CT image is the critical test for the successful application of fibrinolysis to patients. The mean time from admission to CT image was 33 minutes, higher than other studies [16]. In our study, one patient had a symptomatic intracranial haemorrhage detected in the CT image. The mean "onset-to-hospital" time was 69 minutes. Similar studies [9,12-14,16] reported this time ranged from 36 to 84 minutes. The mean "door-to-needle" time was 82 minutes. Similar studies [9,11-12,16] reported this time ranged from 53 to 106 minutes. The mean "onset-to-needle" time was 150 minutes. Similar studies [9-12,14,16] reported this time ranged from 122 to 157 minutes.

There were no technical failures preventing a clinical response being given to the patient within the established limits, although there were delays in receiving the CT image and problems with sound and images from the videoconferences that increased response times.

The acceptance of the TeleStroke system by health care professionals is essential for the incorporation of eServices to be successful. To assess the reliability of the questionnaire, we obtained a value of 0.97 for Cronbach alpha, higher than the recommended value of 0.7. Results of the TAM questionnaire indicate a very positive perception of questionnaire items, particularly the Intention to Use and Subjective Norms dimensions.

The authors have identified two important limitations in this study: the small sample size and the lack of patient outcomes related to mortality and functional status. These issues will be covered in a future study.

As current evidence is limited, a cost-effectiveness analysis may be required into the economic impact of telemedicine in acute stroke treatment [37] following the methodology applied in a recent study [38]. Currently, we are working on adapting the TeleStroke system to be deployed in the rest of the Andalusian Public Health System. A new line of work is the development of mobility technology to capture and transmit videoconference and CT data to allow swifter diagnosis and treatment of fibrinolysis in acute stroke patients [39].

Taking all of these discussion factors into account, we can state that the implementation of a telemedicine service for the treatment of acute stroke patients at Virgen del Rocío University Hospital is feasible, and it has been proven that TeleStroke facilitates the accessibility to neurological specialist care and fibrinolytic treatment in hospitals where a fully equipped stroke center is not available.

Acknowledgments

This work has been conducted within the framework of the R&D project "Desarrollo e Implantación de un nuevo dispositivo asistencial para la atención a pacientes con ictus basado en la Telemedicina (Teleictus) / Develop and implement a new treatment device for stroke patient care based on telemedicine: TeleStroke", financed by the Spanish Ministry of Science and Innovation.

We thank all who participated in this project and particularly those who directly or indirectly contributed to the publication of this article from San Juan de Dios Hospital and José María Díaz Domínguez Hospital.

We thank the Network for Innovation in Medical Technologies and Health (Red ITEMAS in Spanish), dossier number RD09/0077/00025, promoted by the Carlos III Health Institute.

Conflicts of Interest

Francisco J. Fernández works at Everis Group. The remaining authors have no conflict of interest.

Multimedia Appendix 1

Process model TeleStroke.

[JPG File, 210KB - ijmr_v1i2e15_app1.JPG]

References

1. World Health Organization. World health statistics. Geneva, Switzerland: WHO; 2008.
2. Regidor E, Gutiérrez-Fisac JL, Alfaro M. Health Indicators. Spain: Ministry of Health and Social Policy; 2009.
3. Stroke Unit Trialists' Collaboration. Organised inpatient (stroke unit) care for stroke. Cochrane Database Syst Rev 2007(4):CD000197. [doi: 10.1002/14651858.CD000197.pub2] [Medline: 17943737]
4. Ribo M, Molina CA, Pedragosa A, Sanclemente C, Santamarina E, Rubiera M, et al. Geographic differences in acute stroke care in Catalunya: impact of a regional interhospital network. Cerebrovasc Dis 2008;26(3):284-288. [doi: 10.1159/000147457] [Medline: 18648202]
5. Klein GO. Standardization of health informatics--results and challenges. Methods Inf Med 2002;41(4):261-270. [Medline: 12425236]
6. Handschu R, Littmann R, Reulbach U, Gaul C, Heckmann JG, Neundörfer B, et al. Telemedicine in emergency evaluation of acute stroke: interrater agreement in remote video examination with a novel multimedia system. Stroke 2003 Dec;34(12):2842-2846 [FREE Full text] [doi: 10.1161/01.STR.0000102043.70312.E9] [Medline: 14615620]
7. Shafqat S, Kvedar JC, Guanci MM, Chang Y, Schwamm LH. Role for telemedicine in acute stroke. Feasibility and reliability of remote administration of the NIH stroke scale. Stroke 1999 Oct;30(10):2141-2145 [FREE Full text] [Medline: 10512919]
8. Wang S, Lee SB, Pardue C, Ramsingh D, Waller J, Gross H, et al. Remote evaluation of acute ischemic stroke: reliability of National Institutes of Health Stroke Scale via telestroke. Stroke 2003 Oct;34(10):e188-e191 [FREE Full text] [doi: 10.1161/01.STR.0000091847.82140.9D] [Medline: 14500929]
9. Audebert HJ, Kukla C, Clarmann von Claranau S, Kühn J, Vatankhah B, Schenkel J, TEMPiS Group. Telemedicine for safe and extended use of thrombolysis in stroke: the Telemedic Pilot Project for Integrative Stroke Care (TEMPiS) in Bavaria. Stroke 2005 Feb;36(2):287-291 [FREE Full text] [doi: 10.1161/01.STR.0000153015.57892.66] [Medline: 15625294]
10. Meyer BC, Raman R, Hemmen T, Obler R, Zivin JA, Rao R, et al. Efficacy of site-independent telemedicine in the STRokE DOC trial: a randomised, blinded, prospective study. Lancet Neurol 2008 Sep;7(9):787-795 [FREE Full text] [doi: 10.1016/S1474-4422(08)70171-6] [Medline: 18676180]
11. Pedragosa A, Alvarez-Sabín J, Molina CA, Brugués J, Ribó M. [Endovenous thrombolysis in a district hospital using the telestroke system]. Rev Neurol 2011 Aug 1;53(3):139-145 [FREE Full text] [Medline: 21748711]
12. Schwamm LH, Rosenthal ES, Hirshberg A, Schaefer PW, Little EA, Kvedar JC, et al. Virtual TeleStroke support for the emergency department evaluation of acute stroke. Acad Emerg Med 2004 Nov;11(11):1193-1197. [doi: 10.1197/j.aem.2004.08.014] [Medline: 15528584]
13. Waite K, Silver F, Jaigobin C, Black S, Lee L, Murray B, et al. Telestroke: a multi-site, emergency-based telemedicine service in Ontario. J Telemed Telecare 2006;12(3):141-145. [doi: 10.1258/135763306776738611] [Medline: 16638235]
14. Wang S, Gross H, Lee SB, Pardue C, Waller J, Nichols FT, et al. Remote evaluation of acute ischemic stroke in rural community hospitals in Georgia. Stroke 2004 Jul;35(7):1763-1768 [FREE Full text] [doi: 10.1161/01.STR.0000131858.63829.6e] [Medline: 15166386]
15. Wiborg A, Widder B, Telemedicine in Stroke in Swabia Project. Teleneurology to improve stroke care in rural areas: The Telemedicine in Stroke in Swabia (TESS) Project. Stroke 2003 Dec;34(12):2951-2956 [FREE Full text] [doi: 10.1161/01.STR.0000099125.30731.97] [Medline: 14631092]
16. Carmen Jiménez M, Tur S, Legarda I, Vives B, Gorospe A, José Torres M, et al. [The application of telemedicine for stroke in the Balearic Islands: the Balearic Telestroke project]. Rev Neurol 2012 Jan 1;54(1):31-40 [FREE Full text] [Medline: 22187210]
17. Meyer BC, Lyden PD, Al-Khoury L, Cheng Y, Raman R, Fellman R, et al. Prospective reliability of the STRokE DOC wireless/site independent telemedicine system. Neurology 2005 Mar 22;64(6):1058-1060. [doi: 10.1212/01.WNL.0000154601.26653.E7] [Medline: 15781827]
18. Audebert HJ, Wimmer ML, Hahn R, Schenkel J, Bogdahn U, Horn M, TEMPIS Group. Can telemedicine contribute to fulfill WHO Helsingborg Declaration of specialized stroke care? Cerebrovasc Dis 2005;20(5):362-369. [doi: 10.1159/000088064] [Medline: 16141717]
19. Hess DC, Wang S, Hamilton W, Lee S, Pardue C, Waller JL, et al. REACH: clinical feasibility of a rural telestroke network. Stroke 2005 Sep;36(9):2018-2020 [FREE Full text] [doi: 10.1161/01.STR.0000177534.02969.e4] [Medline: 16051892]
20. Fisher M. Developing and implementing future stroke therapies: the potential of telemedicine. Ann Neurol 2005 Nov;58(5):666-671. [doi: 10.1002/ana.20659] [Medline: 16240338]
21. Giordano R, Clark M, Goodwin N. Perspectives on telehealth and telecare. London: WSD Action Network; 2011.
22. Panzarasa S, Maddè S, Quaglini S, Pistarini C, Stefanelli M. Evidence-based careflow management systems: the case of post-stroke rehabilitation. J Biomed Inform 2002 Apr;35(2):123-139. [Medline: 12474426]
23. Mans R, Schonenberg H, Leonardi G, Panzarasa S, Cavallini A, Quaglini S, et al. Process mining techniques: an application to stroke care. Stud Health Technol Inform 2008;136:573-578. [Medline: 18487792]

24. Parra Calderón CL, Framiñán Torres JM, Pérez González P, Montes Worboys M, de la Higuera González JM, Melero Bellido JM. Metodologia de Modelado de Procesos en Servicios de Salud utilizando Telemedicina. I+S: Informática y Salud 2005 Mar;51:16-25.
25. Rojo MG, Rolón E, Calahorra L, García FO, Sánchez RP, Ruiz F, et al. Implementation of the Business Process Modelling Notation (BPMN) in the modelling of anatomic pathology processes. Diagn Pathol 2008;3 Suppl 1:S22 [FREE Full text] [doi: 10.1186/1746-1596-3-S1-S22] [Medline: 18673511]
26. Ruffolo M, Curia R, Gallucci L. Process Management in Health Care: A System for Preventing Risks and Medical Errors. In: van der Aalst WMP, Benetallah B, Casati F, Curbera F, editors. Business Process Management. Berlin/Heidelberg: Springer-Verlag; 2005:334-343.
27. Svagård I, Farshchian BA. Using business process modelling to model integrated care processes. Experiences from a European Project. In: Lecture Notes in Computer Science. Berlin: Springer; 2009:922-925.
28. White SA, Miers D. BPMN Modeling and Reference Guide: Understanding and Using BPMN. 2008. URL: http://www.google.com/search?tbs=bks:1&q=isbn:0977752720 [accessed 2012-11-05]
29. World Health Organization. eHealth for Health-care Delivery: Strategy 2004-2007. Geneva: WHO; 2007.
30. Health Level Seven International. URL: http://www.hl7.org/ [accessed 2012-11-05] [WebCite Cache ID 67vpvsobp]
31. Digital Imaging and Communications in Medicine. URL: http://medical.nema.org/ [accessed 2012-05-08] [WebCite Cache ID 67VQqG8ww]
32. Davis FD. Perceived usefulness, perceived ease of use and user acceptance of information technology. MIS Q 1989;13(3):319-339.
33. Hu PJ, Chau PYK, Tam KY. Examining the technology acceptance model using physician acceptance of telemedicine technology. J Manag Inform Syst 1999;16:91-112.
34. Holden RJ, Karsh BT. The technology acceptance model: its past and its future in health care. J Biomed Inform 2010 Feb;43(1):159-172 [FREE Full text] [doi: 10.1016/j.jbi.2009.07.002] [Medline: 19615467]
35. Joint Photographic Experts Group. URL: http://www.jpeg.org/ [accessed 2012-05-10] [WebCite Cache ID 67YT6p2Il]
36. Palomino-García A, Moniche-Álvarez F, de la Torre-Laviana FJ, Cayuela-Domínguez A, Vigil E, Jiménez-Hernández MD. [Factors that affect time delays to fibrinolytic treatment in ischaemic stroke]. Rev Neurol 2010 Dec 16;51(12):714-720 [FREE Full text] [Medline: 21157733]
37. Johansson T, Wild C. Telemedicine in acute stroke management: systematic review. Int J Technol Assess Health Care 2010 Apr;26(2):149-155. [doi: 10.1017/S0266462310000139] [Medline: 20392317]
38. Nelson RE, Saltzman GM, Skalabrin EJ, Demaerschalk BM, Majersik JJ. The cost-effectiveness of telestroke in the treatment of acute ischemic stroke. Neurology 2011 Oct 25;77(17):1590-1598. [doi: 10.1212/WNL.0b013e318234332d] [Medline: 21917781]
39. Walter S, Kostpopoulos P, Haass A, Helwig S, Keller I, Licina T, et al. Bringing the hospital to the patient: first treatment of stroke patients at the emergency site. PLoS One 2010;5(10):e13758 [FREE Full text] [doi: 10.1371/journal.pone.0013758] [Medline: 21060800]

Abbreviations

BPM: Business Process Modeling
BPMN: Business Process Modeling Notation
CT: Computerized Tomography
DICOM: Digital Imaging and Communication in Medicine
eHCD: eHealth for Health Care Delivery
HL7: Health Level Seven
JMDDH: José María Díaz Domínguez Hospital
NIHSS: National Institute of Health Stroke Scale
RH: Referral Hospital
SJDH: San Juan de Dios Hospital
SD: Standard Deviation
TAM: Technology Acceptance Model

Edited by G Eysenbach; submitted 11.05.12; peer-reviewed by J Brice, S Davis, D Jacobs; comments to author 30.07.12; revised version received 03.08.12; accepted 21.09.12; published 15.11.12.

Please cite as:
Parra C, Jódar-Sánchez F, Jiménez-Hernández MD, Vigil E, Palomino-García A, Moniche-Álvarez F, De la Torre-Laviana FJ, Bonachela P, Fernández FJ, Cayuela-Domínguez A, Leal S
Development, Implementation, and Evaluation of a Telemedicine Service for the Treatment of Acute Stroke Patients: TeleStroke
Interact J Med Res 2012;1(2):e15
URL: http://www.i-jmr.org/2012/2/e15/
doi:10.2196/ijmr.2163
PMID:23612154

Original Paper

Improving Interoperability in ePrescribing

Sten-Erik Öhlund[1,2,3], BSC, PhLic; Bengt Åstrand[1], MPharm, PhD; Göran Petersson[1], MD, PhD

[1]eHealth Institute, Linnaeus University, Kalmar, Sweden

[2]Information Systems, Department of Management and Engineering, Linköping University, Linköping, Sweden

[3]Logica Sverige AB, Kalmar, Sweden

Corresponding Author:
Sten-Erik Öhlund, BSC, PhLic
Logica Sverige AB
Skeppsbrogatan 49
Kalmar, 391 23
Sweden
Phone: 46 733982117
Fax: 46 480 85573
Email: sten.erik.ohlund@logica.com

Related Article:

This is a corrected version. See correction statement: http://www.i-jmr.org/2012/2/e19/

Abstract

Background: The increased application of eServices in health care, in general, and ePrescribing (electronic prescribing) in particular, have brought quality and interoperability to the forefront. The application of standards has been put forward as one important factor in improving interoperability. However, less focus has been placed on other factors, such as stakeholders' involvement and the measurement of interoperability. An information system (IS) can be regarded to comprise an instrument for technology-mediated work communication. In this study, interoperability refers to the interoperation in the ePrescribing process, involving people, systems, procedures and organizations. We have focused on the quality of the ePrescription message as one component of the interoperation in the ePrescribing process.

Objective: The objective was to analyze how combined efforts in improving interoperability with the introduction of the new national ePrescription format (NEF) have impacted interoperability in the ePrescribing process in Sweden, with the focus on the quality of the ePrescription message.

Methods: Consecutive sampling of electronic prescriptions in Sweden before and after the introduction of NEF was undertaken in April 2008 (pre-NEF) and April 2009 (post-NEF). Interoperability problems were identified and classified based on message format specifications and prescription rules.

Results: The introduction of NEF improved the interoperability of ePrescriptions substantially. In the pre-NEF sample, a total of 98.6% of the prescriptions had errors. In the post-NEF sample, only 0.9% of the prescriptions had errors. The mean number of errors was fewer for the erroneous prescriptions: 4.8 in pre-NEF compared to 1.0 in post-NEF.

Conclusions: We conclude that a systematic comprehensive work on interoperability, covering technical, semantical, professional, judicial and process aspects, involving the stakeholders, resulted in an improved interoperability of ePrescriptions.

(Interact J Med Res 2012;1(2):e17) doi:10.2196/ijmr.2089

KEYWORDS

eHealth, Electronic prescribing, Electronic prescription, Information quality, Interoperability

Introduction

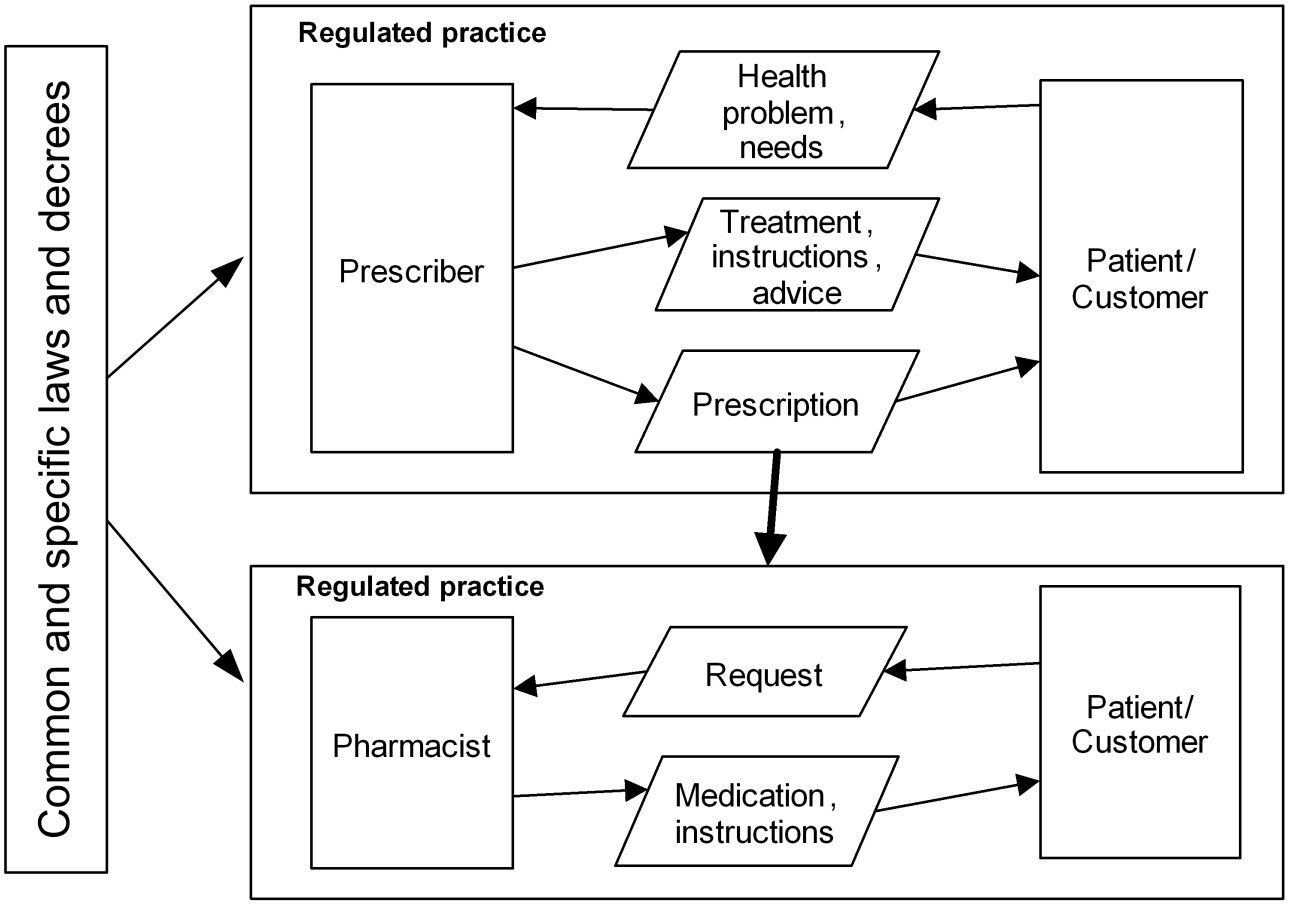
The increased use of eServices in health care in general, and ePrescribing (electronic prescribing) in particular, has placed a focus on quality [1] and interoperability [2]. In Sweden, ePrescribing has increased notably during the 2000s. In 2011, ePrescriptions constituted more than 90% of all filled prescriptions [3].

ePrescribing

ePrescribing is a co-operation between prescriber, patient, and pharmacist for the purpose of medical treatment of the patient. A prescription is a regulated, social act, authorizing a pharmacist to dispense a medical drug to a patient [4]. ePrescribing has been analyzed using a Generic Regulation Model (GRM) (Figure 1) [4,5].

The interaction between the patient and the prescriber results in a prescription having multiple functions, such as an authorization to a pharmacy to dispense a medical drug according to certain rules, a directive to the patient for medical treatment, and a commitment on behalf of the health care organization (a reimbursement commitment specific to Sweden) to the patient that the patient will receive reimbursement for the prescribed medical drug. The prescription also provides a basis for further information processing by authorities and researchers.

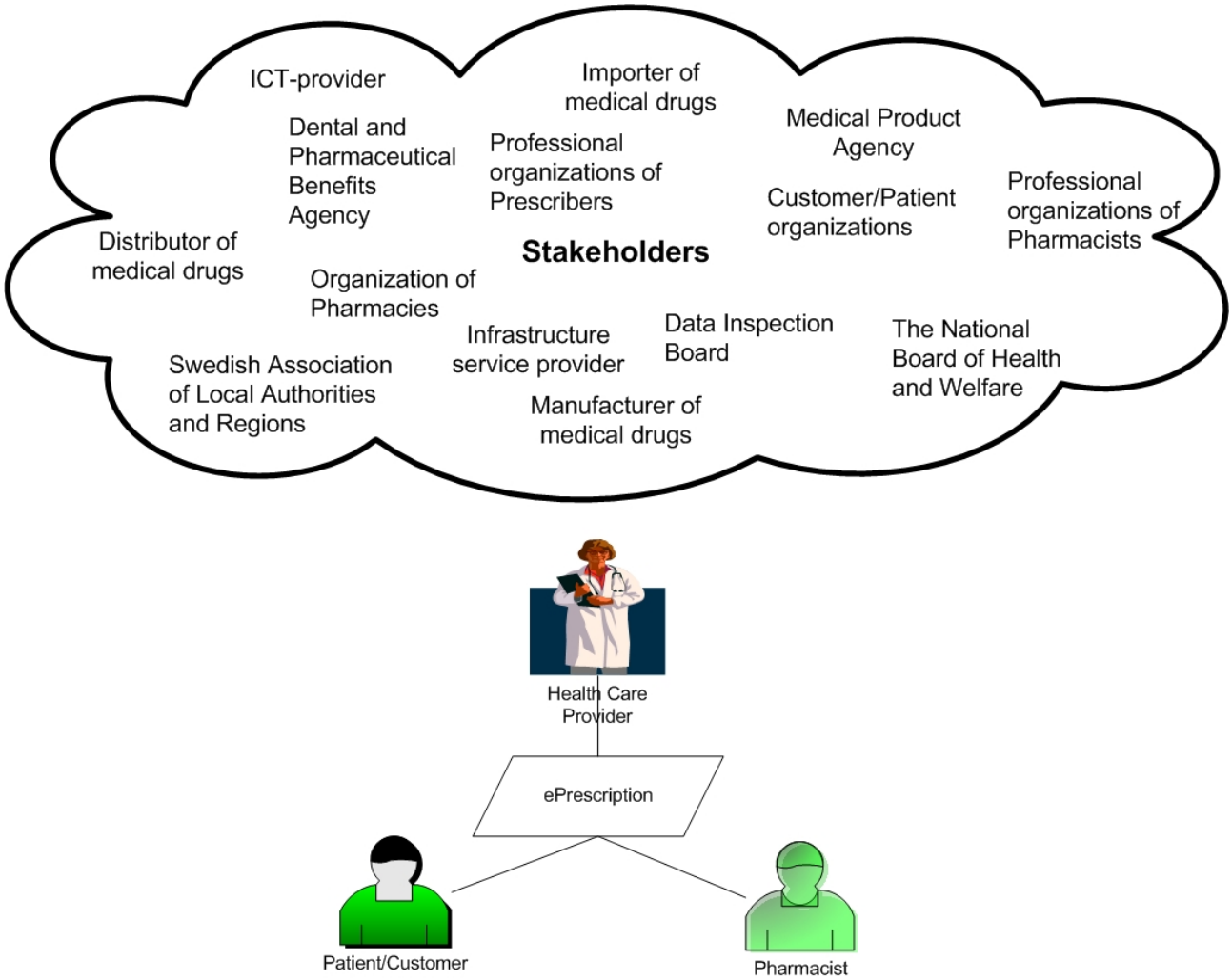
Figure 1. The Generic Regulation Model applied to ePrescribing [4].



Stakeholders in ePrescribing

The ePrescribing process involves a large number of stakeholders having interest in, being influenced by, and who influence the ePrescribing process and the content of the ePrescription (Figure 2). The multiple functions inherent in the prescription and the different stakeholders involved make the communication regarding ePrescriptions complex. As an illustration, the Medical Products Agency approves the medical drugs to be sold on the Swedish market, determines which drugs are to be considered generic exchangeable drugs, and acts as the statutory authority for prescription regulations; whereas, the Dental and Pharmaceutical Benefits Agency decides which medical drugs are to be reimbursed and determines procurement and sales prices. Furthermore, the National Board of Health and Welfare regulates the health care sector and provides information about prescribers and their prescription rights, as well as granting practitioner licenses to prescribers and pharmacists. Basic information about the medical drugs to be prescribed is provided by the Medical Product Agency and by the manufacturers and importers of medical drugs. This information is communicated both to the prescriber organizations and their Electronic Health Record (EHR) systems, and to the pharmacies and their dispensing systems via a state-owned infrastructure service provider. This service provider also provides e-services for ePrescriptions, dispensing, reimbursement of prescriptions, and statistics.

Figure 2. Stakeholders in ePrescribing.



ePrescribing as Communication

Information Systems Actability Theory [6] considers an Information System (IS) to be an instrument for technology-mediated work communication. In addition to the technical aspects, user interaction, communication between users applying IS as an instrument of communication, and the overall influence of information and actions involved in the process, impact the ePrescribing process. The ability to interoperate in ePrescribing is closely related to interaction, communication, and process quality (Figure 3).

A sender interacts with an IS to communicate something to a receiver, which in turn interacts with an IS to read and interpret a message. Interaction quality criteria can be defined for this activity, ie, what the users are doing *with* the system. In this process, there is also communication between the sender and the receiver about what the users are doing *through* the system. Communication quality criteria referr to the formulation and communication of messages by a sender, as well as the reading and interpreting of messages by a receiver. Process quality is concerned with what the users do *outside* the system, ie, the effects of IS usage on work practice.

Interaction quality criteria for a prescriber might imply that the vocabulary of the system is intelligible and in line with terminology of the profession or regulated practice—that it is obvious what the user can do in the system and that consequences of different actions are transparent. For a prescriber, this could imply that the consequences approving or cancelling a prescription are clear and that navigation between the various parts of the system is easy.

For a pharmacist, communication quality criteria might refer to relevant prescription information being easily available for dispensing, that the information is accurate, that it is obvious who the sender is, and that the intention of the prescriber is unambiguous.

In general, process quality criteria refer to the requirement that the information from the system is useful on behalf of its users; ie, that the information has a meaningful use. In the process of ePrescribing, eg, the system should support process objectives, such as patient safety, correct reimbursement processes, clear instructions for the patients, but should also support the objectives of other stakeholders, such as achieving correct statistics for researchers and authorities.

The quality of the communication between prescriber and pharmacist is dependent on many factors, among them the quality of the communicated message. The quality of the communicated message is dependent on the quality of formulating an ePrescription (part of communication quality)

and by interaction quality of the EHR system. Finally, how the communicated message is presented and made available to the pharmacists in their dispensing systems affects the overall communication quality. Here, we elucidate the communication of ePrescription messages between the EHR systems prescription modules, the ePrescription service system, and the quality of this message (box marked IS) with regard to the requirements that have been established for this communication (Figure 4). In this study interaction, quality and communication quality aspects in EHR and dispensing systems are not considered, although they affect the quality of the communicated message.

Figure 3. Different layers of quality according to the Information Systems Actability Theory [6].

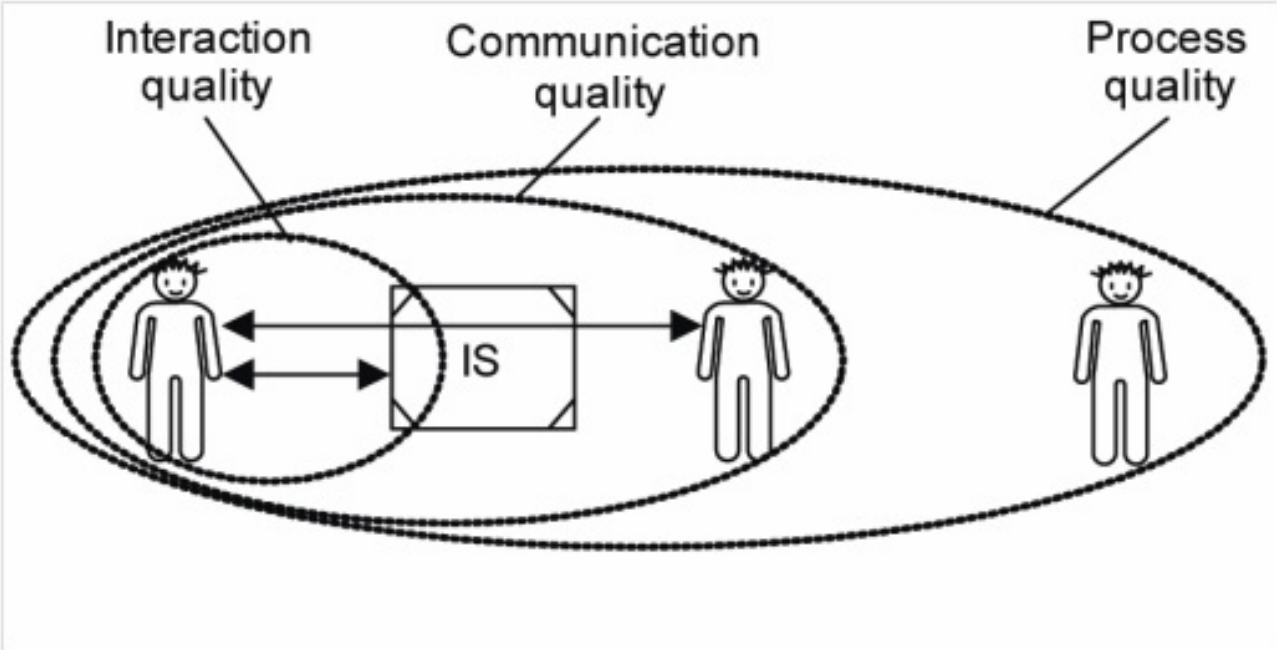
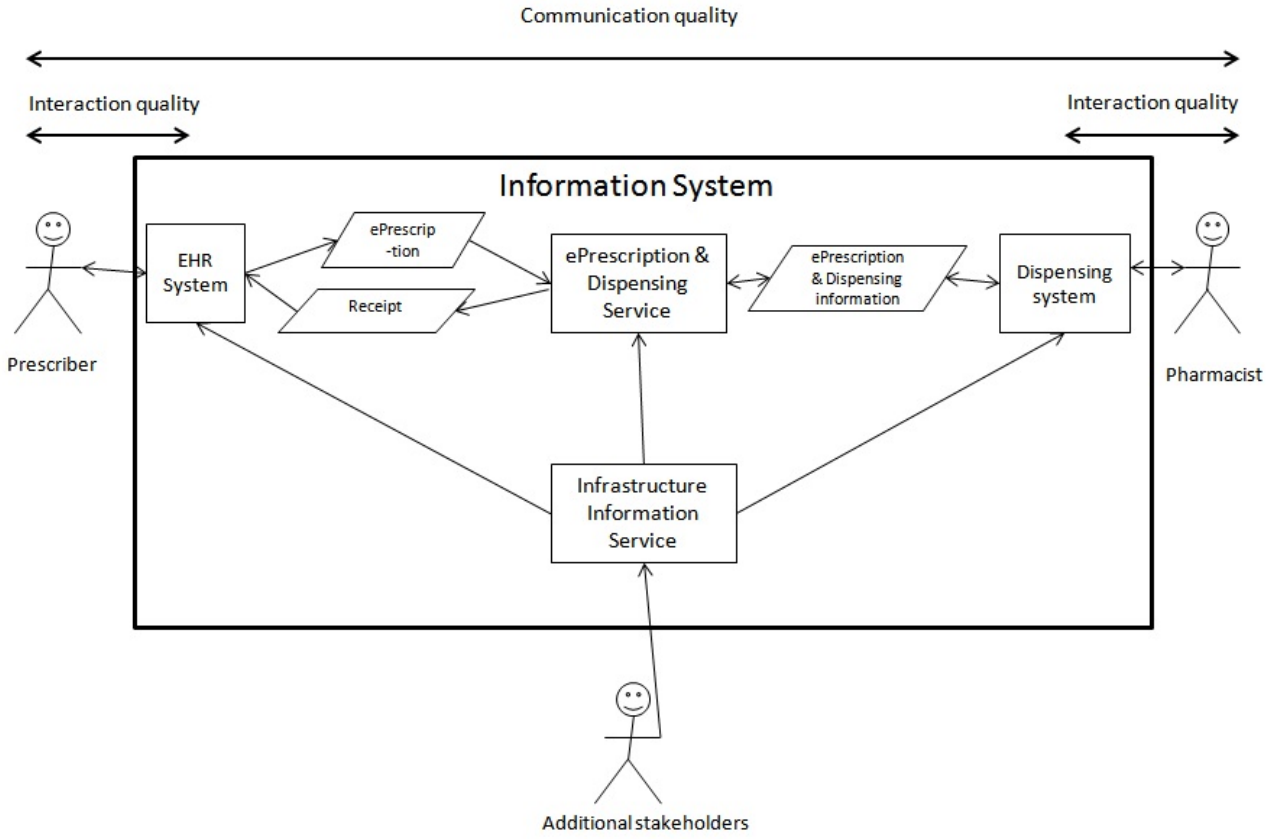


Figure 4. Overview of IS and stakeholders involved in ePrescribing communication.



Communication Quality in ePrescribing

The quality of the ePrescription message is one part of communication quality in the ePrescribing process between the prescriber and the pharmacist. The stakeholder's involvement in providing infrastructure information influences the communication quality, such as the information about the population, medical drugs, prices, reimbursement rules, prescribers, and pharmacist rights. This information is provided through various information services, which can be grouped into infrastructure information services, and ePrescription and dispensing services. These services are provided to the prescriber and pharmacy systems and have an important influence on the communication quality. The interaction quality of the EHR and dispensing systems are important to consider in assessing communication quality [7].

Interoperability in ePrescribing

The various definitions of interoperability involve different perspectives on interoperation. Some definitions focus on the ability of systems to interoperate [8], and others focus on the ability of people to interoperate (individual, organizational level) by using systems to achieve a certain goal [9,10].

We regard interoperability as the capability of the entire process, involving people, systems, procedures, and organizations, to interoperate using IS in order to achieve its objectives. Here, we focus on one vital aspect of the interoperation in the ePrescribing process: the communication of the ePrescription message. To understand the complexity of the ePrescription message, it is necessary to analyze it as one component of communication in the ePrescribing process involving many stakeholders and IS. Thus, the content of the ePrescription message is dependent on stakeholder involvement both in the infrastructure services and in the actual formulation and interpretation of the prescription message by prescriber, pharmacist, and systems involved.

European Interoperability Framework (EIF) [10] defines four levels of interoperability: legal, organizational, semantic, and technical interoperability (Table 1).

Table 1. Levels of interoperability.

Level of interoperability	Description
Legal	Alignment of legislation concerning the interoperation between different organizations, which affects how and what can be communicated
Organizational	How different organizational processes are integrated and how information exchange is managed
Semantic	Processing of information in a meaningful way, provided that information in the communicated message is precisely defined, agreed, and understood by all the stakeholders involved
Technical	Technical prerequisites linking different systems, such as communication protocols, message format, services, interface specification, etc.

All these levels of interoperability influence the actual implementation of the ePrescription message and service, which is the central point where the interoperation is achieved. To achieve interoperability, significant emphasis in eHealth has been put on the application of standards and common terminology, such as ISO 13606 [11], HL7 [12], and Snomed CT [13]. Although these standards are important building blocks in achieving interoperability, less attention has been put on other factors for improving interoperability, such as measuring interoperability and stakeholder involvement in order to address other aspects of interoperability, which is particularly relevant for interorganizational interoperability.

The Implementation of a New National ePrescription Format in Sweden

The introduction in 2009 of a new National ePrescription Format (NEF) in Sweden was the result of an effort to improve interoperability in ePrescribing. Consequently, this provided a unique opportunity to study NEF's effect in terms of changes on the interoperability in ePrescribing.

The implementation of ePrescription messages sent from health care organizations to pharmacies in Sweden has evolved during three decades. Varying communication standards with different message specifications (based on the UN standard MEDPRE [15] and the pre-standard ENV 13607 [7]) have been introduced and applied. The infrastructure has constantly evolved, from point-to point message communication of prescriptions by local health care organizations to local pharmacies, towards centralized communication on both sides. ePrescription have evolved from a mere electronic transfer of a message to management of the prescription during its entire life cycle with repeated refills. A growing number of health care regions have been involved and the number of prescription systems has increased. In 2009, there were 16 EHR-systems with ePrescription modules and one web-based prescription system sending ePrescriptions from 21 health care regions in Sweden.

Interoperability problems in ePrescribing were observed in an increasing number of issues by the support team at the Apoteket AB (National Corporation of Swedish Pharmacies—the state-owned pharmacy chain, which, at that time, was a monopoly). With the increased volume of ePrescriptions forecasted in 2006, from 30% to 80% (of new prescriptions) during the subsequent years, the handling cost of poor quality was expected to increase considerably. There was no automated control of incoming prescriptions other than a failure to store the ePrescription in the database. The system did not report any information about the possible cause of errors. Testing and approval of EHR systems to send ePrescriptions were rudimentary, and the process and organizational aspects of this were unclear, as was the process for maintenance and development of the ePrescribing process. Focus was on managing the technical communication, while the process of how to handle errors in content was pushed to the pharmacist at the end of the process.

To meet these challenges, a joint project between Apoteket AB and the regional health care providers was initiated in 2006 with the purpose to improve patient safety and to decrease the cost of deficient quality in ePrescribing. The project subsequently implemented the new NEF-format [14] in Sweden, together with a stricter test procedure than previously. Also, the automatic control of format and prescription rules was introduced. From May 31, 2009, all ePrescriptions in Sweden were issued in the NEF-format.

Objective

The objective of the present study was to analyze the manner in which the combined efforts in improving interoperability with the introduction of NEF affected interoperability in the ePrescribing process in Sweden, with a focus on the quality of the ePrescription message.

Methods

The present study was an intervention study. The intervention consisted of the combined efforts in implementation of NEF. We measured interoperability problems prior to and after the intervention: pre-NEF and post-NEF.

In the pre-NEF study period, interoperability errors in pre-NEF prescriptions were validated against the format specification and prescriptions rules valid for the pre-NEF study period. In the post-NEF study period, interoperability errors in post-NEF prescriptions were measured against the format specification and prescription rules valid for the post-NEF study period. The prescription rules did not change between the study periods.

In the two study periods, we compared changes in adherence to the agreed format specification and prescription rules based on legislation and agreed praxis. Consequently, the focus was on communication quality between health care and pharmacy using ePrescription as an instrument of communication for the medical treatment of a patient. Also, the assessment of communication quality was limited to the formal and documented requirements on the ePrescription message.

The hypothesis was that adherence to the agreed format specification and prescription rules should be improved with the introduction of NEF, resulting in fewer interoperability errors in the post-NEF period.

Information System Actability theory and theories about interoperability were used to analyze the ePrescribing process, the implementation of NEF, and its results.

Design

Consecutive sampling was applied on all incoming ePrescriptions during two periods: April 3, 2008, to May 3, 2008 (pre-NEF), and April 3, 2009, to May 3, 2009 (post-NEF). To be able to demonstrate a significant change of 1% between the two study periods, the required sample size was estimated to be approximately 1,450,000. The calculation was made by using a sample size calculator developed by Rollin F. Brant [15].

Consecutive sampling of all ePrescriptions during the two study periods was considered to be the best choice for handling historic changes in the drug database, including all ePrescriptions during the study periods. Electronic Data Interchange For Administration, Commerce and Transport (EDIFACT) prescriptions were present only in the pre-NEF period and were, therefore, excluded from the study.

During the two study periods, all of the prescribing systems used in Sweden were expected to be represented with a fair amount of ePrescriptions. During one month of sampling, ePrescriptions from the majority of the prescribers would be represented, except those making only a few prescriptions per year.

Definition of the ePrescribing Process

In this study, ePrescribing was regarded as a process that starts with a communication between prescriber and patient and that is related to medical treatment. The process also includes communication between patient and pharmacy, followed by the completion phase with the aim to evaluate the result of the medical treatment. The ePrescribing process is described and further developed in [4] as two interconnecting loops of generic phases following the Generic Exchange Model (GEM) [5]: initiation, agreement, fulfillment, and completion (Figure 5). Thus, ePrescribing is a complete process for medical treatment using ePrescriptions as an instrument for communication to achieve the process objectives.

In this study, we analyzed the interoperability problems in the ePrescription communication in ePrescribing. We did not study other interoperability problems in ePrescribing connected with or related to actions and consequences in the stakeholder processes influencing the medical treatment of the patient.

A prescription set contains a number of prescriptions for a patient by a prescriber at a certain point in time. A prescription set is equal to a prescription message. A prescription refers to the prescription of a medical drug.

Intervention: Combined Efforts for Implementation of NEF

The intervention refers to all of the combined efforts associated with the implementation of NEF (Table 2).

Figure 5. ePrescribing process seen as a process of initiation, agreement, fulfillment, and completion in two exchange situations: Prescriber – Patient and Pharmacy – Patient/Customer.

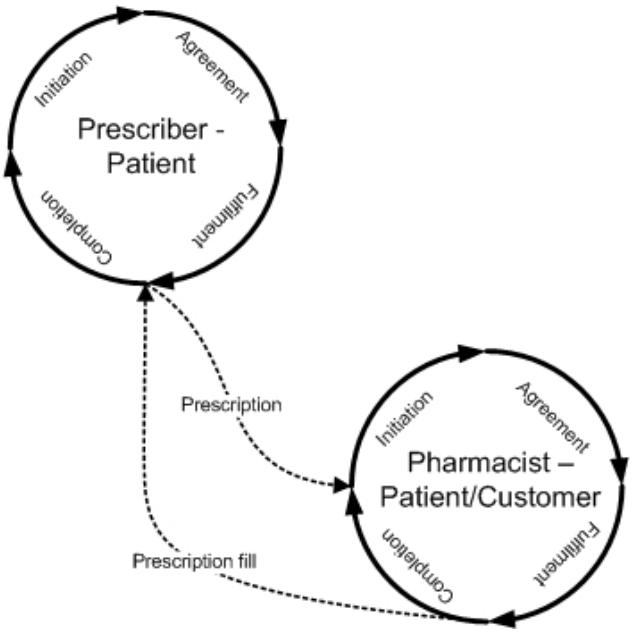


Table 2. Summary of actions taken with the introduction of NEF.

Action	Description
Phasing out of old formats	Phasing out of the United Nations Electronic Data Interchange For Administration, Commerce and Transport (UN/EDIFACT) format.
Definition of terms	Extensive definition of the usage of terms in the ePrescription message with minimal changes in the previous ePrescription message.
New features in ePrescription	New features of the ePrescription message: unique prescription identifications, version, and name of EHR system.
Format control	Applying a new strict Extensible Markup Language (XML) schema complying with the ePrescriptions format to validate incoming ePrescriptions.
Validation of prescription rules	"Online" validation of prescription rules and the completeness of ePrescriptions in the communication process.
Improved feedback to prescriber	New, improved feedback from the pharmacy systems to the prescriber, including validation results.
New test and approval procedures	Applying new and more rigorous test procedures before approving an EHR system for the sending of ePrescriptions to a pharmacy.

During the intervention, stakeholders from the health care providers, pharmacies, and software vendors were involved. The stakeholders participated in the work to achieve a more rigorous and developed message specification, an improved testing and approval process, the implementation of automatic control of incoming prescriptions messages, and to secure a more developed feedback mechanism, using a new extended acknowledge receipt message with status and error codes.

There were common national and regional meetings, including a referral and revision procedure of the NEF documentation. The new format specification was clarified with supplementary documentation regarding the interpretation of the specification.

Regarding the scope of the testing of the EHR-systems, it was decided to focus on the communicated message and that the tests included quality controls in the EHR-systems of prescriptions before being approved and sent. The implementation plans of the regional health care centers, Apoteket AB, and EHR vendors, were coordinated.

An important question was how to manage the different errors identified in the new process by the various stakeholders. Three different overall validation status of the ePrescription were defined: Accepted, Rejected, and Accepted with warning.

Another question that was addressed was the manner in which the health care regions should handle and communicate the rejection of a prescription. A third question pertained to the appropriate and legal actions to be taken by the pharmacist in handling ePrescriptions with errors that are accepted with warning. The messages linked to each error status code were also discussed extensively and were revised to be sufficiently comprehensive so as to be directly communicated to the prescriber.

One major challenge was to manage all the 16 EHR-systems to be changed, tested, and approved for the new format. Apoteket AB and the health care regions made a common effort to put pressure on the software vendors and on the health care regions

to provide implementation plans to be able to implement NEF in time.

One difficult issue was deciding which organization should be responsible for administration and coordinating the development of the ePrescribing process. The test and approval process of the EHR systems were clarified, with the right to appeal for software vendors to the project organization if they were not satisfied with a decision. Both these tasks were at the end taken by a state-owned infrastructure company.

Overall, the intervention can be regarded as an effort to improve interoperability on four levels: legal, organizational, semantic, and technical (Table 1).

Analysis of Interoperability Problems in ePrescription Communication

The analysis of the ePrescription messages was made applying a specifically developed software that analyzed all collected ePrescriptions sent to Apoteket AB in Sweden during the two study periods. In brief, a test procedure was set up in the two study periods accessing all sampled ePrescriptions in the former XML-format (pre-NEF) and in the new NEF XML-format (post-NEF) respectively [2]. The collected electronic prescriptions were validated on the basis of an XML-schema that agreed with the previous format specification (pre-NEF) and according to the new NEF (for the post-NEF prescriptions) and then against the prescriptions rules implemented in the software.

The prescription rules implemented in the software were derived either from legislation or agreed praxis between the parties of the exchange. The prescription rules were validated by pharmacists and legal experts. The legal prescription rules did not change with the introduction of NEF. With NEF, prescription rules were formalized in a way suitable for implementation of controls of ePrescriptions. A control of the availability of the drugs on the Swedish market at the time of prescribing was undertaken using a historical drug database built for the selected test periods of approved and marketed drugs in Sweden.

No information about patients or prescribers, or any information that could be traced back to an individual was collected from the prescriptions. No controls were made of the medical content or adequacy of the prescription.

Classification of Errors

Two major classes of errors were identified: format errors and prescription rule errors. Format errors consist basically of valid terms and structure, data type errors, enumeration code errors, and structural errors (sequence, mandatory information, and cardinality). Format errors correspond to semantic interoperability errors (valid terms, codes, structures), where syntactic interoperability errors are the major part. The format errors were more precisely defined as a deviation from the XML-schema reflecting the different format specifications. For the post-NEF sample, the published XML-schema from the NEF-project was applied. For the pre-NEF sample, an XML-schema was constructed on the same basis and principles as the NEF-schema, but this schema adhered to the pre-NEF format specification. Format errors captured in the study are summarized in Table 3.

Table 3. Summary of format errors captured in ePrescriptions.

Format errors	Description
Incorrect code enumeration	Incorrect qualification code according to format specification
Element not defined in the specification	XML-tags not defined in the specification
Incorrect sign or format	Violating pattern constraints, such as using forbidden characters or wrong date-format
Override of maximum length	Excessive number of characters in a given field
Incomplete structure	Missing mandatory fields in a structure
Invalid data type or missing values	Incorrect data type or missing values in field (minimum length, minimum value, missing value)

Prescription rule errors were defined as legislative rules for a correct and complete prescription and also rules for agreed praxis for handling reimbursement rules, rules for communicating to the pharmacy in special cases of identification of the patient, correct references to drug identity, and valid packages for prescribed drug. Prescription rule errors correspond to legal, organizational (process), and semantic interoperability errors.

In all, 24 prescription rules were implemented and used in the validation. Prescription rule errors captured in the study are summarized in Table 4. Certain prescription rules are aggregated into one rule or collection of rules for improved readability. See Appendix 1 for a description of prescription rules used and actions taken when errors occur.

Table 4. Summary of prescription rule errors captured in ePrescriptions.

Prescription rule errors	Description
Incomplete prescriber information	Missing name, address, or telephone number
Invalid prescriber code	Incorrect format on the prescriber code
Missing workplace code	Without workplace code. The prescription can be dispensed only if the customer pays the full price for the medical drug.
Invalid reimbursement status for prescribed drug	The prescriber (or the system by default) has affirmed that the prescribed drug is valid for reimbursement, when the drug in question is not a reimbursement drug.
Incomplete or erroneous patient information	For example that the personal identification number is incorrect, or the name is missing.
Invalid drug identity	The drug identity in the prescription is not found in the database of approved and marketed drugs in Sweden at the point of issue of the prescription.
Prescription not valid for controlled substances	The prescription does not follow the specific prescription rules for these types of drugs.
Invalid combination of packages	The packages combined in the prescription for a multiple choice of a prescribed medical drug is not of the same medical product according to the drug database.
Missing directions for patient use	Text is missing when a medical drug is present in the prescription.

Statistics

Descriptive statistics were generated from databases using SQL-queries. Pearson chi-square, uncorrected for continuity, was calculated to test that no change in interoperability errors occurred. A high number of Pearson chi-square would indicate a significant improvement in interoperability in the post-NEF sample. $P<.05$ was regarded to be significant.

Results

Sampled Prescriptions

The pre-NEF sample comprised 1,270,399 prescription sets. The number of prescriptions (prescribed drugs) was 1,910,982. The mean number of prescribed drugs in each prescription set was 1.50. The post-NEF sample comprised 1,479,588 prescription sets. The number of prescriptions (prescribed drugs) was 2,204,444. The mean number of prescribed drugs in each prescription set was 1.49 (Table 5).

Table 5. Sampled prescriptions—pre-NEF and post-NEF.

Prescriptions	Pre-NEF[a]	Post-NEF
Prescription sets	1,270,339	1,479,588
Prescriptions	1,910,982	2,204,444
Mean prescribed number of prescriptions per prescription set	1.5	1.5

[a] EDIFACT prescriptions were excluded.

Dispensing Fills and Refills

According to Swedish prescription rules, the prescriber may assign, for each prescribed drug, the number of dispensing fills and refills allowed during one year. The most common case was one single fill; the second most common included four fills/refills indicating that this represents the usual treatment for one year.

Prescribed Reimbursement

The majority (95% pre-NEF and 92% post-NEF) of the prescriptions were asserted by the prescriber to be valid for reimbursement (Table 6).

Table 6. Number of prescriptions with prescribed reimbursement and mean prescribed reimbursement per prescription in pre-NEF and post-NEF samples.

	Pre-NEF		Post-NEF	
Reimbursement type	No.	Mean	No.	Mean
With reimbursement	1,810,942	94.8	2,022,957	92.8
Without reimbursement	94,971	5.0	181,487	8.2
Incorrect or missing value[a]	4225	0.2	0	-

[a] Prescriptions in the pre-NEF sample that either used old classification codes for a reimbursement type that was no longer valid or that had a missing value.

Errors per Prescription and Prescription Set

The following is a summary of the errors per prescription and prescription set (Tables 7 and 8):

- The total number of errors found in pre-NEF prescriptions was 5,970,737. The number of pre-NEF prescription sets that had at least one error was 1,253,134. The percentage of pre-NEF prescription sets with at least one error was 98.6% (1,253,134/1,270,399).
- The mean of pre-NEF prescription errors was 3.1 (5,970,737/1,910,982).
- The mean of pre-NEF prescription set errors was 4.7 (5,970,737/1,270,399).
- The mean number of errors for pre-NEF prescription sets *with* errors was 4.8 (5,970,737/1,253,134). No errors were found in 17,205 (1,270,339−1,253,134) pre-NEF prescription sets.
- The total number of errors found in post-NEF prescriptions was 13,735. The number of post-NEF prescription sets that had at least one error was 13,735.
- The percentage of post-NEF prescription sets with at least one error was 0.9% (13,735/1,479,588).
- The mean number of errors for post-NEF prescription sets *with* errors was 1.0 (13,735/13,735). No errors were found in 1,465,853 (1,479,588−13,735) post-NEF prescriptions sets. No post-NEF prescription sets that had more one error.
- The mean of post-NEF prescription errors was 0.006 (13,735/2,204,444).
- The mean of post-NEF prescription set errors was 0.009 (13,735/1,479,588).

Table 7. Summary of pre-NEF and post-NEF prescription and prescription set errors.

	Pre-NEF	Post-NEF
Total prescription sets	1,270,399	1,479,588
Prescription sets with error	1,253,134	13,735
Prescription sets with no error	17,205	1,465,853
Prescription sets with error, %	98.6	0.9
Mean error prescription sets	4.7	0.006
Mean error prescriptions	3.1	0.009

Table 8. Number of errors and mean error per prescription set.

	Pre-NEF			Post-NEF		
Error type	No. of errors	%	Mean error prescription set	No. of errors	%	Mean error prescription set
Format error	5,824,675	97.6	4.6	1273	9.3	0.0009
Prescription rule error	146,062	2.4	0.1	12,462	90.7	0.0084
Total	5,970,737	100	4.7	13,735	100	0.0093

Format errors in the pre-NEF prescriptions were the most common errors (5,824,675). Prescription rule errors in the pre-NEF sample were also common in absolute terms with 146,062 prescriptions rule errors but relatively few compared to format errors. Format errors in the post-NEF prescriptions were much less frequent compared to pre-NEF prescriptions with only 1273 errors. Format errors were relatively fewer in post-NEF compared to the pre-NEF sample. Prescription rule errors had decreased considerably to 12,462 errors in the post-NEF sample, although they had not decreased in the same proportion as format errors. Prescription rule errors have in the post-NEF sample become the most common error. To test the null-hypothesis, a chi-square test was made on the two samples (Table 9). The Pearson chi-square, uncorrected for continuity, was 2,626,673.01, $P < .0001$.

Format Errors

The distribution of format errors in the two samples were compared (Table 10). To test the null-hypothesis, a chi-square test was made on the two samples (Table 11). The Pearson chi-square, uncorrected for continuity, was 2,673,508.8. $P < .0001$

Prescription Rule Errors

The distribution of prescriptions rule errors were compared (Table 12). The largest improvement in the post-NEF sample was a decrease of Incorrect account number for patient fee, from 125,471 to 138. The second largest prescriptions rule error in the pre-NEF sample was decreased from 10,829 to 0 in the post-NEF sample. Errors that increased in the post-NEF sample were Invalid reimbursement status for prescribed drug, Invalid drug identity, Invalid multiple choice, Missing direction for patients use, and Local pharmacy destination required. To test the null-hypothesis, a chi-square test was made on the two samples (Table 13). The Pearson chi-square, uncorrected for continuity, was 141,147.86, $P <.0001$.

Table 9. Chi-square test of null-hypothesis with no significant improvement in interoperability errors.

Sample	No. of prescriptions with error	No. of prescriptions with no error	Total
Pre-NEF	1,253,134	17,265	1,270,399
Post-NEF	13,735	1,465,853	1,479,588
Total	1,266,869	1,483,118	2,749,987

Table 10. Number of format errors (XML-Schema validation errors) in pre-NEF and post- NEF prescriptions grouped by type of error.

Format error type	Pre-NEF	Post-NEF
Incorrect code enumeration	1,704,100	26
Element not defined in the specification	1,175,861	20
Incorrect sign or format	1,131,238	522
Override of maximum length	904,278	61
Incomplete structure	311,871	524
Invalid data type (not integer)	240,432	9
Override of minimum length	204,447	108
Override of minimum value	149,962	0
No amount in patient fee	2486	3
Total	5,824,675	1273

Table 11. Chi-square test of null-hypothesis with no significant improvement in interoperability.

Sample	No. of prescriptions with format error	No. of prescriptions with no format error	Total
Pre-NEF	1,252,337	18,062	1,270,399
Post-NEF	1166	1,478,422	1,479,588
Total	1,253,503	1,496,484	2,749,987

Table 12. Number of prescription rule errors grouped by type and the pre-NEF and the post-NEF sample.

	Pre-NEF		Post-NEF	
Prescription rule error type	No.	%	No.	%
Incorrect account number for the patient fee	125,471	85.9	138	1.1
Incomplete prescriber information	10,829	7.4	0	0.0
Invalid prescriber code	6279	4.3	425	3.4
Missing workplace code	1184	0.8	132	1.1
Invalid reimbursement status for prescribed drug	1007	0.7	7,589	60.9
Incomplete or erroneous patient information	895	0.6	7	0.0
Invalid drug identity	366	0.3	3,735	30.0
Prescription not valid for controlled substances	16	0.0	5	0.0
Invalid multiple choice	14	0.0	273	2.2
Missing directions for patient use	1	0.0	2	0.0
Local pharmacy destination required	0	0.0	156	1.3
Total	146,062	100.0	12,462	100.0

Table 13. Chi-square test of null-hypothesis with no significant improvement in prescription rule error.

Sample	No. prescriptions with prescription rule error	No. prescriptions without prescription rule error	Total
Pre-NEF	144,104	1,126,295	1,270,399
Post-NEF	12,172	1,467,416	1,479,588
Total	156,276	2,593,711	2,749,987

Distribution of Errors per Prescribing System

With the introduction of NEF, the tracking of each message from the prescribing system creating the prescription was made possible, which was not possible with the pre-NEF sample (Figure 6). Consequently, we do not have any comparisons between the study periods. Only data from the post-NEF study are presented here. With the introduction of NEF, it was possible to measure each system's interoperability errors (Figure 7).

Duplicate Prescriptions

With the introduction of NEF, a unique identification (UUID) was introduced for each prescription, allowing rejection of the so-called technical duplicates. A technical duplicate can occur when, for example, prescriptions are being re-sent in the case of communication failures or delays. In the post-NEF, this made it possible to measure the mean number of duplicated prescriptions from different prescribing systems (Figure 8).

Figure 6. Number of ePrescription messages sent per prescribing system.

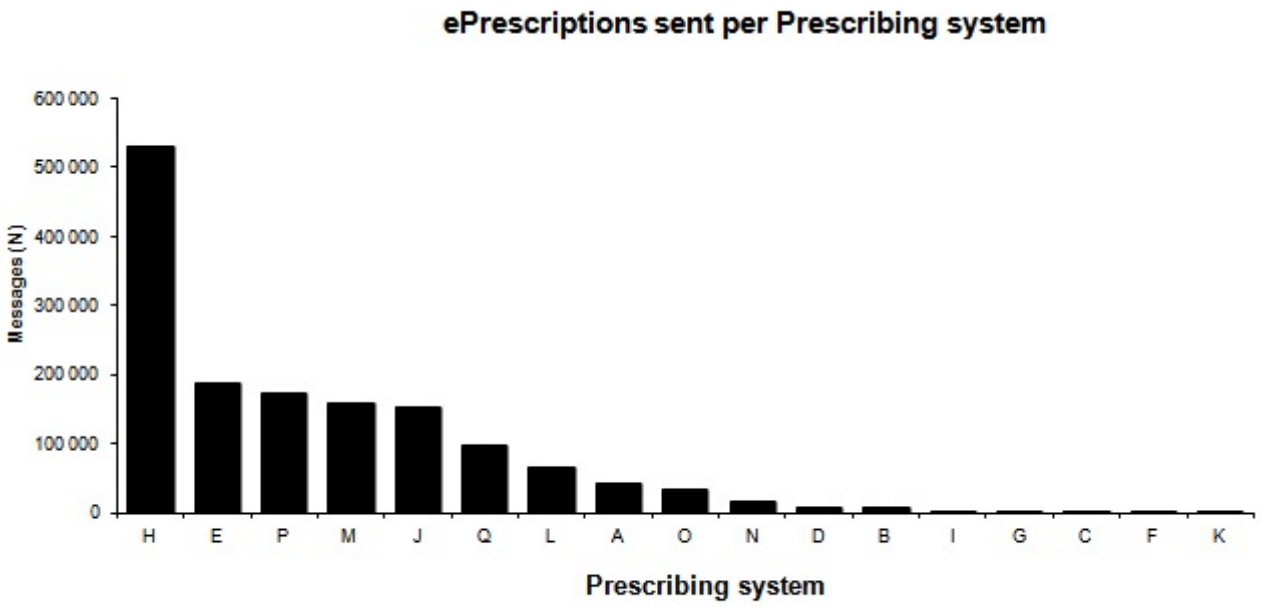


Figure 7. Mean prescription set errors (post-NEF) per Prescribing system.

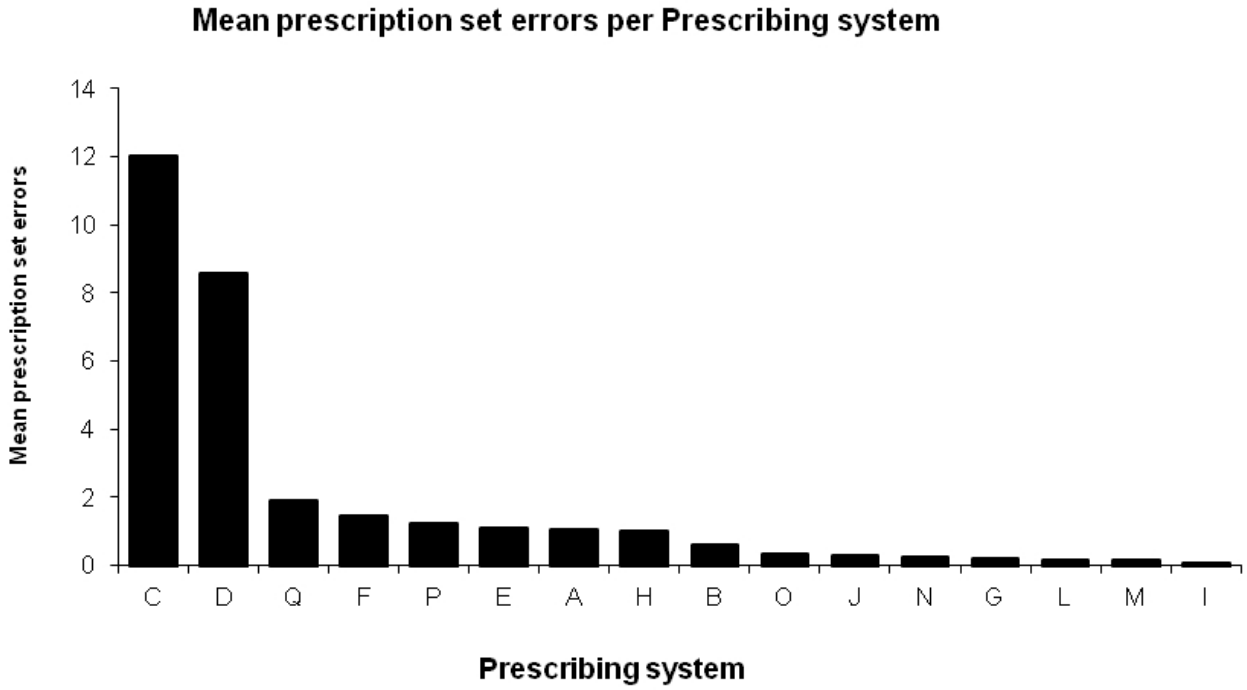
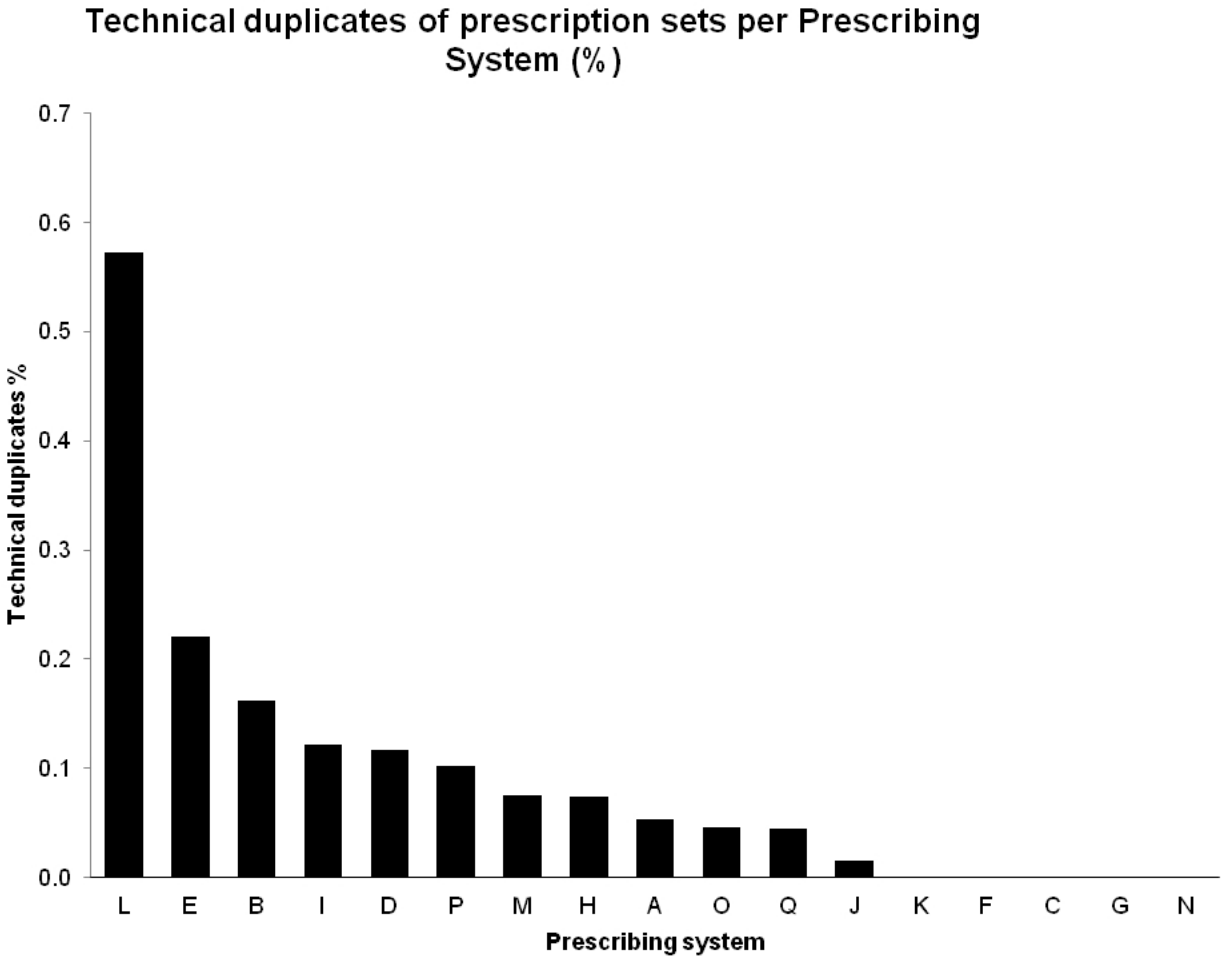


Figure 8. Technical duplicates of post-NEF prescription sets per prescribing system.



Discussion

Principal Results

The implementation of NEF substantially improved the interoperability in ePrescriptions. We have studied interoperability on four levels: legal, organizational, semantic and technical interoperability.

On the level of legal interoperability, the implementation of NEF decreased the number of ePrescriptions that were not in alignment with the legislation on prescriptions. The majority of the prescription rules errors captured concerned legal rules.

Organizational interoperability was also improved with NEF. Apart from implementing new processes such as testing and approval processes with clarification of organizational responsibilities, the process of handling certain errors was also defined, which was made possible with feedback in the acknowledge receipt. With NEF it was clarified which type of error should lead to a rejection and thus should be the responsibility of the prescriber organization to handle, and other types of errors that could be handled by the pharmacies and therefore could be accepted with a warning message. One example is the prescription rule error, missing workplace code, which was accepted with warning because it was agreed that this error, which relates to reimbursement, should not stop the dispensing of a prescription. It can be corrected by the pharmacist after contact with the prescriber. NEF changed the responsibilities in the process of error handling at both the prescriber and pharmacy organizations. Moreover, NEF made it possible to identify the EHR system sending a prescription allowing for a more systematic follow-up of errors to improve the ePrescribing process.

Semantic interoperability improved most, where the improvement in syntactic interoperability, which is part of semantic interoperability, was the most striking. A common and clearer definition of terms used in the message explains the improvement in the semantic interoperability. The improvement in the amount of errors like incorrect code enumeration, incomplete structure, and element not defined in the specification, have greatly benefited from a clearer format specification formalized in an improved XML-schema. Other examples of improvement in prescription rule errors are incorrect account number for the patient fee, incomplete prescriber information, invalid prescriber code, and missing workplace code. The first three errors resulted in a rejection of the prescription, the last one in an acceptance with warning. The improvement might be explained by a greater effort by EHR-vendors and by the health care regions to provide prescriptions with more complete information in order to avoid a rejection of prescriptions. In the pre-NEF situation, corrections

of this kind of information were done by personnel at the pharmacies.

Another aspect of semantic interoperability is the reference to objects that are defined in the infrastructure information, such as a drug register with identities of prescribed medical drugs enabling access to attributes necessary for prescribing and reimbursement, but also workplace, pharmacy, and prescriber register. Thus, semantic interoperability depends also on information sources external to the prescription and the EHR and dispensing system. Some prescription rule errors related to infrastructure information increased in proportion in the NEF-sample such as invalid drug identity. This could be explained by more frequent changes during the sample period regarding new and withdrawn drugs in combination with low frequency of updating the drug register in the EHR system. Thus, managing infrastructure information is critical in achieving and maintaining interoperability.

The technical interoperability improved too, with phasing out EDIFACT and providing a XML-schema to improve the format controls both early in the ePrescribing process, creating the prescription and later when receiving the ePrescription to the pharmacy system. The implementation of format and prescription rule controls in the ePrescribing process, particularly at the receiving end, helped to improve the interoperability of the EHR systems. Furthermore, the control and feedback process that was implemented with NEF required a faster response from the receiving process and thus made it more beneficial to use a synchronous mode of communication, like Web services. In the pre-NEF asynchronous communication, long response time was not a problem as there was no feedback. However, with NEF, new challenges arise to provide a faster response to the prescriber, which will involve not only the ePrescription services but also the technical infrastructure in health care.

The new feature with unique identifiers made it possible to measure the number of technical duplicate prescriptions for the first time. During the sample period, the rejection of duplicates had not yet been implemented. Technical duplicates are a medical risk, which could lead to drug overdose. Improvements in technical interoperability have important effects on the overall interoperability, and vice versa, the overall requirements of semantic, organizational, and legal interoperability will influence the requirements for technical interoperability.

Limitations

In this study, we have investigated the communication quality with regard to documented requirements on the ePrescription message before and after an intervention. The study of communication quality has addressed only that portion of communication quality concerned with the quality of the ePrescription message. Assessing the quality of EHR and pharmacy systems has not been within the scope of this study. Moreover, the effect of interoperability errors on the work practice, ie, its influence on the process quality, for example on medication errors, has not been studied. Other studies have addressed prescription errors from a pharmaceutical point of view [1,16-21].

Future Research

There is a need to develop practical theories and methods that can assist in creating a greater awareness and understanding to address the objective of improving the interoperation between and within different sectors and organizations using IS as an instrument for communication between different stakeholders. Without theories and methods, it is easy to fall prey to technical solutions with promise of easy ways of achieving interoperability, or that good initiatives are not implemented taking into consideration the need to involve stakeholders and grasp all levels of interoperability.

In order to improve interoperability in the overall ePrescribing process, it is necessary to analyze the different information flows and the stakeholders' roles and influence on communication and process quality. Further studies are necessary to assess the interaction quality of both prescription modules in EHR systems and prescription handling in the dispensing systems.

To improve interoperability in the ePrescribing process, the entire architecture of IS and of the stakeholders' roles in this process need to be analyzed to assess interoperability problems and to identify areas that are important to address.

To achieve a continuous improvement of interoperability, it is necessary to establish a continuous measurement of interoperability problems as a basis for improvements. How interoperability errors influence medical errors is an important topic to study in the future.

Conclusion

The introduction of NEF has considerably improved interoperability in electronic prescriptions in Sweden. This study showed that systematic and comprehensive work on interoperability, covering technical, semantic, professional, judicial, and process aspects, may lead to an important improvement in interoperability.

Acknowledgments

This study was funded by Logica Sweden AB, Apoteket AB, and the eHealth Institute at the Linnaeus University, Kalmar, Sweden.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Prescription rules and error handling.

[JPG File, 124KB - ijmr_v1i2e17_app1.jpg]

References

1. Åstrand B, Montelius E, Petersson G, Ekedahl A. Assessment of ePrescription quality: an observational study at three mail-order pharmacies. BMC Med Inform Decis Mak 2009;9:8 [FREE Full text] [doi: 10.1186/1472-6947-9-8] [Medline: 19171038]
2. Öhlund SE, Åstrand B, Petersson G. Interoperability in action - the case of electronic prescribing. In: Proceedings of the 15th International Symposium of Health Information Management Research - ISHIMR 2011, 8-9 September 2011, Zurich, Switzerland. Zurich, Switzerland; 2011 Presented at: Symposium of Health Information Management; 2011; Zurich, Switzerland.
3. Bridell G, Nordling M. Elektronisk recepthantering. In: Läkemedelsboken. Uppsala: Läkemedelsverket; 2011.
4. Öhlund SE, Goldkuhl G. Towards a socio-pragmatic understanding of ePrescribing. In: Proceedings of the 5th Intl Conference on Action in Language, Organisations and Information Systems (ALOIS-2008). 2008 Presented at: 5th Intl Conference on Action in Language, Organisations and Information Systems (ALOIS-2008); 2008; Venice URL: http://www.vits.org/publikationer/dokument/638.pdf
5. Goldkuhl G. Generic regulation model: the evolution of a practical theory for e-government. Transforming Government: People, Process and Policy 2011;5(3):249-267. [doi: 10.1108/17506161111155397]
6. Goldkuhl G. Actability Criteria for Design and Evaluation: Pragmatic Qualities of Information Systems. International Journal of Information Systems and Social Change 2011;2(3):1-15. [doi: 10.4018/jissc.2011070101]
7. Hellström L, Waern K, Montelius E, Åstrand B, Rydberg T, Petersson G. Physicians' attitudes towards ePrescribing--evaluation of a Swedish full-scale implementation. BMC Med Inform Decis Mak 2009;9:37 [FREE Full text] [doi: 10.1186/1472-6947-9-37] [Medline: 19664219]
8. IEEE. IEEE Standard Computer Dictionary. In: A Compilation of IEEE Standard Computer Glossaries. New York: IEEE; 1990.
9. Department of Defense. Department of Defense Directive, number 4630. 2004. Interoperability and Supportability of Information Technology (IT) and National Security Systems (NSS) URL: http://www.dtic.mil/whs/directives/corres/pdf/463005p.pdf [accessed 2010-01-14] [WebCite Cache ID 5mm2zkyOL]
10. European Commission. European Interoperability Framework (EIF) for European public services - Annex 2 to the Communication from the Commission to the European Parliment, the Council, the European Economic and Social Committee and the Committee of Regions 'Towards interoperability for European public services'. Bruxelles 2010.
11. International Organization for Standardization. ISO 13606 Health informatics - Electronic health record communication. ISO 2008.
12. Health Level Seven - HL7. URL: http://www.hl7.org/ [accessed 2010-01-14] [WebCite Cache ID 5mm39sonx]
13. The International Health Terminology Standards Development Organization. SNOMED CT URL: http://www.ihtsdo.org/ [accessed 2010-01-14] [WebCite Cache ID 5mm3FWixq]
14. Apotekens Service AB. Nationellt E-receptformat (in Swedish). 2007. NEF URL: http://www.apotekensservice.se/ [accessed 2010-01-14] [WebCite Cache ID 5mm3KdQ1f]
15. Brant R. Rollin Brant's Home Page. Canada: Department of Community Health Sciences, University of Calgary; 2008. URL: http://www.stat.ubc.ca/~rollin/ [accessed 2010-01-07] [WebCite Cache ID 5mbUe4NCS]
16. Åstrand B, Hovstadius B, Antonov K, Petersson G. The Swedish National Pharmacy Register. Stud Health Technol Inform 2007;129(Pt 1):345-349. [Medline: 17911736]
17. Ekedahl A. Problem prescriptions in Sweden necessitating contact with the prescriber before dispensing. Res Social Adm Pharm 2010 Sep;6(3):174-184. [doi: 10.1016/j.sapharm.2009.09.001] [Medline: 20813331]
18. Devine EB, Hansen RN, Wilson-Norton JL, Lawless NM, Fisk AW, Blough DK, et al. The impact of computerized provider order entry on medication errors in a multispecialty group practice. J Am Med Inform Assoc 2010;17(1):78-84 [FREE Full text] [doi: 10.1197/jamia.M3285] [Medline: 20064806]
19. Ammenwerth E, Schnell-Inderst P, Machan C, Siebert U. The effect of electronic prescribing on medication errors and adverse drug events: a systematic review. J Am Med Inform Assoc 2008;15(5):585-600 [FREE Full text] [doi: 10.1197/jamia.M2667] [Medline: 18579832]
20. Schedlbauer A, Prasad V, Mulvaney C, Phansalkar S, Stanton W, Bates DW, et al. What evidence supports the use of computerized alerts and prompts to improve clinicians' prescribing behavior? J Am Med Inform Assoc 2009;16(4):531-538 [FREE Full text] [doi: 10.1197/jamia.M2910] [Medline: 19390110]
21. Spina JR, Glassman PA, Simon B, Lanto A, Lee M, Cunningham F, et al. Potential safety gaps in order entry and automated drug alerts: a nationwide survey of VA physician self-reported practices with computerized order entry. Med Care 2011 Oct;49(10):904-910. [doi: 10.1097/MLR.0b013e318222a6f5] [Medline: 21666510]

Edited by K El Emam; submitted 21.02.12; peer-reviewed by K Keshavjee, F Lau, T Eguale; comments to author 18.04.12; revised version received 08.05.12; accepted 06.10.12; published 22.11.12.

Please cite as:
Öhlund SE, Åstrand B, Petersson G
Improving Interoperability in ePrescribing
Interact J Med Res 2012;1(2):e17
URL: http://www.i-jmr.org/2012/2/e17/
doi:10.2196/ijmr.2089
PMID:23612314

Original Paper

Born to Yawn? Understanding Yawning as a Warning of the Rise in Cortisol Levels: Randomized Trial

Simon BN Thompson[1], BA(Hons), MPhil, PhD(Stroke), PhD(Dem); Phil Bishop[2], BSc(Hons), MSc

[1]Psychology Research Centre, & Dementia Institute, Bournemouth University, Poole, United Kingdom

[2]Psychology Research Centre, Bournemouth University, Poole, United Kingdom

Corresponding Author:
Simon BN Thompson, BA(Hons), MPhil, PhD(Stroke), PhD(Dem)
Psychology Research Centre
& Dementia Institute
Bournemouth University
Talbot Campus
Poole House (P315), Fern Barrow
Poole, BH12 5BB
United Kingdom
Phone: 44 1202961558
Fax: 44 1202961558
Email: simont@bournemouth.ac.uk

Abstract

Background: Yawning consistently poses a conundrum to the medical profession and neuroscientists. Despite neurological evidence such as parakinesia brachialis oscitans in stroke patients and thermo-irregulation in multiple sclerosis patients, there is considerable debate over the reasons for yawning with the mechanisms and hormonal pathways still not fully understood. Cortisol is implicated during yawning and may link many neurological disorders. Evidence was found in support of the Thompson cortisol hypothesis that proposes cortisol levels are elevated during yawning just as they tend to rise during stress and fatigue.

Objectives: To investigate whether saliva cortisol levels rise during yawning and, therefore, support the Thompson cortisol hypothesis.

Methods: We exposed 20 male and female volunteers aged between 18 and 53 years to conditions that provoked a yawning response in a randomized controlled trial. Saliva samples were collected at the start and again after the yawning response, or at the end of the stimuli presentations if the participant did not yawn. In addition, we collected electromyographic data of the jaw muscles to determine rest and yawning phases of neural activity. Yawning susceptibility scale, Hospital Anxiety and Depression Scale, General Health Questionnaire, and demographic and health details were also collected from each participant. A comprehensive data set allowed comparison between yawners and nonyawners, as well as between rest and yawning phases. Collecting electromyographic data from the yawning phase is novel, and we hope this will provide new information about neuromuscular activity related to cortisol levels. Exclusion criteria included chronic fatigue, diabetes, fibromyalgia, heart conditions, high blood pressure, hormone replacement therapy, multiple sclerosis, and stroke. We compared data between and within participants.

Results: In the yawning group, there was a significant difference between saliva cortisol samples ($t_{10} = –3.071$, $P = .01$). Power and effect size were computed based on repeated-measures t tests for both the yawning and nonyawning groups. There was a medium effect size for the nonyawners group ($r = .467$) but low power (36%). Results were similar for the yawners group: medium effect size ($r = .440$) and low power (33%).

Conclusions: There was significant evidence in support of the Thompson cortisol hypothesis that suggests cortisol levels are elevated during yawning. A further longitudinal study is planned to test neurological patients. We intend to devise a diagnostic tool based on changes in cortisol levels that may assist in the early diagnosis of neurological disorders based on the data collected.

Trial Registration: International Standard Randomized Controlled Trial Number (ISRCTN): 61942768; http://www.controlled-trials.com/ISRCTN61942768/61942768 (Archived by WebCite at http://www.webcitation.org/6A75ZNYvr)

(Interact J Med Res 2012;1(2):e4) doi:10.2196/ijmr.2241

KEYWORDS

Clinical practice; cortisol; electromyography; neural pathway; saliva; yawning

Introduction

Yawning consistently poses a conundrum to neurologists and neuroscientists [1]. Increasingly, evidence is found to link neurological disorders through the commonality of yawning episodes and contagious yawning. Despite discrete incidences (such as parakinesia brachialis oscitans) in brain stem ischemic stroke patients, there is considerable debate over the reasons for yawning, with the mechanism of yawning still not fully understood [2]. Cortisol is implicated during yawning and may link many neurological disorders. Evidence was found in support of the Thompson cortisol hypothesis [3,4] that proposes cortisol levels are elevated during yawning just as cortisol levels are known to be raised in instances of stress and fatigue [5].

There have been several explanations about the yawning mechanism. Yawning is a physiological behavior that has been described as a transition between wakefulness and sleep [6]. According to Walusinski [7], yawns exteriorize the activity of the motor centers of the brain stem (cranial nerves V, VII, IX, X, XI, and XII) and of the spinal cord under the control of the hypothalamic paraventricular nucleus. The hypothalamic paraventricular nucleus is a point of integration between the central and peripheral autonomic systems. Walusinski [7] comprehensively presents several disorders due to deregulation of yawning:

- anhedonia (frustration because of an incomplete or inharmonious development of a yawn possibly due to unconscious inhibition of the letting go that underlies a complete yawn)
- the disappearance of a yawn (indicating the activity state of the dopaminergic neurons of the hypothalamic paraventricular nucleus, which are necessary for yawning)
- excessive yawning (possibly linked to hunger and arousal) and famously illustrated in Charcot's *Leçons du Mardi de la Salpêtrière* [8] by his patient who yawned eight times in a minute.

Yawning is a powerful reflex that may serve to evacuate the palatine tonsillar fossae [9]. This is a possible explanation because the strong reflex does not have any immediate urgency, is reflected in our circadian rhythm, and is allocated to times that cause us minimal inconvenience. However, McKenzie [9] suggests that, due to our social sanctions in the Western world to generally suppress the yawning reflex, perhaps we are leading to endemic tonsillitis.

Spontaneous yawning is present in humans from the early stages of development [10]. It has been observed in infants and newborns and in fetuses of 12- to 14-weeks' gestational age. The time course of yawning seems to differ with age [6], with adults yawning in the early morning and late evening [11]; in young adults, yawning seems to be linked to a low level of vigilance, increasing before and after the sleep episode [12]. Yawning is also contagious and can be elicited by seeing or even hearing someone else yawn [13]. Yet yawning has also been observed in other species [14,15], which has led to suggestions that yawning may serve communicative as well as physiological functions.

Some authors do believe that the physiological explanations given by some researchers do not adequately explain the reasons why we yawn. For example, Guggisberg and colleagues [16] argue that research tends to support yawning as a communicative function. Gallup [17] suggests it is likely that there is not just one theory to explain the functions of yawning, and it is unlikely that yawning serves primarily as a communicative function, since experimental evidence of contagious yawning is observed in only a small number of species in one lineage (primates). We tend to agree with this notion and believe the evidence of yawning in patients with neurological disorders and stroke (such as parakinesia brachialis oscitans) tends to suggest that there are specific mechanisms for yawning that are excited under special circumstances. However, it is acknowledged that yawning may be elicited because of empathy in *Homo sapiens* [18,19].

The debate for clarity continues [20]. Some authors argue that physiological explanations are imprecise and that there is evidence that neural networks responsible for empathy and social skills may be implicated and activated during the yawning episode. There is a problem with comparing some studies due to methodological differences and inadequacies [21].

However, evidence from neurological patients has led to a new line of enquiry that focuses on thermoregulation. Corey and colleagues [22] examined physiological measurements taken before, during, and after yawns in humans. They concluded their data are most consistent with the brain-cooling hypothesis and advocate that yawning increases blood flow. Indeed, it is known that painful headaches [23] and thermoregulatory disorders [24] may arise from excessive yawning. Corey and colleagues [22] suggest that the yawning experienced during these times may be due to circulatory dysfunction.

Gallup and Gallup [25] reported on repetitive yawning in patients with multiple sclerosis, showing that thermoregulatory dysfunction is a symptom of multiple sclerosis. Furthermore, yawning seems to provide symptom relief in patients with multiple sclerosis. Gallup and Eldakar [26] also showed that the incidence of yawning in humans is associated with seasonal climate variation.

Researchers are constantly striving to find commonality in disorders via their metabolic or neuronal pathways. It is interesting to note in past years how the treatment of Parkinson disease could be modified because of the exploration of dopaminergic and serotonic pathways [1,27,28]. Uncertainty in the functions of some neurotransmitters and their possible multiple implications in chemical pathways presents a complicated picture that is not unlike the clinical signature of Alzheimer disease in those with comorbid Down syndrome [29]. Having Alzheimer disease together with Down syndrome does not necessarily result in the clinical symptoms of dementia in later life [30].

The overlap between symptoms and neurochemical pathways may be more apparent than initially thought. New evidence has emerged of overlapping pathways involving DISC1, a scaffold protein that interacts with multiple neurodevelopmental, cytoskeletal, and signaling proteins [31], and Huntingdon disease [32]. In time, it is hoped that yawning and its role in neurological disorders may be understood by exploring its presence as a symptom in different neurological disorders.

Indeed, Collins and Eguibar [33] stated that antagonist interaction studies have now clearly defined at least 3 distinct neural pathways involved in the induction of yawning. Scientists are beginning to understand the hierarchical order through which these different neurotransmitter systems interact to regulate yawning. So far, the following neurotransmitters and neurohormones have been implicated: acetylcholine, dopamine, glutamate, serotonin, oxytocin, gamma-aminobutyric acid, opioids, adrenergics, nitric oxide, adrenocorticotropic hormone, and alpha-melanocyte stimulating hormone. Yet advanced techniques, such as functional magnetic resonance imaging, are yet to yield conclusive evidence to assist in fully explaining the yawn [34].

Yawning has often been associated with fatigue, stress, and exposure to cold [1]. During exposure to cold, the cortisol level in humans rises dramatically, except when they are exposed quickly (as in the cold face test), when there are reduced cortisol rises, perhaps due to vagal inhibition [35]. It is suspected that exposure to extreme cold temperature gives rise to a similar stress-like response with respect to cortisol levels in humans.

Cortisol is known to be present and elevated during stressful situations. Blood cortisol levels are directly related to salivary cortisol levels [36], which have been documented in various paradigms. The cortisol level and stress correlation is curvilinear. However, in preterm infants, cortisol levels may be lower during the heel-stick pain procedure [37], and in girls whose parents had depressive problems, cortisol levels were blunted [38]. In animal models, the cortisol level profile is also similar to that in humans during stressful situations [39]. Cortisol levels appear higher after being subjected to stress-induced situations [40].

What is unknown is the cortisol level during yawning. The cortisol level may be higher when yawning occurs after exposure to cold than after exposure to a stressful situation. Are cortisol levels elevated when neurologically impaired patients yawn, perhaps in those with multiple sclerosis?

The link between fatigue and hormonal changes is well documented. A greater level of neuromuscular fatigue and larger responses in serum hormone concentrations have been seen, for example, after hypertrophic variable resistance loadings [41]. This has led to identifying markers of fatigue [42], particularly following postmatch professional rugby [43] and in young athletes [44]. Elevated salivary cortisol levels have also been seen in elite tennis players [45]. Sleep deprivation and fatigue have been linked with salivary cortisol levels; in this instance, cortisol levels are lowered [46].

Cortisol is a lipophilic steroid with low molecular weight. Following binding with adrenocorticotropic hormone to membrane receptors and cells of the adrenal cortex, cortisol is then synthesized and released into the blood stream. Since most (about 95%) is bound to large proteins, such as albumin, only the small fraction of unbound free cortisol is thought to be biologically active *and* enters cells by passive diffusion. This makes it feasible to measure the free cortisol fraction in all bodily fluids, for example, saliva [47].

Levels of cortisol are regulated by the hypothalamic-pituitary-adrenal axis, which is a complex set of interactions between the hypothalamus (known to regulate body temperature) and the pituitary and adrenal glands. The hypothalamic-pituitary-adrenal axis also assists in digestion, the immune system, sexuality, mood, and energy usage [48]. It is implicated in stress, trauma, and particular disorders such as fibromyalgia and chronic fatigue syndrome.

Curiously, the compound glycyrrhizic acid, found in licorice, has been found to increase the activity of cortisol in the kidney [49]. This is thought to be due to inhibition of the enzyme 11B-hydroxysteroid dehydrogenase type 2, which normally inactivates cortisol in the kidney; hence, licorice tends to inhibit this enzyme and in turn deregulates, resulting in an increase of, cortisol levels. Anecdotally, we asked 1 of our study participants to eat licorice after providing a saliva sample and observed a rise in cortisol levels. This would need to be investigated further to discern significance in this finding.

Methods

We recruited 20 male and female volunteers aged between 18 and 53 years from students at Bournemouth University, Poole, UK, using a computerized recruitment system (Sona Systems, Tallinn, Estonia) and Facebook. Consent from all participants was properly obtained according to code of conduct and research guidelines. Participants were exposed, under randomized controlled trials guidelines, to three conditions intended to provoke a yawning response: photos of people yawning; boring text about yawning; and a short video of a person yawning. Comparisons were made with people exposed to the same conditions but who did not yawn.

We collected saliva samples at the start and again after a yawning response, together with electromyographic data of the jaw muscles via surface-placed electrodes to determine rest and yawning phases of neural activity. If there was no yawning response, then we took a second saliva sample at the end of the experimental paradigm. A yawning susceptibility scale (questionnaire designed for this study), Hospital Anxiety and Depression Scale (HADS) [48,50], General Health Questionnaire-28 (GHQ-28) [51-56], and demographic and health details were also collected from each participant.

Exclusion criteria were chronic fatigue, diabetes, fibromyalgia, a heart condition, high blood pressure, hormone replacement therapy, multiples sclerosis, and stroke. Between- and within-participant comparisons were made using *t* tests, and correlations were calculated using the SPSS package, version 19 (IBM Corporation, Somers, NY, USA). This enabled a comparison to be made between yawner and nonyawner

participants, as well as between rest status and yawning episodes.

Ethics

Bournemouth University research and ethics approval was granted (BU-PS5/10/11; PS1/3/12). Professional code of conduct, confidentiality, and safety issues were addressed and approved in the ethics submission. Data collected were made anonymous, coded, securely stored, and destroyed after completion of the study analysis. Protective measures were put in place for collection and analysis of saliva samples, and the right of participants to withdraw from the study was made clear to all participants.

Funding

This research received funding of £4000 from the host institution, Bournemouth University to support the purchase of essential equipment and materials. In addition, £4344 was received from Santander plc for travel expenses incurred to assist the first author to gain essential information for the selection and analysis of salivary cortisol kits.

Results

In saliva cortisol sample 1, the mean for nonyawners (n=9) was 3.3889 (SD 1.43479) and for yawners (n=11) was 2.9727 (SD 1.94889). In sample 2, the means were 4.5778 (SD 1.93589) for nonyawners and 3.9273 (SD 2.36309) for the yawners (Table 1).

There were no significant differences between groups in terms of age (t_{18} = –0.071, P = .94), HADS depression scores (t_{18} = 0.890, P = .39), GHQ-28 scores (t_{18} = 0.663, P = .52), or HADS anxiety scores (t_{18} = 0.484, P = .63). Age, depression, GHQ-28 total, and anxiety scores were not significantly correlated with sample 1 (saliva cortisol) (Table 2).

Table 1. Descriptive statistics of participants.

Yawn	Mean	SD	n
Sample 1			
Nonyawn	3.3889	1.43479	9
Yawn	2.9727	1.94889	11
Total	3.1600	1.70615	20
Sample 2			
Nonyawn	4.5778	2.93589	9
Yawn	3.9273	2.36309	11
Total	4.2200	2.58428	20

Table 2. Correlations between independent variables (n = 20) for sample 1.

Variable	Age	General total	Anxiety	Depression	Sample 1
Age					
Pearson correlation	1	–.428	–.709[a]	–.582[a]	.093
P value (2-tailed)		.06	.0	.01	.70
General total					
Pearson correlation	–.428	1	.425	.790[a]	–.185
P value (2-tailed)	.06		.06	.0	.44
Anxiety					
Pearson correlation	–.709[a]	.425	1	.750[a]	–.011
P value (2-tailed)	.0	.06		.0	.96
Depression					
Pearson correlation	–.582[a]	.790[a]	.750[a]	1	–.005
P value (2-tailed)	.01	.0	.0		.98
Sample 1					
Pearson correlation	.093	–.185	–.011	–.005	1
P value (2-tailed)	.70	.44	.98	.98	

[a] Correlation is significant at the 0.01 level (2-tailed).

Results were normally distributed, as illustrated by the Shapiro-Wilk analysis (Table 3). Results for sample 1 (saliva cortisol taken at the start) suggest that there were no significant differences between those who yawned and those who did not, as confirmed by an independent-samples *t* test (t_{18} = 0.532, *P* = .60). An independent-samples *t* test suggests that there were no significant differences in sample 2 (saliva cortisol taken either after yawning or after the last stimuli presentation if the participant did not yawn) between those who yawned and those who did not (t_{18} = 0.550, *P* = .59).

Table 3. Normality testing of data.

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
Sample	Statistic	*df*	*P* value	Statistic	*df*	*P* value
Sample 1						
Nonyawn	.199	9	.20[b]	.901	9	.28
Yawn	.218	11	.15	.919	11	.31
Sample 2						
Nonyawn	.234	9	.17	.836	9	.05
Yawn	.152	11	.20[b]	.939	11	.51

[a] Lilliefors significance correction.

[b] This is a lower bound of the true significance.

There were no significant differences in saliva cortisol between sample 1 and sample 2 for those who did not yawn during the experiment. This was confirmed using a repeated-measures *t* test (t_8 = –1.710, *P* = .13; Table 4).

Table 4. Paired-samples test for nonyawners and yawners (sample 1 – sample 2).

Study group	Mean	SD	SEM	95% CI[a] of the difference	*t* Test	*df*	*P* value (2-tailed)
Nonyawners	–1.18889	2.08533	.99511	–2.79182 to .41404	–1.710	8	.13
Yawners	–.95455	1.03089	.31082	–1.64710 to –.26199	–3.071	10	.01

[a] Confidence interval.

However, there was a significant difference in saliva cortisol between sample 1 and sample 2 among the yawners (t_{10} = –3.071, *P* = .01; Table 4).

Power and Effect Size

We computed power and effect size based on repeated-measures *t* tests for both the yawning and nonyawning groups. There was a medium effect size for the nonyawners (*r* = .467) but low power (36%). Results were similar for the yawners: medium effect size (*r* = .440) and low power (33%).

Saliva Cortisol Samples

Correlations between participants' HADS anxiety scores and sample 1 (saliva cortisol) were not significant (*r* = –.11, n = 20, *P* = 0.96); nor were correlations between participants' HADS depression scores and sample 1 (saliva cortisol; *r* = –.005, n = 20, *P* = .98).

Correlations between participants' GHQ-28 total score and sample 1 (saliva cortisol) were not significant (*r* = –.185, n = 20, *P* = .44); nor were correlations between age and sample 1 (saliva cortisol; *r* = .93, n = 20, *P* = .70).

Therefore, it was unnecessary to run an analysis of covariance, as none of the covariates were sufficiently correlated.

Normal Distribution Test

Sample 1 (saliva cortisol) was normally distributed across the whole group of participants (W (20) = 0.939, *P* = .23). It was normally distributed for the nonyawners (W (9) = 0.901, *P* = .26) and for the yawners (W (110) = 0.919, *P* = .31).

Sample 2 (saliva cortisol) was normally distributed (W (200) = 0.929, *P* = .15) across the whole group. It was normally distributed for the nonyawners (W (90) = 0.836, *P* = .53) and for the yawners (W (110) = .0939, *P* = .51).

t Tests: Between Groups

For the yawners versus nonyawners, there was no significant difference between the groups for sample 1 (saliva cortisol; t_{180} = 0.532, *P* = .60) or for sample 2 (saliva cortisol; t_{18} = 0.550, *P* = .59).

t Tests: Within Groups

There was no significant difference between sample 1 (saliva cortisol) and sample 2 (saliva cortisol) in the nonyawning group (t_8 = –1.710, *P* = .13). However, in the yawning group, there was a significant difference between samples (t_{10} = –3.071, *P* = .01).

Discussion

Several interesting findings have emerged from the study, which are consistent with the original hypothesis. Among those who yawned, there was a significant difference in cortisol levels between sample 1 and sample 2. A *t* test confirmed that there were no significant differences in salivary cortisol levels between those who yawned and those who did not for the first baseline sample (sample 1). Furthermore, there were no significant differences in a repeated-measures *t* test between sample 1 and sample 2 for those who did not yawn.

It was a concern while designing the study that age, recent anxiety and depression levels, and general health could all potentially be factors affecting participants' baseline cortisol levels. If this were the case, these factors could also have contributed toward the change in cortisol levels, and the experiment could have been measuring interference from these, rather than a change due to the independent variable (whether or not participants yawned during the experiment). However, correlation analysis suggested that these did not play a significant part in cortisol levels, with nonsignificant correlations across all variables. There were no significant differences in baseline cortisol levels between those who yawned and those who did not.

Age, recent anxiety and depression levels, or general health could also have had a potential impact on cortisol levels, which could have affected whether a participant yawned or not during the study. The *t* tests for each of the above variables confirmed that there were no significant differences in these factors between groups.

Although we used a relatively low sample size, inspection of data (Table 3) suggests that the differences between groups in terms of cortisol change between sample 1 and 2 were not vast. We intend to investigate these findings further by conducting the experiment on a larger scale, with a proposed 100 participants. Calculation using G*Power suggests that for a power size of 80%, 27 participants will be required for each group; therefore, 100 participants should permit random allocation of 50 participants per group. Investigation of participants with different neurological disorders, such as multiple sclerosis, is also planned. We hope to achieve an understanding of yawning and its role in neurological disorders, together with the potential development of a diagnostic test for the early identification of neurological sequelae.

As well as being of interest to clinical scientists and practitioners, yawning is clearly of interest in the media. A newspaper article has cited interest by the US Department of Homeland Security, which warned that apparently innocuous yawning behavior in passengers could signal a would-be terrorist [57]. Although this reported observation can be argued as perhaps being rather tangential to the pursuit of an explanation of yawning, this tends to highlight the importance of other contributing factors such as the social milieu and cultural norms.

Social and physiological factors as well as cortisol activity are all important considerations, not only because they may potentially provide the answer to why we yawn but also because they may help in the development of a potential diagnostic test. The research team led by the first author at Bournemouth University is interested in determining whether we are truly born to yawn as a protective indicator of untoward neurological dysfunction. Yawning is perhaps a warning, neurologically speaking.

Acknowledgments

Source of funding: Bournemouth University.

Ethics: Bournemouth University Research & Ethics approval granted: BU-PS5/10/11-PS1/3/12. Professional code of conduct, confidentiality, and safety issues have been addressed and approved in the Ethics submission.

Conflicts of Interest

None declared.

References

1. Thompson SB. The dawn of the yawn: is yawning a warning? Linking neurological disorders. Med Hypotheses 2010 Dec;75(6):630-633. [doi: 10.1016/j.mehy.2010.08.002] [Medline: 20797820]
2. Thompson SBN. The Dawn of the yawn: Is yawning a warning, neurologically? Webmedcentral Neurol 2011;2(2):1-8.
3. Thompson SB. Born to yawn? Cortisol linked to yawning: a new hypothesis. Med Hypotheses 2011 Nov;77(5):861-862. [doi: 10.1016/j.mehy.2011.07.056] [Medline: 21864988]
4. Thompson SBN, Zisa L. Yawning: Thompson Cortisol Hypothesis discussed. Webmedcentral Physiol 2011;2(8):1-19.
5. Thompson SBN, Zisa L. Ill-health, stress, cortisol levelsyawning. In: Thompson SBN. editor. Psychology of Trauma: Clinical Reviews, Case Histories, Research. Portsmouth: Blackwell-Harvard-Academic; 2012:125-132.
6. Baenninger R, Binkley S, Baenninger M. Field observations of yawning and activity in humans. Physiol Behav 1996 Mar;59(3):421-425. [Medline: 8700941]
7. Walusinski O. Yawning in diseases. Eur Neurol 2009;62(3):180-187 [FREE Full text] [doi: 10.1159/000228262] [Medline: 19602891]
8. Charcot JM. Leçons du Mardi de la Salpêtrière [Tuesday lessons at the Salpêtrière]. Paris: Lecrosnier & Babé; 1889.

9. McKenzie AA. The tonsillar evacuation hypothesis of yawning behaviour. S Afr J Sci 1994;90:64-66.
10. Giganti F, Salzarulo P. Yawning throughout life. Front Neurol Neurosci 2010;28:26-31. [doi: 10.1159/000307072] [Medline: 20357459]
11. Provine RR. Yawning. Am Sci 2005;93:532-539.
12. Provine RR, Hamernik HB, Curchack BB. Yawning: relation to sleeping and stretching in humans. Ethology 1987;76:152-160.
13. Provine RR. Yawning as a stereotyped action pattern and releasing stimulus. Ethology 1986;72:109-122.
14. Miller ML, Gallup AC, Vogel AR, Vicario SM, Clark AB. Evidence for contagious behaviors in budgerigars (Melopsittacus undulatus): an observational study of yawning and stretching. Behav Processes 2012 Mar;89(3):264-270. [doi: 10.1016/j.beproc.2011.12.012] [Medline: 22209955]
15. Campbell MW, de Waal FB. Ingroup-outgroup bias in contagious yawning by chimpanzees supports link to empathy. PLoS One 2011;6(4):e18283 [FREE Full text] [doi: 10.1371/journal.pone.0018283] [Medline: 21494669]
16. Guggisberg AG, Mathis J, Schnider A, Hess CW. Why do we yawn? Neurosci Biobehav Rev 2010 Jul;34(8):1267-1276. [doi: 10.1016/j.neubiorev.2010.03.008] [Medline: 20382180]
17. Gallup AC. Why do we yawn? Primitive versus derived features. Neurosci Biobehav Rev 2011 Jan;35(3):765-769. [doi: 10.1016/j.neubiorev.2010.09.009] [Medline: 20883719]
18. Norscia I, Palagi E. Yawn contagion and empathy in Homo sapiens. PLoS One 2011;6(12):e28472 [FREE Full text] [doi: 10.1371/journal.pone.0028472] [Medline: 22163307]
19. Walusinski O. Le bâillement: quand, comment, pourquoi? [Yawning: when, how, why?]. Méd Someil 2006;3(10):29-37.
20. Guggisberg AG, Mathis J, Schnider A, Hess CW. Why do we yawn? The importance of evidence for specific yawn-induced effects. Neurosci Biobehav Rev 2011 Apr;35(5):1302-1304. [doi: 10.1016/j.neubiorev.2010.12.004] [Medline: 21168437]
21. Campbell MW, de Waal FBM. Methodological problems in the study of contagious yawning. In: Walusinski O, editor. Walusinski O, editor. The Mystery of Yawning in Physiology and Disease. Frontiers of Neurology and Neuroscience. Volume 28. Basel: S Karger Pub; 2010:120-127.
22. Corey TP, Shoup-Knox ML, Gordis EB, Gallup GG. Changes in Physiology before, during, and after Yawning. Front Evol Neurosci 2011;3:7 [FREE Full text] [doi: 10.3389/fnevo.2011.00007] [Medline: 22319494]
23. Jacome DE. Primary yawning headache. Cephalalgia 2001 Jul;21(6):697-699. [Medline: 11531903]
24. Gallup AC, Gallup GG. Yawning and thermoregulation. Physiol Behav 2008 Sep 3;95(1-2):10-16. [doi: 10.1016/j.physbeh.2008.05.003] [Medline: 18550130]
25. Gallup AC, Gallup GG, Feo C. Yawning, sleep, and symptom relief in patients with multiple sclerosis. Sleep Med 2010 Mar;11(3):329-330. [doi: 10.1016/j.sleep.2009.12.007] [Medline: 20163983]
26. Gallup AC, Eldakar OT. Contagious yawning and seasonal climate variation. Front Evol Neurosci 2011;3:3 [FREE Full text] [doi: 10.3389/fnevo.2011.00003] [Medline: 21960970]
27. Stocchi F. Dopamine receptor agonists in the treatment of advanced Parkinson's disease. Parkinsonism Relat Disord 2009 Dec;15 Suppl 4:S54-S57. [doi: 10.1016/S1353-8020(09)70836-3] [Medline: 20123558]
28. Steur EN, Ballering LA. Moclobemide and selegeline in the treatment of depression in Parkinson's disease. J Neurol Neurosurg Psychiatry 1997 Oct;63(4):547 [FREE Full text] [Medline: 9343144]
29. Thompson SBN. Dementia and Memory: A Handbook for Professionals and Students. Aldershot, England: Ashgate; 2006:127-160.
30. Thompson SBN. The Central Executive System in people with Down's syndrome and dementia. Clin Gerontol 2000;21(3):3-32.
31. Camargo LM, Collura V, Rain JC, Mizuguchi K, Hermjakob H, Kerrien S, et al. Disrupted in Schizophrenia 1 Interactome: evidence for the close connectivity of risk genes and a potential synaptic basis for schizophrenia. Mol Psychiatry 2007 Jan;12(1):74-86. [doi: 10.1038/sj.mp.4001880] [Medline: 17043677]
32. Boxall R, Porteous DJ, Thomson PA. DISC1 and Huntington's disease: overlapping pathways of vulnerability to neurological disorder? PLoS One 2011;6(1):e16263 [FREE Full text] [doi: 10.1371/journal.pone.0016263] [Medline: 21298101]
33. Collins GT, Eguibar JR. Neuropharmacology of yawning. In: Walusinski O, editor. Walusinski O, editor. The Mystery of Yawning in Physiology and Disease. Frontiers of Neurology and Neuroscience. Volume 28. Basel: S Karger Pub; 2010:90-106.
34. Nahab FB. Exploring yawning in neuroimaging. In: Walusinski O, editor. Walusinski O, editor. The Mystery of Yawning in Physiology and Disease. Frontiers of Neurology and Neuroscience. Volume 28. Basel: S Karger Pub; 2010:128-133.
35. Marca RL, Waldvogel P, Thörn H, Tripod M, Wirtz PH, Pruessner JC, et al. Association between Cold Face Test-induced vagal inhibition and cortisol response to acute stress. Psychophysiology 2011 Mar;48(3):420-429. [doi: 10.1111/j.1469-8986.2010.01078.x] [Medline: 20667035]
36. Aardal-Eriksson E, Karlberg BE, Holm AC. Salivary cortisol--an alternative to serum cortisol determinations in dynamic function tests. Clin Chem Lab Med 1998 Apr;36(4):215-222. [doi: 10.1515/CCLM.1998.037] [Medline: 9638346]
37. Grunau RE, Holsti L, Haley DW, Oberlander T, Weinberg J, Solimano A, et al. Neonatal procedural pain exposure predicts lower cortisol and behavioral reactivity in preterm infants in the NICU. Pain 2005 Feb;113(3):293-300. [doi: 10.1016/j.pain.2004.10.020] [Medline: 15661436]

38. Bouma EM, Riese H, Ormel J, Verhulst FC, Oldehinkel AJ. Self-assessed parental depressive problems are associated with blunted cortisol responses to a social stress test in daughters. The TRAILS Study. Psychoneuroendocrinology 2011 Jul;36(6):854-863. [doi: 10.1016/j.psyneuen.2010.11.008] [Medline: 21185125]
39. Beerda B, Schilder MBH, Van Hooff JARAM, De Vries HW, Mol JA. Behavioural and hormonal indicators of enduring environmental stress in dogs. Anim Welf 2000;9(1):49-62.
40. Karlson B, Eek F, Hansen AM, Garde AH, Ørbæk P. Cortisol variability and self-reports in the measurement of work-related stress. Stress Health 2011;26(2):e11-e24.
41. Walker S, Taipale RS, Nyman K, Kraemer WJ, Häkkinen K. Neuromuscular and hormonal responses to constant and variable resistance loadings. Med Sci Sports Exerc 2011 Jan;43(1):26-33. [doi: 10.1249/MSS.0b013e3181e71bcb] [Medline: 20473217]
42. Bresciani G, Cuevas MJ, Molinero O, Almar M, Suay F, Salvador A, et al. Signs of overload after an intensified training. Int J Sports Med 2011 May;32(5):338-343. [doi: 10.1055/s-0031-1271764] [Medline: 21380974]
43. McLellan CP, Lovell DI, Gass GC. Markers of postmatch fatigue in professional Rugby League players. J Strength Cond Res 2011 Apr;25(4):1030-1039. [doi: 10.1519/JSC.0b013e3181cc22cc] [Medline: 20703169]
44. Locke S, Osborne M, O'Rourke P. Persistent fatigue in young athletes: measuring the clinical course and identifying variables affecting clinical recovery. Scand J Med Sci Sports 2011 Feb;21(1):90-97. [doi: 10.1111/j.1600-0838.2009.00983.x] [Medline: 19895383]
45. Gomes RV, Coutts AJ, Viveiros L, Aoki MS. Physiological demands of match-play in elite tennis: a case study. Eur J Sport Sci 2011;11(2):105-109. [doi: 10.1080/17461391.2010.487118]
46. Carev M, Karanović N, Bagatin J, Matulić NB, Pecotić R, Valić M, et al. Blood pressure dipping and salivary cortisol as markers of fatigue and sleep deprivation in staff anesthesiologists. Coll Antropol 2011 Jan;35 Suppl 1:133-138. [Medline: 21648323]
47. Kirschbaum C. University of California, San Francisco, Research Network on SES & Health. 2011. Salivary Cortisol: Technical Issues URL: http://www.macses.ucsf.edu/Research/allostatic/pdf/FAQs-salivcort.pdf [accessed 2012-08-25] [WebCite Cache ID 6ABM0kIWj]
48. Dellwo A. about.com. 2011 Jun 28. HPA Axis URL: http://chronicfatigue.about.com/od/cfsglossary/g/hpa_axis.htm [accessed 2012-08-25] [WebCite Cache ID 6ABM2DvmU]
49. Farese S, Kruse A, Pasch A, Dick B, Frey BM, Uehlinger DE, et al. Glycyrrhetinic acid food supplementation lowers serum potassium concentration in chronic hemodialysis patients. Kidney Int 2009 Oct;76(8):877-884. [doi: 10.1038/ki.2009.269] [Medline: 19641483]
50. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983 Jun;67(6):361-370. [Medline: 6880820]
51. Abiodun OA. A validity study of the Hospital Anxiety and Depression Scale in general hospital units and a community sample in Nigeria. Br J Psychiatry 1994 Nov;165(5):669-672. [Medline: 7866683]
52. Goldberg D. Manual of the General Health Questionnaire. 3rd edition. Windsor: NFER Publishing; 1978.
53. Banks MH. Validation of the General Health Questionnaire in a young community sample. Psychol Med 1983 May;13(2):349-353. [Medline: 6878521]
54. Lobo A, Pérez-Echeverría MJ, Artal J. Validity of the scaled version of the General Health Questionnaire (GHQ-28) in a Spanish population. Psychol Med 1986 Feb;16(1):135-140. [Medline: 3961039]
55. Goldberg D. Use of the general health questionnaire in clinical work. Br Med J (Clin Res Ed) 1986 Nov 8;293(6556):1188-1189. [Medline: 3096424]
56. Bridges KW, Goldberg DP. The validation of the GHQ-28 and the use of the MMSE in neurological in-patients. Br J Psychiatry 1986 May;148:548-553. [Medline: 3779225]
57. Golgowski N. Mail Online. 2012 Mar 17. Yawning in the Line at Security? You're a Potential Terrorist, Warns Homeland Security URL: http://www.dailymail.co.uk/news/article-2116336/Goose-bumps-yawning-Homeland-Security-lists-potential-signs-spotting-terrorist.html?ITO=1490 [accessed 2012-08-25] [WebCite Cache ID 6ABM4TkIE]

Abbreviations

GHQ-28: General Health Questionnaire-28
HADS: Hospital Anxiety and Depression Scale

Edited by G Eysenbach; submitted 26.06.12; peer-reviewed by N Jones, H Chinnery; comments to author 20.07.12; revised version received 23.07.12; accepted 27.07.12; published 20.09.12.

Please cite as:
Thompson SBN, Bishop P
Born to Yawn? Understanding Yawning as a Warning of the Rise in Cortisol Levels: Randomized Trial
Interact J Med Res 2012;1(2):e4
URL: http://www.jmir.org/2012/2/e4/
doi:10.2196/ijmr.2241
PMID:23611879

Original Paper

Development of the Bullying and Health Experiences Scale

Tanya Beran[1*], BA MSc PhD; Lauren Stanton[2*], PhD; Ross Hetherington[3*], PhD; Faye Mishna[4*], PhD; Shaheen Shariff[5*], PhD

[1]University of Calgary, Department of Community Health Sciences, Calgary, AB, Canada

[2]Catholic District School Board of Eastern Ontario, Ancaster, ON, Canada

[3]University of Toronto, Departments of Pediatrics and Public Health Sciences, Toronto, ON, Canada

[4]University of Toronto, Factor-Inwentash Faculty of Social Work, Toronto, ON, Canada

[5]McGill University, Department of Integrated Studies, Montreal, QC, Canada

[*]all authors contributed equally

Corresponding Author:
Tanya Beran, BA MSc PhD
University of Calgary
Department of Community Health Sciences
3330 Hospital Dr. NW
Calgary, AB, T2N 4N1
Canada
Phone: 1 1 403 220 5667
Fax: 1 1 403 210 7507
Email: tnaberan@ucalgary.ca

Abstract

Background: Until recently, researchers have studied forms of bullying separately. For 40 years, research has looked at the traditional forms of bullying, including physical (eg, hitting), verbal (eg, threats), and social (eg, exclusion). Attention focused on cyberbullying in the early 2000s. Although accumulating research suggests that bullying has multiple negative effects for children who are targeted, these effects excluded cyberbullying from the definition of bullying.

Objective: This paper responds to the need for a multidimensional measure of the impact of various forms of bullying. We used a comprehensive definition of bullying, which includes all of its forms, to identify children who had been targeted or who had participated in bullying. We then examined various ways in which they were impacted.

Methods: We used an online method to administer 37 impact items to 377 (277 female, 100 male) children and youth, to develop and test the Bullying and Health Experience Scale.

Results: A principal components analysis of the bullying impact items with varimax rotation resulted in 8 factors with eigenvalues greater than one, explaining 68.0% of the variance. These scales include risk, relationships, anger, physical injury, drug use, anxiety, self-esteem, and eating problems, which represent many of the cognitive, psychological, and behavioral consequences of bullying. The Cronbach alpha coefficients for the 8 scales range from .73 to .90, indicating good inter-item consistency. Comparisons between the groups showed that children involved in bullying had significantly higher negative outcomes on all scales than children not involved in bullying.

Conclusions: The high Cronbach alpha values indicate that the 8 impact scales provide reliable scores. In addition, comparisons between the groups indicate that the 8 scales provide accurate scores, with more negative outcomes reported by children involved in bullying compared to those who are not involved in bullying. This evidence of reliability and validity indicates that these scales are useful for research and clinical purposes to measure the multidimensional experiences of children who bully and are bullied.

(Interact J Med Res 2012;1(2):e13) doi:10.2196/ijmr.1835

KEYWORDS

mental health; school bullying; cyberbullying; peer victimization; psychosocial impact; children

Introduction

Encountering bullying behavior is one of the most distressing experiences for children and adolescents, particularly when such behaviors occur over a prolonged period of time [1,2]. Bullying behavior is strongly associated with a number of detrimental cognitive, psychological, health, and behavioral outcomes that can persist into adulthood [3,4]. Although researchers have studied bullying since the 1970s [5], few measures of children's functioning related to bullying exist. Considering that its newest form, cyberbullying, occurs electronically, it seems reasonable to consider developing a bullying and health experiences scale that could be administered electronically. The purpose of this study was to develop, pilot, and evaluate an online, multidimensional instrument to assess the health experiences of children who are bullied and/or bully others.

Health Experiences Associated with Bullying

A substantial body of research supports specific deleterious correlates of bullying [6]. In terms of cognitive functioning, victims of bullying report lower school attachment and lower academic achievement compared to children not involved in bullying [7,8].

Researchers have also documented psychological and health factors related to bullying. Findings from a national study that examined 123,227 children 11, 13, and 15 years of age residing in 28 countries found bullying was strongly related to psychological symptoms, including nervousness, sleep difficulty, and feelings of rejection, loneliness, and helplessness [9]. Other studies have found that frequent exposure to bullying is significantly related to internalizing disorders, with signs and symptoms of anxiety, depression, social withdrawal, diminished self-esteem or diminished sense of self-worth, and suicide ideation [1,10,11]. Fekkes [10] extended previous research inquiries into the relationship between mental health symptoms and bullying using longitudinal data. Findings suggest that children who were regularly bullied at the beginning of the school year were more likely to develop new mental health and health-related symptoms throughout the year, including depression, anxiety, bed-wetting, abdominal pain, and tension [10]. Similar to findings related to traditional bullying, children who are bullied online are significantly more likely to experience psychological symptoms than children not involved in bullying [12].

In addition to these cognitive and psychological symptoms, many researchers have demonstrated a number of behavioral problems related to bullying. Similar to victims of traditional bullying, victims of cyberbullying are likely to display externalizing behaviors, such as drug and alcohol use [12]. In addition, they are more likely to report carrying a weapon to school compared to children who have not experienced online bullying [12]. Other studies have found that frequent exposure to bullying experiences is linked with higher reports of eating disorders, such as bulimia nervosa, among female youth [13].

Although they may be disliked by the peers they target, aggressive children experience average levels of popularity [5]. Additionally, they do not necessarily exhibit low levels of self-esteem [14], are not highly anxious [5,14], nor are they highly depressed [14]. Nevertheless, the negative effects of bullying are not restricted to victimized children. Research has also documented negative health experiences for children who perpetrate the bullying. For example, they tend to have cognitive difficulties such as poor academic achievement [15,16]. They also tend to require mental health services [17]. They may exhibit conduct problems, such as vandalism [18], smoking [16], and drinking [16,18]. In addition, Holmes and Brandenburg-Ayres [19] reported that early bullying experiences significantly predicted later gang membership and incarceration.

In addition to children who bully and those who are victimized, there exists a group of children who both bully and are victimized (bully/victims). Hanish and Guerra [20] described these children as having the most disturbed functioning and note that they "are more likely to have emotional, behavioral, social, academic, and family problems." Children who both bully others and are bullied are considered the most vulnerable to negative developmental outcomes [21].

Despite years of research, measures typically assess one or two of the cognitive, psychological, or behavioral dimensions discussed. The purpose of our research, therefore, is to create and test a multidimensional measure of the health experiences associated with bullying.

The Internet as a Method of Data Collection

The Internet is an expedient means of transferring information and is recognized as a valid method of collecting qualitative and quantitative data [22,23]. It is also becoming an effective means of delivering general and personal information about health [24,25]. It is superior to traditional paper-based survey methods because it allows greater access to respondents through mass distribution and eliminates data entry errors arising from transcribing responses, which are entered into a database [26]. In addition, online data collection yields higher rates of completion than mail-out surveys [27]. Although it is not possible to verify the identity of the respondent, the validity of data from any method of data collection is always in question. Recall bias, misinterpretation of questions, fabrication of information, and so on are inherent to all data collection methods. It may be advantageous to use electronic methods with certain populations about certain topics of study. Given that youths ages 12-17 years use the Internet more than any other age group [28], they are likely to have access to and skill in using it for a variety of purposes. Therefore, it seems appropriate to use the Internet to ask youths questions about their behaviors while on the Internet. Using the Internet as a means of data collection is likely to provide valid and generalizable results about youths' Internet experiences [29].

Power and dominance exerted through aggression are endemic in human societies. Although researchers have investigated this phenomenon for decades, multidimensional measures of children's functioning related to bullying have not yet been developed. The purpose of our study was to comprehensively examine the health experiences of all forms of bullying using a web-based survey, and to determine whether it yields reliable and valid information when compared across groups of children who are bullied, bully others, neither, or both. We predict that

children who are bullied and/or bully others are more likely to experience problems across cognitive, psychological, and behavioral domains of functioning than children who are not bullied.

Methods

Sample

We planned to recruit a convenience sample of 200 children with an open online survey. We selected the Kids Help Phone website for recruitment because of its widespread use among youth [30]. Based on the average of 15,000 unique visits to the Kids Help Phone home page per week, and a response rate of 1%, we estimated it would take two weeks to recruit our sample. We actually recruited 377 subjects in two weeks. The completion rate was 100%. The sample included 377 children (n = 277 girls, n = 100 boys) from age 10 to 17 years (M = 13.90, SD = 1.84). Most children lived in urban (n = 307) compared to rural areas (n = 70); were born in Canada (n = 319) compared to outside of Canada (n = 58, all children were living in Canada at the time of the survey); and spoke English at home (n = 280) compared to another language (n = 97).

Procedure

The Kids Help Phone website has been in existence since 1996 and provides professional counseling and referrals to people aged 5 to 20 years on any problem or concern. Through word-of-mouth and advertising, many children in Canada have become aware of this resource. Participants were recruited by placing a click-through badge with a message about the survey on the webpage. Visitors were shown a distorted password that they typed in. This password served as a security check to prevent automated programs from accessing the questionnaire. In cases where someone responded twice, we included only the first response.

The Bullying and Health Experiences Scale

Respondents first read the following definition:

> *There are lots of ways to hurt someone. A person who bullies wants to hurt the other person. A person who bullies does it because they can. They may be older, stronger, bigger, or have other students on their side.*

Then respondents were provided with examples of forms of bullying, as shown in Table 1.

Table 1. Forms of bullying and examples provided to respondents.

Form of Bullying	Examples
Physical	Hitting, kicking, or spitting
Verbal	Name calling, mocking, humiliating, or hurtful teasing
Social	Leaving someone out, gossiping, or spreading rumors
Electronic	Bullying on Facebook, MSN, email, or text messaging
Racial	Saying hurtful things about someone whose skin is a different color
Sexual	Kissing, hugging, grabbing, pinching, and saying something sexual
Sexual preference	Teasing someone for being gay whether they are or not

Children indicated whether and how often they experienced any of the listed forms of bullying within the last month using a Likert scale with anchors from “no” to “several times a week.” A similar question was administered about bullying others.

Children were then administered 37 items asking about their cognitive, psychological, and behavioral experiences. These items were obtained from a review of the research and various measures of children’s functioning. These were rated on a Likert scale according to their frequency using the following anchors: “never,” “only once or twice,” “two or three times a month,” “about once a week,” and “several times a week.” All items included a “no” response option. The questions constituted 8 subscales, as shown in Table 2.

Table 2. Subscales for rating behavioral experiences survey.

Subscale	Number of questions	Example
Positive relationships towards peers.	6	“Other students are kind and helpful to me.”
Anger	6	“I felt angry.”
Anxiety	7	“I worry a lot.”
Self-esteem	3	“I like myself.”
Risk behaviors	5	“I have been in a physical fight.”
Physical injury caused by being bullied	3	“I was hurt by someone at school, enough to need bandages or a doctor.”
Eating problems	4	“I threw up because I was upset.”
Drug use	3	“I have used drugs.”

Results

For the questions measuring bullying victimization and perpetration, the Cronbach alphas were .77 and .71, respectively, indicating good internal consistency. To summarize, 31.3% of children reported being victimized in some way at least 2-3 times in the past month, and 11.1% reported bullying others at least 2-3 times in the past month. This is consistent with previous research [31]. Pearson product moment correlation coefficients among types of bullying perpetrated (r = 0.06 to 0.64, Mean = 0.28) and experienced (r = 0.016 to 0.54, Mean = 0.33) were low, suggesting that there was little overlap across different forms of bullying.

We analyzed the factor structure of children's reports of health experiences. The Bartlett's test of sphericity (Chi-square = 9346.06, $P < .001$) and the Kaiser-Meyer-Olkin measure of sampling adequacy (0.88) indicated that the sample was sufficient to evaluate the correlations and detect factors. The items were all subjected to principal components analysis with varimax rotation. The reliability was then determined by examining the internal consistency of the items that loaded under each factor. These analyses resulted in an 8-factor solution that explained 68.0% of the variance (see Table 3). These factors have eigenvalues greater than 1 and are named *anger* (Cronbach alpha = .87, 6 items), *relationships* (Cronbach alpha = .86, 6 items), *physical injury* (Cronbach alpha = .90, 3 items), *risk* (Cronbach alpha = .84, 5 items), *anxiety* (Cronbach alpha = .82, 7 items), *self-esteem* (Cronbach alpha = .92, 3 items), *drug use* (Cronbach alpha = .86, 3 items), and *eating problems* (Cronbach alpha = .73, 4 items). A few of the items had split loadings; however, reliability analyses indicated that they were consistent with other items within the factor, as shown in the table. When an item had a loading of greater than 0.40 on two components, the higher loading was used to assign it to a component [32].

Table 3. Types of health experiences determined from principal components analysis

	Anger	Relation-ships	Physical-injury	Risk	Anxiety	Self-esteem	Drug use	Eating problems
I like myself.	.049	-.114	.023	-.010	.087	.901	.028	.028
I am able to do things as well as most other people.	.016	-.048	.064	.017	.069	.910	.005	.138
I have lots of good qualities.	.034	-.077	.013	.054	.087	.919	.060	.101
I like school.	-.306	.545	.040	-.119	.055	.067	-.335	-.108
I feel alone or left out at school.	-.060	.587	-.092	.035	-.308	-.243	.071	-.103
Other students are kind and helpful to me.	-.154	.773	-.117	-.084	-.038	.064	-.159	-.117
Other students like hanging out with me.	-.124	.844	-.113	.032	-.132	-.049	-.037	-.044
Other students accept me for who I am.	-.164	.793	-.110	-.040	-.128	-.040	-.056	-.154
I make friends easily.	-.084	.791	-.023	-.117	-.154	-.142	.127	-.033
I worry a lot.	.310	-.285	-.077	-.134	.592	.082	.155	.257
I worry about my family.	.124	-.099	.010	.001	.719	.013	.007	.034
I worry about doing well in school.	-.013	.057	.044	-.092	.778	.034	-.027	-.180
I worry about making friends.	.113	-.444	.054	.112	.560	.169	-.090	-.014
I worry about the future.	.015	-.185	.014	.079	.751	.039	.026	.124
My worries kept me up at night.	.399	-.182	.035	-.022	.519	.194	.031	.428
I have trouble catching my breath.	.292	-.195	.166	-.081	.358	.149	.165	.433
I was in a physical fight.	.338	-.060	.412	.546	.052	-.009	.147	-.037
I was in a physical fight where the other person needed bandages or a doctor.	.204	.026	.507	.621	-.039	.046	.263	.036
I was in trouble with the law.	.098	-.048	.228	.680	-.079	.072	.291	.255
I took weapons to school.	.085	-.084	.116	.834	-.023	.018	.134	.164
I damaged property.	.275	-.095	.136	.598	.051	-.037	.253	.083
I was hurt by someone at school, enough to need bandages or a doctor.	.067	-.121	.831	.160	.061	.021	.073	.210
I was hurt by someone while going to and from school, enough to need bandages or a doctor.	.034	-.146	.855	.233	-.011	.072	.114	.141
I was hurt by someone in my neighborhood, enough to need bandages or a doctor.	.089	-.080	.862	.170	.078	.012	.120	.139
I smoked cigarettes.	.181	-.073	.156	.153	.038	.073	.833	.070
I drank alcohol.	.165	.020	.107	.360	.048	-.018	.704	.268
I have used drugs.	.091	-.057	.144	.313	-.008	.044	.812	.156
I felt angry.	.697	-.277	-.049	.027	.235	-.027	.071	.152
I yelled/screamed/stomped my feet when angry.	.795	-.104	.045	.059	.062	.030	.010	.172
I swore at others when angry.	.737	-.039	.027	.120	.085	-.007	.180	.183
I broke things when angry.	.727	-.076	.158	.244	.044	.045	.165	.073
I hurt someone when angry.	.614	-.140	.230	.393	-.069	.039	.047	.046
I got angry more easily with others.	.748	-.217	.055	.087	.145	.063	.071	-.002
I threw up because I was upset.	.209	-.093	.237	.169	-.033	.052	.381	.613
I could not eat because I was so upset.	.191	-.196	.072	.104	.128	.150	.122	.663
I used medicine to help me lose weight.	.080	-.027	.313	.333	-.062	.083	.083	.622
I over ate to the point where I became sick.	.153	-.263	.349	.222	.035	.042	.101	.433

We compared the scores on these 8 dimensions across several profiles, including bully, victim, non-bully/victim, and bully/victim. First, we classified children as bullies if they reported bullying others 2-3 times per month or more, *and* being bullied no more than 1-2 times per month. Similarly, we coded children as victims if they reported being bullied 2-3 times per month or more, *and* bullying others not more than 1-2 times per month. Children who bullied and were bullied 2-3 times per month or more were often coded as bully/victims. We considered children who bullied and were bullied no more than 1-2 times per month as non-bully/victims. This classification system is typically used in the research [33].

The number of children in each group is shown in Table 4. The mean scores were calculated by taking the sum of all the items that loaded under each factor and dividing by the number of items for each factor. Thus, the possible range of mean values in Table 4 is from 1 to 5. Partial η^2 provided effect size estimates in analyses of variance (ANOVAs) and were interpreted using Cohen's [34] criteria (.01 = small, .09 = medium, .25 = large). Accordingly, there was a small difference for risk, anger, physical injury, anxiety, and eating problems across the bully classifications. Also, there was a moderate to large difference in relationship experiences. Self-esteem and drug use, however, did not significantly differ across bully or victim groups.

Table 4. Mean scores of health experiences subscales for bullies, victims, bully/victims, and non-bully/victims.

	Bully(n = 19)	Victim(n = 95)	Bully/Victim(n = 23)	Non-bully/Victim(n = 240)	F*df (3, 373)	*Partial Eta 2*
Risk	1.58	1.64	1.61	1.29	6.98	.05
Relationships	3.50	2.67	2.98	3.68	27.38	.18
Anger	3.46	3.29	3.54	2.68	8.83	.07
Physical injury	1.07	1.46	1.49	1.14	6.95	.05
Anxiety	3.56	3.78	3.88	3.30	7.56	.06
Self-esteem	3.33	3.25	3.29	3.17	0.18	.00
Eating problems	1.88	2.11	2.23	1.59	10.29	.08
Drug use	1.53	1.86	1.58	1.63	1.08	.01

*All F values are significant at $P < .001$ with the exception of self-esteem and drug use.

Post hoc analyses using Tukey's HSD showed several significant differences ($P < .001$). Victimized children engaged in more risk behaviors, experienced fewer positive relationships and more physical injury, and reported higher levels of anger, anxiety, and eating problems than did children not involved in bullying. Children who admitted to bullying behaviors were more likely to report anger than were non-bully/victims. Bully/victims were more likely to report anger, anxiety, eating problems, and poor peer relationships compared to non-bully/victims. Finally, non-bully/victims were the least likely to report negative experiences.

Discussion

Based on an extensive review of the literature, we created a 37-item multidimensional scale. Through web-based administration of these items, and analyses of factorial structure and reliability, we obtained evidence of their usefulness when assessing children's health experiences. The subscales measure risk behaviors, relationship experiences, anger, physical injury, anxiety, drug use, self-esteem, and eating problems. These areas address many of the cognitive, psychological, and behavioral problems of children involved in bullying, as identified in previous research. The reliability coefficients are adequate and scores were significantly elevated for children who were victimized and/or perpetrated bullying.

Bullying Roles

Significant differences in health experiences emerged among groups according to the type of involvement in bullying. Specifically, children who reported victimization also reported engaging in risk behaviors, such as weapons possession. This is consistent with findings published by Ybarra [12]. Children may be exposed to these risk behaviors through acts of bullying, or they may participate in these risk behaviors as a coping reaction to being bullied. Participation in these risky behaviors probably increases the likelihood of physical injury, as was found in our study. Children who participated in risky behaviors also reported fewer positive relationships with peers, which is consistent with research showing that children who are bullied may experience interpersonal victimization and social skills difficulties [12]. Several studies have substantiated that children who are bullied experience high levels of anger and anxiety [1,9,10,11]. Other health problems, such as eating problems involving vomiting or limiting food intake, were more often reported by children who were bullied compared to children not involved in bullying. This finding is also substantiated by previous research [13].

Similar to children who were bullied, children who perpetrated bullying reported elevated levels of anger, as shown in previous research [35]. They did not report high levels of anxiety, which is also consistent with other studies [14]. In addition, they reported experiencing positive relationships, which supports research suggesting that bullies perceive themselves as liked by their peer group [5]. Indeed, bullying attracts attention from the peer group, which bullies may interpret as positive attention. Bullies are, moreover, likely to report similar levels of self-esteem as non-bully/victims [14], which is consistent with our findings. Contrary to our expectations, children who reported

bullying others did not report high levels of risk behaviors. The absence of reported risky behaviors in this group could be related to the relatively young age of our sample. For example, a multinational study of the prevalence of substance use indicates that peak onset of alcohol and cannabis use is in mid- to late adolescence, and at age 18 for all other drugs [36].

We also included children who were both victims and perpetrators of bullying. Several health concerns were higher for these children than those not involved in bullying. That is, the bully/victim was likely to experience poor peer relationships, anger, anxiety, and eating problems, as suggested by previous research [37]. We found, however, that these children did not report the most negative outcomes, as expected. Rather, victims did. This could be due to the relatively strong self-esteem and positive relationships that the bully/victim reported, which may buffer negative outcomes. Moreover, children who are bullied and victimized typically receive the lowest levels of social support [38], which may explain the negative outcomes documented in most research. In our study, however, they may have been receiving support from the Kids Help Phone website. This support may have lessened the detrimental correlates of bullying, resulting in lower reporting of these negative behaviors.

Significance and Implications

We chose not to use items about academic achievement and educational experiences because they did not appear to reliably or accurately measure the constructs of our scale. In addition, we did not include questions about depression. We recommend further research in item development on questions pertaining to academic achievement and depression. Moreover, considering that no single source could feasibly assess all aspects of functioning, we recommend that additional instruments be administered to provide a more comprehensive profile, particularly when designing or providing interventions or support. In addition, identifying children as bullies or victims based on experiencing any or all types of bullying created heterogeneous groups. However, given the large number of groups that would have been formed for each type and combinations of types of bullying, it was not possible to compare functioning across all groups. Generalizability of the results may be limited because participants were recruited from the Kids Help Phone website, which they may have been visiting because they were already troubled or seeking support. Although the prevalence of bullying we report is similar to other research, our participants may be unique on other unknown dimensions. Participants filled out a questionnaire via a computer using the Internet. Accordingly, this study was restricted to children who have access to a computer and the Internet. Because the survey was administered online, participants could not ask for clarification or help from the individual administering the questionnaire, which could have affected the accuracy of some responses. Finally, in common with all anonymous online research, it was not possible to verify the information provided by participants.

Despite these limitations, this study provides strong evidence for the usefulness of this multidimensional web-based survey on health experiences related to bullying. It further adds to the growing evidence of the detrimental impact of bullying among multiple profiles of children and youth, whether in the role of bully, victim, or bully/victim.

Acknowledgments

This research was supported by a Standard Research Grant #410-2007-0671 from the Social Sciences and Humanities Research Council of Canada.

Conflicts of Interest

None declared.

References

1. Brunstein Klomek A, Marrocco F, Kleinman M, Schonfeld, Gould MS. Bullying, depression, and suicidality in adolescents. J Am Acad Child Adolesc Psychiatry 2007 Jan;46(1):40-49. [doi: 10.1097/01.chi.0000242237.84925.18] [Medline: 17195728]
2. Frisén A, Jonsson AK, Persson C. Adolescents' perception of bullying: Who is the victim? Who is the bully? What can be done to stop bullying? Adolescence 2007;42(168):749-761. [Medline: 18229509]
3. Fosse GK, Holen A. Childhood maltreatment in adult female psychiatric outpatients with eating disorders. Eat Behav 2006 Nov;7(4):404-409. [doi: 10.1016/j.eatbeh.2005.12.006] [Medline: 17056418]
4. Smith JD, Cousins JB, Stewart R. Antibullying interventions in schools: Ingredients of effective programs. Canadian Journal of Education / Revue canadienne de l'éducation 2005;28:739-762. [doi: 10.2307/4126453]
5. Olweus D. Aggression in the Schools: Bullies and Whipping Boys. Washington: John Wiley & Sons Inc; 1978.
6. Hawker DS, Boulton MJ. Twenty years' research on peer victimization and psychosocial maladjustment: A meta-analytic review of cross-sectional studies. J Child Psychol Psychiatry 2000 May;41(4):441-455. [Medline: 10836674]
7. Beran TN, Hughes G, Lupart J. A model of achievement and bullying: Analyses of the Canadian National Longitudinal Survey of Children and Youth data. Educational Research 2008;50(1):25-39. [doi: 10.1080/00131880801920379]

8. Spriggs AL, Iannotti RJ, Nansel TR, Haynie DL. Adolescent bullying involvement and perceived family, peer and school relations: Commonalities and differences across race/ethnicity. J Adolesc Health 2007 Sep;41(3):283-293 [FREE Full text] [doi: 10.1016/j.jadohealth.2007.04.009] [Medline: 17707299]
9. Due P, Holstein BE, Lynch J, Diderichsen F, Gabhain SN, Scheidt P, et al. Bullying and symptoms among school-aged children: international comparative cross sectional study in 28 countries. Eur J Public Health 2005 Apr;15(2):128-132 [FREE Full text] [doi: 10.1093/eurpub/cki105] [Medline: 15755782]
10. Fekkes M, Pijpers FI, Fredriks AM, Vogels T, Verloove-Vanhorick SP. Do bullied children get ill, or do ill children get bullied? A prospective cohort study on the relationship between bullying and health-related symptoms. Pediatrics 2006 May;117(5):1568-1574 [FREE Full text] [doi: 10.1542/peds.2005-0187] [Medline: 16651310]
11. Mishna F, Scarcello I, Pepler D, Wiener J. Teacher's understanding of bullying. Canadian Journal of Education / Revue canadienne de l'éducation 2005;28:718-738. [doi: 10.2307/4126452]
12. Ybarra ML. Linkages between depressive symptomatology and Internet harassment among young regular Internet users. Cyberpsychol Behav 2004 Apr;7(2):247-257. [doi: 10.1089/109493104323024500] [Medline: 15140367]
13. Fosse GK, Holen A. Childhood maltreatment in adult female psychiatric outpatients with eating disorders. Eat Behav 2006 Nov;7(4):404-409. [doi: 10.1016/j.eatbeh.2005.12.006] [Medline: 17056418]
14. Woods S, Wolke D. Direct and relational bullying among primary school children and academic achievement. Journal of School Psychology 2004;42:135-155. [doi: 10.1016/j.jsp.2003.12.002]
15. Leff SS, Power TJ, Manz PH, Costigan TE, Nabors LA. School-based aggression prevention programs for young children: Current status and implications for violence prevention. Sch Psych Rev 2001;30:344-362.
16. Nansel TR, Overpeck M, Pilla RS, Ruan WJ, Simons-Morton B, Scheidt P. Bullying behaviors among US youth: Prevalence and association with psychosocial adjustment. JAMA 2001 Apr 25;285(16):2094-2100 [FREE Full text] [Medline: 11311098]
17. DeRosier ME, Kupersmidt JB, Patterson CJ. Children's academic and behavioral adjustment as a function of the chronicity and proximity of peer rejection. Child Dev 1994 Dec;65(6):1799-1813. [Medline: 7859556]
18. Olweus D. Bullying at School: What We Know and What We Can Do. Oxford, UK: Wiley-Blackwell; 1993.
19. Holmes SR, Bradenburg-Ayres SJ. Bullying behavior in school: A predictor of later gang involvement. J Gang Res 1998;5:1-6.
20. Hanish LD, Guerra NG. Aggressive victims, passive victims, and bullies: Developmental continuity or developmental change? Merrill-Palmer Quarterly 2004;50:17-38. [doi: 10.1353/mpq.2004.0003]
21. Parker JG, Asher SR. Peer relations and later personal adjustment: Are low-accepted children at risk? Psychol Bull 1987 Nov;102(3):357-389. [Medline: 3317467]
22. Hess-Biber SN, Leavy P. Handbook of Emergent Methods. New York, NY: The Guilford Press; 2008.
23. Hewson C. Internet-mediated research as an emergent methodits potential role in facilitating mixed methods research. In: Hesse-Biber SJ, editor. Handbook of Emergent Methods. New York, NY: The Guilford Press; 2008:543-570.
24. Brehaut JC, Santesso N, O'Connor AM, Lott A, Lindgaard G, Syrowatka A, et al. Presenting evidence to patients online: What do web users think of consumer summaries of cochrane musculoskeletal reviews? J Med Internet Res 2011;13(1):e5 [FREE Full text] [doi: 10.2196/jmir.1532] [Medline: 21247863]
25. El Emam K, Moreau K, Jonker E. How strong are passwords used to protect personal health information in clinical trials? J Med Internet Res 2011;13(1):e18 [FREE Full text] [doi: 10.2196/jmir.1335] [Medline: 21317106]
26. Reynolds RA, Woods R, Baker JD. Handbook of Research on Electronic Surveys and Measurements. Hershey, PA: IGI Global; 2007.
27. Truell AD, Bartlett JE, Alexander MW. Response rate, speed, and completeness: A comparison of Internet-based and mail surveys. Behav Res Methods Instrum Comput 2002 Feb;34(1):46-49. [Medline: 12060989]
28. Pew Internet. 2006. 2006 URL: http://www.pewinternet.org/ [accessed 2012-06-28] [WebCite Cache ID 68lIpQGQx]
29. Pitkow JE, Kehoie CM. Emerging trends in the www user population. Communications of the ACM 1996;39:106-108. [doi: 10.1145/228503.228525]
30. Kids Help Phone. 2012 URL: http://org.kidshelpphone.ca/en/about-us/facts-and-stats/ [accessed 2012-06-28] [WebCite Cache ID 68lIyfzCR]
31. Craig WM, Pepler DJ. Identifying and targeting risk for involvement in bullying and victimization. Can J Psychiatry 2003 Oct;48(9):577-582. [Medline: 14631877]
32. Kerlinger F, Lee H. Foundations of Behavioral Research. Fort Worth, TX: Harcourt College Publishers; 2000.
33. Olweus D. Bully/victim problems among schoolchildren: Basic facts and effects of a school-based intervention program. In: Rubin KH, Pepler DJ, editors. The Development and Treatment of Childhood Aggression. Hillsdale, N.J: Lawrence Erlbaum Associates, Inc; 1991.
34. Cohen J. Statistical Power Analysis for the Behavioral Sciences. Hillsdale, N.J: Lawrence Erlbaum Associates, Inc; 1988.
35. Sigfusdottir ID, Gudjonsson GH, Sigurdsson JF. Bullying and delinquency: The mediating role of anger. Personality and Individual Differences 2010;48:391-396. [doi: 10.1016/j.paid.2009.10.034]
36. Vega WA, Aguilar-Gaxiola S, Andrade L, Bijl R, Borges G, Caraveo-Anduaga JJ, et al. Prevalence and age of onset for drug use in seven international sites: results from the international consortium of psychiatric epidemiology. Drug Alcohol Depend 2002 Dec 1;68(3):285-297. [Medline: 12393223]

37. Pellegrini AD. Bullying, victimization, and sexual harassment during the transition to middle school. Educational Psychologist 2002;37:151-163. [doi: 10.1207/S15326985EP3703_2]
38. Demaray MK, Malecki CK. Perceptions of the frequency and importance of social support by students classified as victims, bullies, and bully/victims in an urban middle school. Sch Psych Rev 2003;32:471-489.

Edited by G Eysenbach; submitted 20.04.11; peer-reviewed by J Green; comments to author 01.09.11; revised version received 17.07.12; accepted 03.10.12; published 09.11.12.

Please cite as:
Beran T, Stanton L, Hetherington R, Mishna F, Shariff S
Development of the Bullying and Health Experiences Scale
Interact J Med Res 2012;1(2):e13
URL: http://www.i-jmr.org/2012/2/e13/
doi:10.2196/ijmr.1835
PMID:23612028

Original Paper

Use of a Web Portal for Support and Research After a Disaster: Opportunities and Lessons Learned

Geertruid MH Marres[1], MD, PhD; Luke PH Leenen[1], MD PhD FACS; Johannes W van der Slikke[2], MD PhD; Eric Vermetten[1], MD PhD

[1]University Medical Centre Utecht, Central Military Hospital, Major Incident Hospital, Utrecht, Netherlands

[2]VU University Medical Centre, Amsterdam, Netherlands

Corresponding Author:
Geertruid MH Marres, MD, PhD
University Medical Centre Utecht
Central Military Hospital
Major Incident Hospital
Heidelberglaan 100
Utrecht, 3508 GA
Netherlands
Phone: 31 88 7559882
Fax: 31 88 7555015
Email: gmarres@hotmail.com

Abstract

Background: In this report we describe the development and use of a web portal in the aftermath of the 2004 tsunami. This large scale disaster confronted many displaced people with death, despair and need for information and support. Awareness and insight in the emotional impact of disasters can provide opportunities for surveillance and early treatment. Moreover, online support systems can contribute to community building, empowerment of victims and resilience.

Objective: We evaluate the development and use of a multilingual web portal that combined a platform for information, emotional support, self assessment and referral with research opportunities. The rapid development, use, advantages, difficulties and learning points are discussed.

Methods: A multidisciplinary working group from the University Medical Centre Utrecht, the Major Incident Hospital and the Central Military Hospital developed a web portal for tsunami victims. The webportal combined: (1) a forum aimed at community building, (2) self assessment tools that in the same time function as a reseach survey, (3) e-consultation, and (4) an information portal.

Results: Within 3 weeks after the tsunami, the working group launched an open, online service (www.TISEI.org. Tsunami Intrenational Survey on Emotional Impact) to foster community) support in the aftermath of the disaster. It combined four functionalities that were earlier previously only used separately. The portal had over 36.800 unique visitors in the first two years. At least 31% (144/464) percent of the Dutch surviving victims could be reached for a survey through the site. The TISEI-environment was available in 15 languages and visitors came from all over the world. Ninety-five percent of all visitors came from Europe or the United States. Subsequent to immediate disaster support, the web portal also served as a memorial archive for anniversary meetings and follow-up incentives. Difficulties we experienced were lack of funding, time pressure, victim-anonymisation, international collaboration and long term maintenance.

Conclusions: A multilingual website with combined modalities for emotional care and research after a natural disaster proved feasible. Web based services like www.TISEI.org in the aftermath of mass disasters can help community building and deliver low level, patient centred and easily accessible information and care. A multilingual website with combined modalities for emotional care and research after a natural disaster proved feasible. Growing Internet penetration world wide and especially the rapid expansion and influence of online communities enables delivery of care and perform research with the internetInternet as a platform. The unpredictable nature of disaster does put time pressure on the development of online solutions and influenced the yield of our site. This highlights the necessity of developing methods and (inter) national collaborations in advance, secure funding, and learn from earlier initiatives.

(Interact J Med Res 2012;1(2):e18) doi:10.2196/ijmr.1588

KEYWORDS

Disaster medicine; Stress Disorders; Post-Traumatic; Internet; mental health; health surveys; stress, psychological; Online Systems; Self-Help groups

Introduction

On December 26th 2004, a massive undersea earthquake caused a giant shockwave or tsunami that devastated the shorelines of Indonesia, Sri Lanka, India, Thailand and many other countries. More than two hundred thousand people from all over the world were reported missing or dead. Most victims were local citizens but many foreign tourists celebrating Christmas holidays were also hit. Large numbers of survivors were repatriated to their home countries. Among the 500 Dutch visitors in the region, 36 were killed. Several were wounded and were repatriated home in the following days. On January 1st 2005, 23 injured survivors were flown to the Major Incident Hospital (MIH) in Utrecht, the Netherlands. A spontaneous call for a communication platform was heard, in which survivors would hope to connect with the community they were part of as well as other survivors. In the days after the tsunami, numerous web logs and local initiatives were posted. In response to the call for help from survivors, we decided to initiate an online resource providing disaster information, to enable contact among affected persons, and facilitate public connection with mental health professionals.

In disasters, apart from the immediate consequences reflected in the death toll, physical injuries, destruction, and economic consequences, the psychological impact of exposure to a traumatic event can lead to prolonged consequences for mental health. The most notably documented effect of coping with trauma is post-traumatic stress disorder (PTSD) and depression [1,2].

Earlier studies demonstrated that acknowledgement of other survivors can reduce feelings of isolation and have a preventive effect on development of psychological problems [3]. Fellow survivors can provide mutual support, which is often perceived to be more meaningful than help from others who have not experienced the same traumatic events. Online support systems can facilitate this community building by providing a channel for communication amongst victims and health professionals. The value to participants of such virtual forums designed to give and receive emotional support has been documented, although the effect of online communities in " regular " mental health care per se has not been equivocally established with high quality scientific evidence yet [4-7].

The psychological aftermath associated with disasters can be managed via a public health approach. Existing services can extend to target mental health care as well as the distribution and availability of resources. The rapid growth of Internet and influence of online communities already in our lives illustrate its opportunities to foster resilience and deliver care after catastrophic events.

Besides mutual support, the Internet can enable self-assessment [8,9] and self-referral. This can stimulate empowerment and a kind of self-triage, which could prevent development of diseases like PTSD and depression in the longer term. Furthermore it could be an instrument to assess the emotional impact of the Tsunami on a group level. Earlier online initiatives in the aftermath of disasters, like surveys after "9/11"[10,11] show the potential of the Internet for health and disaster related research. Awareness and insight in psychological impact of natural disasters on the survivors and the community as a whole has the potential to improve quality of information and support and to aid in planning early treatment options.

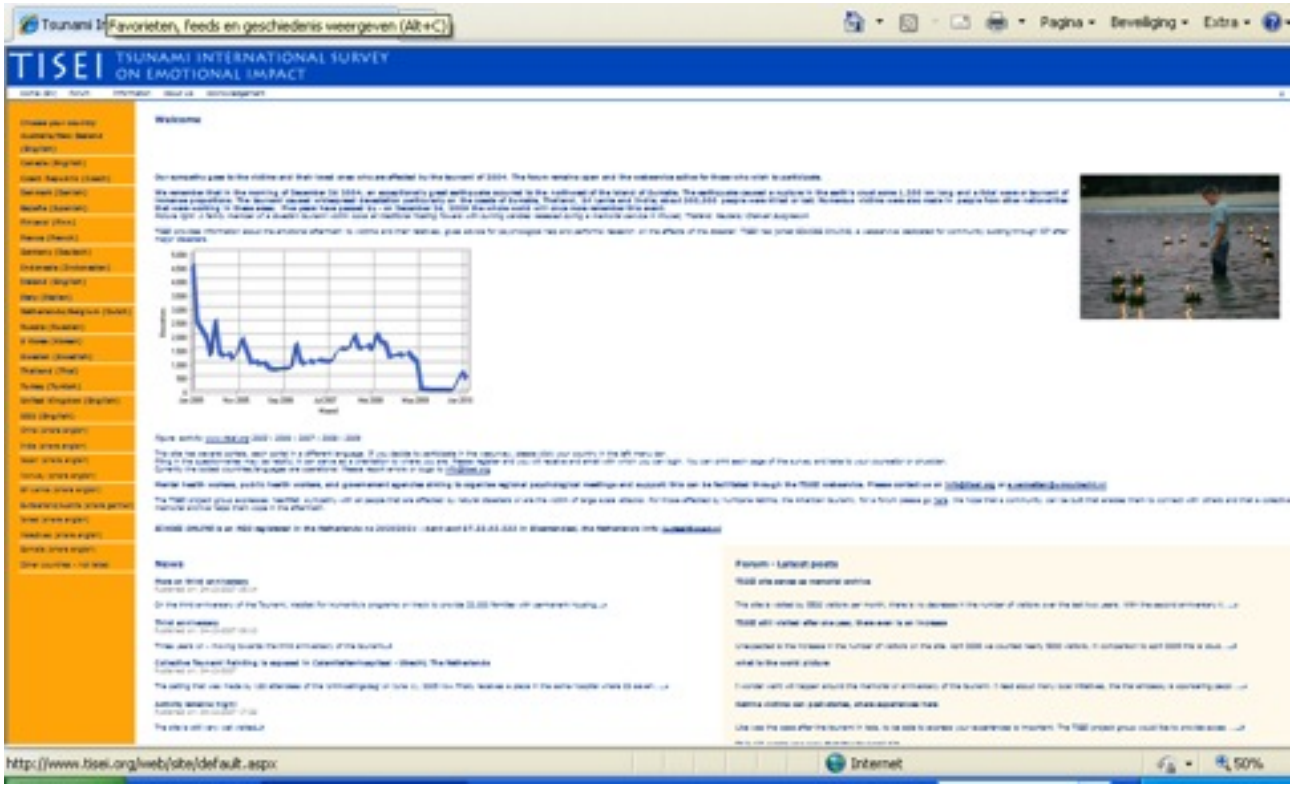
In this paper we report the development and use of a web portal as an aid in the management of emotional impact after the Tsunami disaster. We evaluate the feasibility, launching speed, and the development process of a multilingual site that combined a platform for emotional support, self-assessment and referral with research opportunities. The strengths, the faced difficulties and our learning points are discussed.

Methods

Two days after admission of the patient group to the MIH the chairman of the executive board of the University Medical Centre Utrecht (UMCU) granted support to develop the web portal project. Funding for building and hosting the website was not sufficient at the time we started but was gradually formed along the way. Most of the researchers were voluntary clinicians (psychiatrists, surgeons, psychologists) and IT-specialists. Within 3 weeks after the Tsunami, an open, online service (www.TISEI.org) was launched to foster (community) support in the aftermath of the disaster.

We developed a web service with four functionalities (Figure 1): 1) information portal; 2) forum; 3) self-assessment, and 4) e-consultation. The following describes each of the functionalities in more detail.

Figure 1. Example of one of the pages of TISEI environment (www.TISEI.org), with multilingual portals (15 different languages), the forum and information portal. The self-assessment/survey and e-consult options are visible in the language portals.



Information Portal

The purpose of the portal was to provide disaster-specific and relevant information. Numerous websites provide information about the impact of disasters. By making a specific disaster based service, the site targets victims of this particular disaster and keeps them up to date. In addition we selected services that provided health related information.

Forum

The forum facilitates building a community of support for survivors and their loved-ones. This could be through telling personal experiences of the disaster by choosing the web service as an open diary. This general forum allowed for both public and private conversations. Participants had the option to use the 'back channel' after registering, which enabled them to communicate with their registered peers only. The forum was intended for victims, support groups, helpers and others to share their concerns and express their feelings. A webmaster monitored content and language in the forums. Each language had its own forum.

Self-assessment and "Open" Web-based Survey

This functionality facilitates confidential self-assessment and offers a way to understand feelings through a series of scientifically validated questionnaires.

The survey was tailored to assess the victims' mental health and to obtain a reasonably reliable recommendation to seek psychological help when necessary. The web survey provided feedback and allowed the participants to print the questionnaire to keep for themselves or take to their counselor or physician. The online forms could also be used as an e-Consult tool on the site. Furthermore results allow insight in mental health across the entire gruop and insight in development of symptoms over time (cross-sectional and longitudinal research). The survey began with questions about demographics, pre-exposure health, and specifics about the stay in South East Asia. These were followed by a compilation of the validated, existing questionnaires and non-validated questionnaires that could reveal valuable information related to mental health problems, such as sleeping problems. Among the structured questionnaires were assessments of the impact of events [12], trauma [13], peritraumatic dissociation [14], emotions, general health, sleep [15] and of depression [16]. For those who had lost loved ones, a grief questionnaire was also provided [17]. After 6 months, a quality of life scale was added [18] but this was soon removed due to copyright issues.

Prior to starting the self-assessment, participants had to give informed consent and register to receive a password by email to enter the survey. Further details of the design of the survey module are depicted in Multimedia Appendix 1 according to the relevant parts of the CHERRIES checklist [19].

e-Consultation

This is a functionality for online confidential consultation of professional mental health workers. It offers easily accessible advice on mental health issues, including personal advice about the necessity and location of treatment. It enables victims to seek help in response to symptoms of emotional distress either through online consultation or referral to a counsellor in their own region. The service request would be posted anonymously by email, forwarded to a mobile phone of the on call team, and followed up by email. E-Consult use was only available after registration to the web service and after completion of the web

survey. This part could be hosted separately in different portals to accommodate victims with advice relevant to their region of residency.

Server capacity was arranged with external providers. Building and hosting of the actual site was outsourced to a company that builds Internet solutions that have assisted in earlier projects. The site was hosted in the Netherlands and mirrored to warrant data integrity. The framework of the site was language independent, which made it possible to create a local version within a day and make quick multilingual changes.

All personal information was secured with logins verified by registered email and therefore accessible only for the individual participant. Personal information that was entered in the web survey was anonymous through a pseudo anonymisation procedure that meets the criteria of European law. Only the participant and the researchers could retrieve the results of the survey.

The database was accessible for the research team through a password secured login. For the content of the web survey we adapted the questionnaires to Internet research and applied for permission for copyright measures. Professional translators assisted us to develop multilingual portals.

The entire TISEI-environment was language and region independent. Web platforms for online support were made available in 15 languages including: English, French, Spanish, German, Italian, Russian, Dutch, Danish, Finnish, Swedish, Turkish, Czech, Korean, Thai and Indonesian.

Each language portal could have its own user group that could be modified and shaped to a preferred style, with local information and services—as long as the Web surveys remained similar. Each user group could build a database of Web surveys of its own. By building a local team of consultants, each portal could offer local users in different countries the opportunity to provide regional psychological support through the electronic consultation service.

When the setup was finished, the Institutional Review Board of the UMCU assessed and approved the developed website format and study protocol of the survey according to prevailing ethical standards.

Within two weeks the prototype of the site was tested with different Internet providers and computers. We tested ease of use, clarity of instructions, safety, and user friendliness.

Potential site users and participants in the survey were recruited through several means. We attached keywords to relate to search engines. We issued a press release and international news media and organisations such as the International Society for Traumatic Stress Studies (ISTSS), the World Health Organization, World Psychiatric Association and the United Nations were contacted.

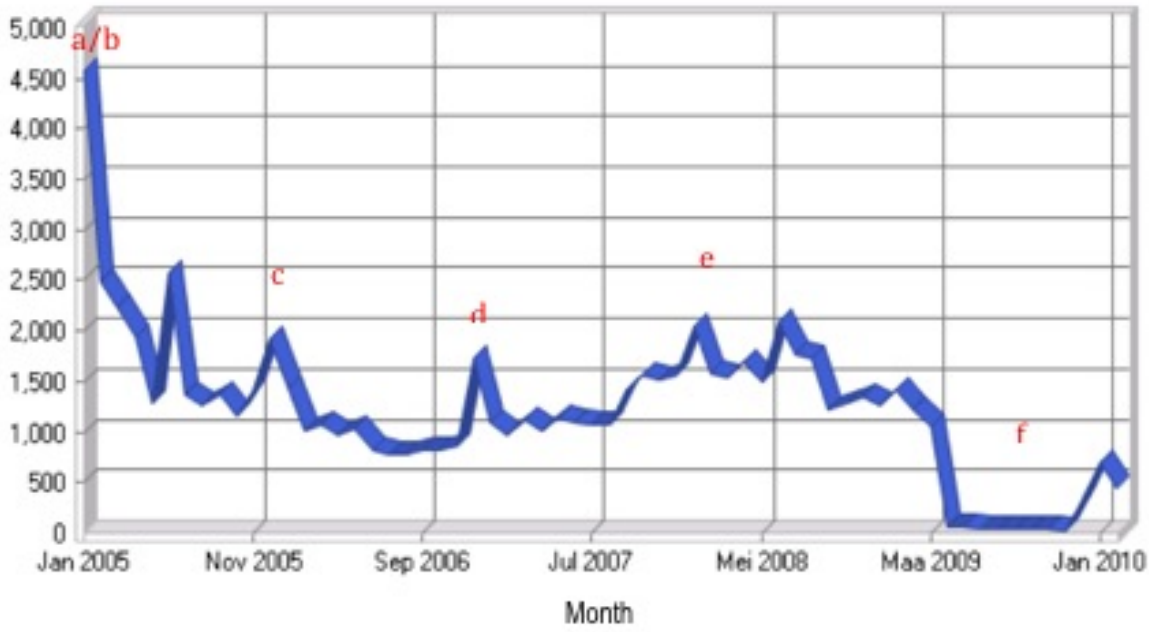
Results

Visitors to the Website

The total number of unique visitors in the first two years was 36, 849 (January 14, 2005 until January 1, 2007). This was based on unique IP addresses entering the site. These numbers peaked in January 2005 with 4.587 visitors within the first 18 days of the website creation. The average number of visitors per month was 1473 varying between 4587/month and 789/month. The first quarter of 2006 – one year after the Tsunami – showed an upward trend in terms of visitors, hits, and bandwidth compared to the eleven months before. This was even more apparent in December 2007 (Figure 2). The European countries and the United States together made up 95% of all visitors to the site. Most of the European visitors were Dutch (Figure 3).

Links on other websites did not result in substantial visitor recruitment; they contributed to less than 2% of the total number of visitors. Most visitors of the TISEI environment directly went to the URL or found the site when searching the keywords “tsunami” (in different languages including English, Dutch, and Turkish), “TISEI”, or names of victims.

Figure 2. Number of visitors per month on the TISEI site.



(a) Tsunami

(b) Launch of website

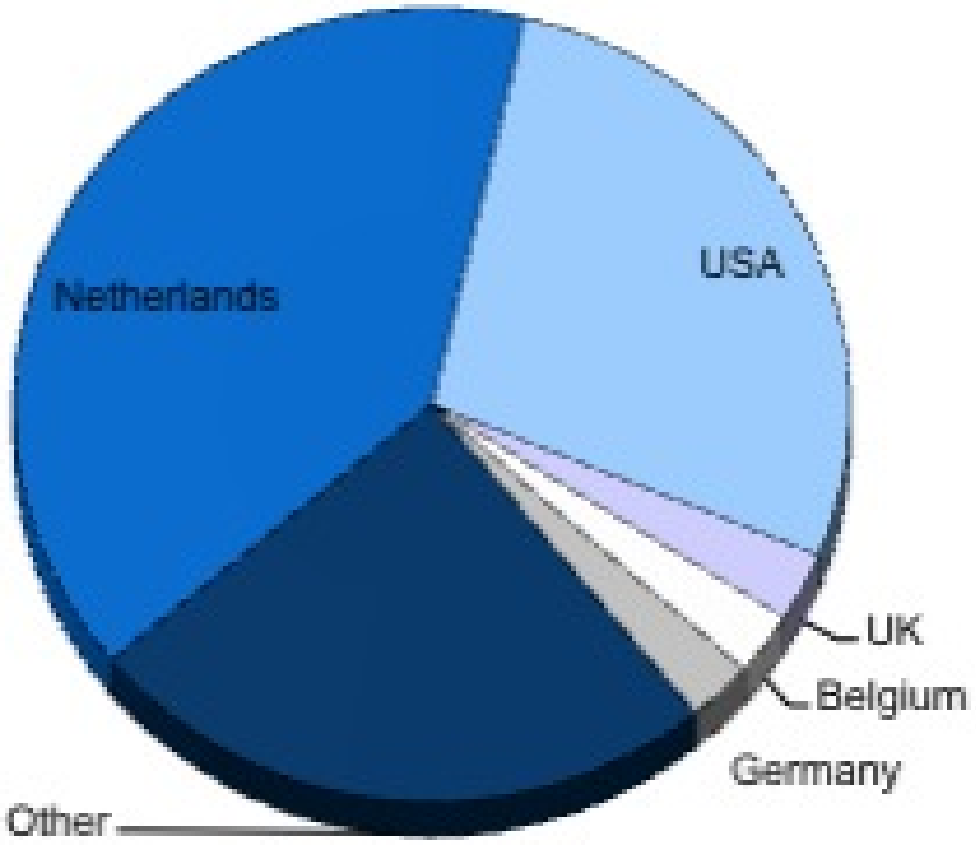
(c) 1 year after the tsunami

(d) 2nd anniversary

(e) Invitation 2nd wave of survey (3 years after tsunami)

(f) Website offline

Figure 3. Activity on the site depicted by country of origin.



Forum

The forum contained many individual stories of victims, as well as responses of people who read the narratives and offered to act as a sounding board or to give explicit help. Examples of these narratives were more extensively described elsewhere [20]. They expressed their symptoms, concerns, complaints, shared evaluation of initiatives they had undertaken for treatment, and provided mutual support to other victims. Experiences and emotions were also expressed by children , illustrated by this contribution on the public part of the site: “I read your story (...) and began to cry. It was so familiar to me! I was also in Thailand during the tsunami, I think it’s so brave of you to talk about it. I would like to talk about it with people our age, I’ve been looking for just that and then I found this Website. It’s good for me to know that I am not the only one with this awful story, I hope we can email (etc) with each other”(www.tisei.org).

It was difficult to come to agreement with international partners to standardize the delivery method for the survey and e-Consult parts of the site. For example Sweden opened a portal but decided to perform a paper and pencil questionnaire instead of participation in the online survey.

For this paper, we only analysed use of the confidential part of the site in our own Dutch portal.

Self Assessment and Survey: the Dutch Sample

The web-based questionnaires were in a password protected area of the site. Informed consent to include survey results in the database measuring emotional impact was obtained from all participants.

The number of Dutch citizens in the Tsunami area at the time of the disaster was estimated to a number of 500, of which 36 died in Asia. Through the Dutch portal, 144 people participated in the TISEI survey and filled in 175 assessments. These people were not directly invited by email but encountered the survey on the site independently.

The response rate was expressed as the number of people completing the questionnaire divided by those who viewed it [21], view rate (ratio of the number of unique site visitors / the number of unique survey visitors) [19], or participation rate (ratio of the number of unique survey page visitors / number of people who agreed to participate) [19] could not be calculated in retrospect due to technical limitations of the site. The completion rate expressed as the number of people completing the questionnaire divided by those who agreed to participate [19] it was 95%.

After three years a secondary measurement was performed in which 39 Dutch survivors participated. However, matching the results of participants over time proved to be difficult due to the method of anonymisation.

e-Consult

As referral or treatment advice is a local or national issue, this part was supported by each country itself with its own “back office” with their own arrangements with actual care providers. In the Dutch sample 31 people used this modality to consult a professional mental health care worker. All questions could be

answered within 24 hrs, except for 3 questions posed in a period of a sick consulent and a non-redirection situation of emails. Besides that, the e-Consult functionality worked well and we did not meet any technical problems with this feature.

After two months a lot of patients expressed PTSD associated symptoms like insomnia, nightmares, irritability, avoidance and flashbacks. Even after one and two years people contacted through e-consult of the TISEI site looking for help for nightmares, anxiety, concentration problems and weight loss related to the Tsunami.

Twenty one persons were referred through the site to a mental health care worker for treatment or further advice. Other people could be helped through online advice and sometimes repetitive contacts. To our surprise, some of the ‘ referrals’ to care providers were done by pears through the forum.

Satisfaction and Feedback of Visitors

Satisfaction about the website was not quantitatively measured on the website but we received numerous emails of users as well as positive reactions on the forum. The patients admitted to the MIH completed an inquiry and were very positive. The site was rated at 8/10. The information portal was valued as the least part. Improvement ideas from users regarded graphical issues and e.g. a counter of messages and more information on e.g. coping with stress and sleeping problems directly on the site in stead of providing links.

Discussion

Opportunities

The TISEI web portal was the first to combine different features of online services used separately in earlier online post disaster initiatives [9]. It fosters community building in conjunction to self assessment questionnaires, e-consultation for referral options, and a research survey. All four different functionalities together can enhance usability and outreach. In this way the portal could serve as a basis for support without interference of health care providers, but with a coupled gateway to professional mental health care and research.

Self-help stimulation through providing adequate, relevant information and a forum serves to empower victims [22,23]. Feelings of self-control can be enhanced though recognition and contact with peers to help to process their own emotions and experiences. Knowing that there are other survivors could reduce feelings of isolation and have a preventive effect on psychological problems [3]. Empowerment after traumatising events is a powerful tool to diminish the severity of mental health complications in the acute phase, which helps to avoid development of chronic complaints. Online communities are growing rapidly and are important for people to exchange experiences and seek advice. The communications on the forum show that the objective of helping victims cope with the impact of a disaster and to help foster resilience was met [20].

Online self-assessment further contributes to this by offering people an easily accessible, flexible, and anonymous tool to discuss, organize, and assess feelings and own mental health. Earlier studies report that self-assessment or a research survey can also be an intervention. These methods provide an opportunity for emotional expression, cognitive processing, restructuring of the experience, and active coping [9]. Through the TISEI site, several participants reported the same experience and expressed their appreciation of interest in the emotional aspects of their response.

Internet applications like the TISEI with a survey module generate research opportunities. Rapid response, tracked standardized answers, cost-effectiveness [24], and global reach are just some examples of its advantages [25]. The identification of individuals at risk for PTSD following a disaster may help organizations prevent both the human and the economic costs of the disease [24]. Several studies showed that sensitive topics, such as psychiatric symptoms are more likely to be reported in self-reported assessments than in interview-based assessments. Web-based data collection has a potential to reduce social bias [26-28]. Studies of responses to disasters have an observational and post event nature. This makes it difficult to identify true causal relationships between exposure and observed outcomes. Reducing selective attrition remains a challenge. As in almost every survey after a disaster, TISEI participants form a volunteer sample of convenience rather than probability sample [29]. Using the Internet to deliver these surveys could diminish such bias, as it also includes people that did not seek professional help in addition to referred patients. Combining research with support groups and self assessment might further improve patient recruitment bias compared to a research site alone [21, 30]. People who are not looking for research or self-assessment may come across the survey while visiting the TISEI site for the forum or other information.

However, these applications also present several logistical as well as scientific challenges. Although surveys like TISEI will not provide causal incidence data, empirical evidence and qualitative research of the impact and adjustment process after disaster exposure will aid clinicians in designing interventions for individuals coping with negative outcome [1]. For an in dept understanding of experiences of particular individuals or groups, qualitative aspects are as important as the quantitative values.

This study shows that at least 29% (144/500) of the Dutch Tsunami victims (31% of 464 survivors) from the Netherlands could be reached for participation in the survey through the TISEI site.

e-Consult

We believe that offering self assessment and/or a survey should be accompanied by e-Consult or phone service for the participants’ safety. Focussing on the details of these topics might exacerbate an already distressed participant [31] and therefore professional help and/or referral options should be arranged for in advance. Patients can then benefit from the information provided through the self-assessment at the same time.

Online consult modalities can offer rapid client centred emotional care. Many clients may not seek help otherwise because of stigma or lack of information. It is an accessible way to pose health related questions to professionals and can assist to seek help in ones own geographic region.

Moreover, the web service was unique in that it was language and region independent. Because the disaster region was a tourist attraction, survivors came from all over the world. To be accessible to this diverse group of survivors it was necessary to translate the TISEI-site into several languages. This differs from many earlier online initiatives such as 9/11 [9, 11] which were made only in English.

Although the TISEI web service was available in 15 languages, 95% of the visitors were of European or US origin. The finding that relatively few survivors from Thailand, Indonesia and Sri Lanka visited the site may be related to a number of factors. First, major global differences in Internet penetration may play a role. Statistics from Internet World Stats [32] reported Internet penetration in the United States in 2006 was 68%. The European average penetration was lower at 36%. Seventy eight percent of Dutch households had Internet access at that time. The areas hit by the tsunami however, had penetration percentages of 9.9 (Asia) and 2.6 (Africa), respectively. Secondly, destruction of the already limited Internet infrastructure by the tsunami will have further lowered Internet availability in Asia and Africa. Thirdly, focus on survival and rebuilding a home and life in this group of survivors might also have been prioritized over personal mental health, in contrast to tourists that returned to their own non-devastated country after the disaster. Given the current and anticipated growth of the Internet, the development of an online multilingual psycho trauma information system holds great promise for the future.

Lessons Learned

Apart from the opportunities of the TISEI site, we encountered a lot of challenges, both organizational and scientific, along the way. There are several lessons to be learned in relation to the topics of penetration, cross country collaboration, funding, setup of the survey, and time pressure.

Penetration

As discussed above, although the site was multilingual and had global potential, most visitors came from Europe and the United States. Apart from the above mentioned reasons, the disappointing global penetration can also be attributed to an inability to establish firm international collaboration and media attention/participant recruitment.

Cross Country Collaborations

Cross country collaborations are mandatory to reach the international population, both for victim recruitment and funding, and for hosting of the portal to accommodate victims with an advice relevant to their region of residency.

For this study, the Swedish, German and Canadian partners came close to collaborating but were stranded on formal bases. It was not clear which organization had ultimate decision power and authority regarding this issue.

Funding

It is no surprise that funding is a crucial but rare prerequisite for innovative acute disaster projects like this. For example in the Netherlands 125 Million EUR was collected on a single promotional night, but this money was dedicated strictly to the rebuilding of the local community. No fundraising was successful for tourist victims or for an international service that had not been proven to be effective yet. Waiting for full and secure funding had interfered with timelines as was also experienced in other projects [9]. Securing quick-response funding, time and people to set up the project was a lesser problem. Long term funding to maintain and manage the website, acquiring data and research on the results proved even more challenging. This interfered with site maintenance, follow up recruitment and result analysis. It was only close to five years later that a grant was awarded to further develop the service for future (Dutch) victims of traumatic events. After three years however, we had difficulties to keep the site online while it could have an important function as a memorial archive, as seen with other sites [33]. These narratives could also be seen as a collective history of the disaster, as these are descriptions of people's unique trauma experiences, which are publicly accessible.

We were somewhat surprised by the increase in visitors in the first trimester of 2006 and 2007 around the anniversary date of the tsunami-disaster. Even after two years the site had over 1000 unique visitors a month. Experience from previous disasters have shown that long term help is essential, particularly during and around the anniversary of an event and needs to be extended past the first year [34-36]. On the TISEI site after 2 years people still used e-Consult. Prospective 5 year follow up of tsunami victims admitted to the MIH also showed that onset of late symptoms can appear after several years. As the e-Consult feature had to be discontinued at that time, patients reported experiencing difficulty finding the right help.

Survey Setup

Due to technical limitations of the initial survey setup, it is difficult to match respondents over time and research possibilities were limited as a consequence. Secure logins were sent to participants by email when they participated in the survey. These could theoretically be used for long-term follow up, but due to the method of anonymisation we used and changes in email addresses, independent entries over time could not be matched to the same patient automatically. Part of them could be matched manually but this was a time consuming effort with incomplete results. For the same reasons, recontraction of survivors was more complicated than expected. Even though multi-timepoint research was possible, true longitudinal follow up was impossible and results could only be analyzed cross-sectionaly. One question list turned out to be an incomplete version and therefore not validated. Not all instruments we wanted to use could be incorporated due to copyright issues.

Unfortunately, the potential for self-selection bias could not be estimated by measuring the response rate, view rate, or participation rate as this was impossible to calculate in retrospect due to technical limitations of the site. Checklists like CHERRIES [19] can not only help to improve papers reporting Web-based surveys but can also help in the starting phase as a checklist for quality of survey set up, together with reports on similar initiatives [9]. At that time we were not aware of the CHERRIES checklist, but participation rates would have been

easy to calculate now if we would have made minor adaptations in the site back then.

Time Pressure

Many difficulties we experienced arose from the time pressure to make the service available immediately after the unexpected disaster. Timeliness is important though, both for support and research. We managed to launch the site within 3 weeks after the tsunami to foster community support in the aftermath of the disaster.

There was little time to develop a template for the site, design a survey study, select and adapt questionnaires, and get approval of Institutional Review Board. A lot of issues such as copyrights, securing patient information and data, and referral systems had to be sorted out on short notice.

As time proved to be a limiting factor in optimizing the design of the website, the study, and the funding, we now started a project to develop a template site. Although disasters are unpredictable in timing and nature, many factors involved in providing a site for online community building, e-help and research will be similar and can be sorted out in advance to maximize possible yield.

In conclusion, the TISEI project, set up as a multilingual website with combined modalities for psychological support and treatment as well as research in the aftermath of the tsunami disaster proved feasible. It could be launched quickly and was operational within 3 weeks after the disaster. It fostered community building combined with self assessment questionnaires, offered in conjunction with e-Consultation. Combining this psychological support with research proved feasible. Self-assessment served as an 'emotional thermometer', and the outcome was fed back to the research participant. All four different functionalities of the site can enhance each other's usability and potential reach. In this way the portal could function as a basis for support primarily without interference of health care providers, but with a coupled gateway to professional mental health care and research. Combination of information and self assessment with offering of treatment or help is mandatory for patient safety. Furthermore, it can help to identify people that need treatment in an early stage.

Web-based services in the aftermath of mass disasters can be an aid in community building and deliver low level, easily available and survivor centred information and support. It has potential for support, care and research with rapid response, cost- and means-effectiveness and global reach.

Time proved one of the most important factors in optimizing design and implementation of the survey in our study and in literature. Patient privacy should be a prime. Yet, for longitudinal analysis the anonymisation procedure should allow for returning participants to be identified. Securing funding and available people to manage the site and its contents proved challenging. Long term funding and maintenance has to be taken into account as even after a few years people look for help and the narratives serve as a collective history memorial.

Despite hurdles and lack of penetration to a global outreach, the growing Internet penetration as well as the rapid expansion and influence of online communities should be an incentive to further optimize care and perform research with the Internet as a platform. The unpredictable nature of disaster puts time pressure on the development of online solutions and influenced the yield of our site. Our lessons of the tsunami web service highlight the necessity of developing methods and international collaborations in advance, secure funding, and expand the potential to other survivors of mass psycho trauma.

Acknowledgments

The initiative for the website was supported by the Major Incident Hospital and was performed through the University Medical Centre Utrecht, the Central Military Hospital and the faculty of Social Sciences if the University of Utrecht. The authors wish to thank Luc Taal for technical assistance in the TISEI project. We thank all other members of the TISEI group; Cobie Groenendijk, Andre Dekker, Cobi Heijnen, Ferdi Unck, and Rolf Kleber for contributing with time and efforts to realize this project. We thank Crossing Channels for resources and setting up the service in collaboration with Microsoft, Hewlett Packard and Intel. We also thank the platform for psychosocial care, Impact Foundation as well as C45, Arq Foundation and Language Institute Babel Utrecht for assistance.

Conflicts of Interest

None declared.

Editorial note: This paper was previously published on the iJMR collection on Google Knol (archived at http://www.webcitation.org/66iTViAmE), last edited Aug 1, 2010, and was openly peer-reviewed.

Multimedia Appendix 1

TISEI; Web-survey of emotional impact; Checklist for Reporting Results of Internet E-Surveys (CHERRIES).

[PDF File (Adobe PDF File), 95KB - ijmr_v1i2e18_app1.pdf]

References

1. Norris FH, Friedman MJ, Watson PJ. 60,000 disaster victims speak: Part II. Summary and implications of the disaster mental health research. Psychiatry 2002;65(3):240-260. [Medline: 12405080]
2. Foa EB, Stein DJ, McFarlane AC. Symptomatology and psychopathology of mental health problems after disaster. J Clin Psychiatry 2006;67 Suppl 2:15-25. [Medline: 16602811]
3. Leon GR. Overview of the psychosocial impact of disasters. Prehosp Disaster Med 2004 Mar;19(1):4-9. [Medline: 15453154]
4. Griffiths KM, Calear AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): Do ISGs reduce depressive symptoms? J Med Internet Res 2009;11(3):e40 [FREE Full text] [doi: 10.2196/jmir.1270] [Medline: 19793719]
5. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
6. Ybarra ML, Eaton WW. Internet-based mental health interventions. Ment Health Serv Res 2005 Jun;7(2):75-87. [Medline: 15974154]
7. Pistrang N, Barker C, Humphreys K. Mutual help groups for mental health problems: a review of effectiveness studies. Am J Community Psychol 2008 Sep;42(1-2):110-121. [doi: 10.1007/s10464-008-9181-0] [Medline: 18679792]
8. Houston TK, Cooper LA, Vu HT, Kahn J, Toser J, Ford DE. Screening the public for depression through the Internet. Psychiatr Serv 2001 Mar;52(3):362-367. [Medline: 11239106]
9. Butler LD, Seagraves DA, Desjardins JC, Azarow J, Hastings A, Garlan RW, et al. How to launch a national Internet-based panel study quickly: lessons from studying how American are coping with the tragedy of September 11, 2001? CNS Spectr 2002 Aug;7(8):597-603. [Medline: 15094696]
10. Butler LD, Blasey CM, Garlan RW, McCaslin SE, Azarow J, Chen XH, et al. Posttraumatic Growth Following the Terrorist Attacks of September 11, 2001: Cognitive, Coping, and Trauma Symptom Predictors in an Internet Convenience Sample. Traumatology. 2005 December 1 2005;11(4):247-267 [FREE Full text]
11. Silver RC, Holman EA, McIntosh DN, Poulin M, Gil-Rivas V. Nationwide longitudinal study of psychological responses to September 11. JAMA 2002 Sep 11;288(10):1235-1244. [Medline: 12215130]
12. Horowitz M, Wilner N, Alvarez W. Impact of Event Scale: a measure of subjective stress. Psychosom Med 1979 May;41(3):209-218 [FREE Full text] [Medline: 472086]
13. Davidson JR, Book SW, Colket JT, Tupler LA, Roth S, David D, et al. Assessment of a new self-rating scale for post-traumatic stress disorder. Psychol Med 1997 Jan;27(1):153-160 [FREE Full text] [Medline: 9122295]
14. Marmar CR, Weiss DS, Metzler TJ. The peritraumatic dissociative experiences questionnaire. In: Wilson JP, Keane TM, editors. Assessing psychological trauma and ptsd. New York: Guilford Press; 1977:412-428.
15. Buysse DJ, Reynolds CF, Monk TH, Berman SR, Kupfer DJ. The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989 May;28(2):193-213 [FREE Full text] [Medline: 2748771]
16. Beck AT, Steer RA, Brown GK. Manual for the Beck Depression Inventory -II. San Antonio: Psychological Corporation; 1996.
17. Boelen PA, Van Den Bout J, De Keijser J, Hoijtink H. Reliability and validity of the Dutch version of the inventory of traumatic grief (ITG). Death Stud 2003 Apr;27(3):227-247. [doi: 10.1080/07481180302889] [Medline: 12703504]
18. EuroQol--a new facility for the measurement of health-related quality of life. The EuroQol Group. Health Policy 1990 Dec;16(3):199-208. [Medline: 10109801]
19. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
20. Vermetten E, Middelkoop, van CJ, Taal L, Carll EK. Online Psychotrauma Intervention in the Aftermath of the Tsunami: A Community-Building Effort. In: Carll EK, editor. Trauma Psychology: Issues in Violence, Disaster, Health,and Illness. Westport, CT: Praeger Publishers; 2007:255-271.
21. Eysenbach G, Wyatt J. Using the Internet for surveys and health research. J Med Internet Res 2002 Nov;4(2):E13 [FREE Full text] [doi: 10.2196/jmir.4.2.e13] [Medline: 12554560]
22. Lemire M, Sicotte C, Paré G. Internet use and the logics of personal empowerment in health. Health Policy 2008 Oct;88(1):130-140. [doi: 10.1016/j.healthpol.2008.03.006] [Medline: 18436330]
23. van Uden-Kraan CF, Drossaert CH, Taal E, Seydel ER, van de Laar MA. Participation in online patient support groups endorses patients' empowerment. Patient Educ Couns 2009 Jan;74(1):61-69 [FREE Full text] [doi: 10.1016/j.pec.2008.07.044] [Medline: 18778909]
24. de Gara F, Gallo WT, Bisson JI, Endrass J, Vetter S. Investment in online self-evaluation tests: A theoretical approach. J Trauma Manag Outcomes 2008;2:3 [FREE Full text] [doi: 10.1186/1752-2897-2-3] [Medline: 18412943]
25. Wyatt JC. When to use web-based surveys. J Am Med Inform Assoc 2000 Aug;7(4):426-429 [FREE Full text] [Medline: 10887170]
26. Krantz JH, Dalal R. Validity of Web-based psychological research. In: Birnbaum MH, editor. Psychological experiments on the Internet. San Diego Academic Press 2000. San Diego: Academic Press; 2000:35-60.

27. Graham AL, Papandonatos GD, Bock BC, Cobb NK, Baskin-Sommers A, Niaura R, et al. Internet- vs. telephone-administered questionnaires in a randomized trial of smoking cessation. Nicotine Tob Res 2006 Dec;8 Suppl 1:S49-S57 [FREE Full text] [Medline: 17491171]
28. Joinson A. Social desirability, anonymity, and Internet-based questionnaires. Behav Res Methods Instrum Comput 1999 Aug;31(3):433-438. [Medline: 10502866]
29. Schlenger WE, Silver RC. Web-based methods in terrorism and disaster research. J Trauma Stress 2006 Apr;19(2):185-193. [doi: 10.1002/jts.20110] [Medline: 16612820]
30. Rhodes SD, Bowie DA, Hergenrather KC. Collecting behavioural data using the world wide web: considerations for researchers. J Epidemiol Community Health 2003 Jan;57(1):68-73 [FREE Full text] [Medline: 12490652]
31. Schlenger WE, Jordan BK, Cadell JM, Ebert L, Faribank JA. Epidemiologic methods for assessing trauma and PTSD. In: Wilson JP, Keane TM, editors. Assessing psychological trauma and PTSD. 2nd ed. New York: Guilford Press; 2004:226-261.
32. Miniwatts Marketing Group. Miniwatts Marketing group. Internet World Stats URL: http://www.internetworldstats.com/stats.htm [accessed 2010-02-21] [WebCite Cache ID 5niX3BJZ9]
33. Virginia Polytechnic Institute, State University. We remember. 2007. Virginia Tech URL: http://www.remembrance.vt.edu/2007/ [accessed 2010-02-21] [WebCite Cache ID 5niVdBaP0]
34. Bryant RA. Recovery after the tsunami: timeline for rehabilitation. J Clin Psychiatry 2006;67 Suppl 2:50-55. [Medline: 16602816]
35. Huizink AC, Slottje P, Witteveen AB, Bijlsma JA, Twisk JW, Smidt N, et al. Long term health complaints following the Amsterdam Air Disaster in police officers and fire-fighters. Occup Environ Med 2006 Oct;63(10):657-662 [FREE Full text] [doi: 10.1136/oem.2005.024687] [Medline: 16644894]
36. Morren M, Dirkzwager AJ, Kessels FJ, Yzermans CJ. The influence of a disaster on the health of rescue workers: a longitudinal study. CMAJ 2007 Apr 24;176(9):1279-1283 [FREE Full text] [doi: 10.1503/cmaj.060626] [Medline: 17452661]

Abbreviations

TISEI: acronym for Tsunami International Survey on Emotional Impact which means ‘wisdom’ in Japanese language
PTSD: Post Traumatic Stress Disorder
ISTSS: International Society of Traumatic Stress Studies
UMCU: University Medical Centre Utrecht
MIH: Major Incident Hospital
IRB: Institutional Review Board
CHERRIES: Checklist for Reporting Results of Internet E-Surveys

Edited by G Eysenbach; submitted 19.06.10; peer-reviewed by J Haagen, S R, D Rutgers; comments to author 15.08.12; revised version received 15.09.12; accepted 21.09.12; published 21.11.12.

Please cite as:
Marres GMH, Leenen LPH, van der Slikke JW, Vermetten E
Use of a Web Portal for Support and Research After a Disaster: Opportunities and Lessons Learned
Interact J Med Res 2012;1(2):e18
URL: http://www.i-jmr.org/2012/2/e18/
doi:10.2196/ijmr.1588
PMID:23612349

Original Paper

Issues Regarding the Implementation of eHealth: Preparing for Future Influenza Pandemics

Junhua Li[1], Mphil; Holly Seale[2], PhD; Pradeep Ray[1,3], PhD; William Rawlinson[4,5], PhD; Lundy Lewis[6], PhD; C. Raina MacIntyre[2,7], PhD

[1]Asia-Pacific Ubiquitous Healthcare Research Centre, The University of New South Wales, Sydney, Australia

[2]School of Public Health and Community Medicine, Faculty of Medicine, University of New South Wales, Sydney, Australia

[3]School of Information Systems, Technology and Management, Australian School of Business, University of New South Wales, Sydney, Australia

[4]Virology Division, Department of Microbiology, South Eastern Sydney and Illawarra Health Service, Sydney, Australia

[5]School of Medical Sciences/School of Biotechnology and Biomolecular Sciences, Faculty of Medicine, University of New South Wales, Sydney, Australia

[6]Information Technology Department, Southern New Hampshire University, Manchester, NH, United States

[7]National Centre for Immunization Research and Surveillance of Vaccine Preventable Diseases (NCIRS), Sydney, Australia

Corresponding Author:
Junhua Li, Mphil
Asia-Pacific Ubiquitous Healthcare Research Centre
The University of New South Wales
APuHC, Room 1039, 1st floor, west wing, Quadrangle Building, kensington Campus, University of New South Wales
Sydney,
Australia
Phone: 61 (2) 9931 9308
Fax: 61 (2) 9662 4061
Email: junhua.li.syd@gmail.com

Abstract

Background: eHealth is a tool that may be used to facilitate responses to influenza pandemics. Prior to implementation of eHealth in the hospital setting, assessment of the organizational preparedness is an important step in the planning process. Including this step may increase the chance of implementation success.

Objective: To identify the preparedness issues in relation to implementation of eHealth for future influenza pandemics.

Methods: One hospital was selected in Australia for this study. We conducted 12 individual interviews to gather a rich data set in relation to eHealth preparedness in the context of the 2009 influenza A (H1N1) pandemic at this major teaching hospital. These participants' views were analyzed according to five main themes: (1) challenges in present practices or circumstances for pandemic responses, which indicates a need for change, (2) healthcare providers' exposure to eHealth, (3) organizational technological capacity to support an IT innovation for medical practices, (4) resource preparedness, and (5) socio-cultural issues in association with eHealth implementation in response to a pandemic.

Results: This article reports a subset of the issues identified during the case study. These issues include, for example, poor sharing of patient health records, poor protection of patient privacy, clinicians' concerns about IT reliability and dissatisfaction with the software in use, clinicians' concerns about IT's impact on professional autonomy versus having inefficient IT support, and inefficient communication across departments in the form of consultation.

Conclusions: Based on discussions with the participants and interpretation of their responses, we assessed the hospital's preparedness status and also identified areas of deficiency. Accordingly, we suggest possible solutions for the areas in need of improvement to facilitate eHealth implementation's success. The study results will also provide policymakers at national, state and local levels with insights to refine relevant public health policies for the planning and management of pandemics from the eHealth perspective.

(Interact J Med Res 2012;1(2):e20) doi:10.2196/ijmr.2357

KEYWORDS

eHealth; influenza pandemic; preparedness assessment; case study

Introduction

An influenza pandemic increases morbidity and mortality across the population, threatening critical infrastructure by removing essential personnel from the workplace for weeks or months [1,2]. The 2009 influenza A (H1N1) pandemic resulted in millions of laboratory confirmed cases and over 18,000 deaths in 200 countries [3]. The pandemic strain H1N1 had similar infectivity as seasonal influenza strains in circulation in previous years. An immense burden was still placed upon health care services [4].

eHealth refers to the application of information and communication technologies (ICT) across the whole range of functions that affect health [5]. This is an emerging field at the intersection of medical informatics, public health, and business [6]. In a pandemic situation, eHealth may facilitate the pandemic response by enhancing surveillance and control activities (eg, rapid case reporting), and by facilitating the exchange of information (eg, efficient documentation and sharing of patient records) [7-10]. However, information system researchers have recognized the problems of sustainability and complexity in eHealth implementations [11,12].

The assessment of organizational preparedness for an innovation can reduce the risk of failure after introduction [13]. Preparedness in the health care context is defined as the degree to which organizations are ready for the implementation of new ICT [13,14]. If motivational forces such as health care providers' dissatisfaction with *status quo* were not present, it would be unlikely that the innovation process would be initiated. Even though adequate motivation was present, sufficient resources would be required to allow and support steps for change. Furthermore, organizational preparedness for change is the strongest predictor of employee commitment to the organization [15]. If staff members do not possess attributes necessary for change (eg, adaptability and growth-orientation) or resist change, the change process is less likely to proceed [16]. A lack of information about health care organization preparedness for new ICT increases uncertainty for decision makers, decreases their ability to make effective decisions that would mitigate ICT innovation risks, and increases the risk of failure at critical times during a pandemic [17].

The Australian Center for Health Research Limited recognized the influenza pandemic as a threat to the hospital system, but there was no data internationally to inform the business continuity and resilience of the hospital sector. With demands for research in this area, a new collaborative project titled *Pandemic influenza, human resources and critical infrastructure dependencies: mitigating the impact on hospitals* was launched in 2009. This project brought together risk analysis, business continuity planning, and complex systems modeling methodologies based on eHealth to predict and mitigate the impact of a pandemic on the function of hospitals. As part of the project outputs, this article reports results from a case study at a major teaching hospital in New South Wales (NSW) Australia, which aimed to assess the organizational preparedness status regarding the implementation of eHealth for future influenza pandemics. The name of the hospital is not mentioned to maintain confidentiality.

Methods

As a research strategy, case studies are used in many situations to contribute to our knowledge of individual, group, organizational, social, political, and related phenomena-it allows investigators to retain the holistic and meaningful characteristics of real-life events [18]. Case studies have a distinctive place in evaluation research [19-22]. One application is to illustrate certain topics within an evaluation in a descriptive manner [18]. In this study, one case was deliberately selected to evaluate eHealth preparedness, following a single-case design [18].

A qualitative research approach was utilized to provide a rich data set in relation to eHealth preparedness assessment, drawing on practical experiences of individuals involved in the 2009 influenza A (H1N1) pandemic response. The Medical and Community Health Research Ethics Advisory Panel, University of New South Wales approved the study protocol (Reference Number: 2011-7-10).

Interview Guide

An interview guide was developed following a review of the literature and incorporated aspects of an integrated eHealth preparedness framework [23]. The guide provided an initial point of departure for the interview-based data collection process and examined the following areas: (1) motivational forces for change that reflect the problems identified by the evaluator and health care providers' dissatisfaction with present practices or circumstances for pandemic responses-pandemic responses require surveillance, control, and performance of medical practices, (2) health care providers' exposure to potential eHealth applications (engagement preparedness), including their perceived benefits and uncertainties of eHealth for a pandemic response, (3) technological preparedness as a reflection of the capacity to support an ICT innovation, (4) resource preparedness including decision makers' knowledge of ICT implementation, supportive policies, and sufficient funding, and (5) societal preparedness in association with eHealth implementation in response to a pandemic. Communication links and partnerships need to be available within and across the organization. Questions from the interview guide (Multimedia Appendix 1) were generated to evaluate preparedness measures at the bottom level of the 5-dimension hierarchical framework, previously described by Li et al [23]. To examine the motivational forces for change, we asked the question, "were there any problems with the performance of medical practices during the influenza A H1N1 pandemic?"

Site Selection Criteria

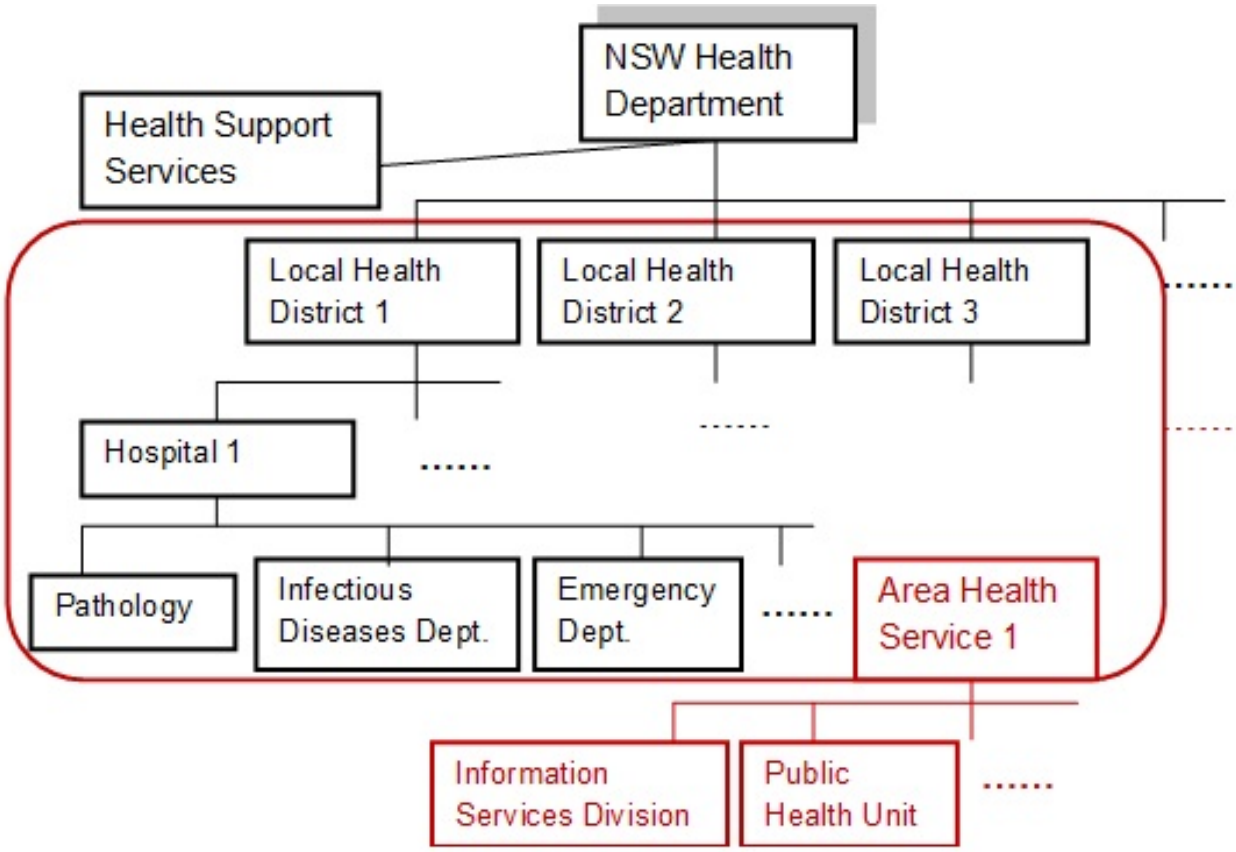
A number of reports and publications from the New South Wales (NSW) Ministry of Health were reviewed to understand the state health system to ensure that a representative site was selected for this case study (Figure 1). The site selected for this study followed these criteria: (1) the hospital is public and affiliated to the NSW Ministry of Health, (2) the hospital must have a large number of admitted acute patients and patients treated in the emergency department each year, (3) the hospital

must have been involved in the 2009 influenza A H1N1 pandemic response, and (4) the hospital must be planning to implement a new or upgrade an existing eHealth system.

Purposive sampling was employed since the interviews required participants' knowledge of *status quo* at the hospital to reveal its motivational, engagement, technological, resource, and societal preparedness for the prospective eHealth system implementation. Due to the nature of the data collected, three groups of participants were involved: (1) clinicians who had experience in diagnosing and reporting cases of influenza A H1N1 and who would be an end-user of the eHealth system, (2) an information technologies (IT) manager who provided IT support services during the H1N1 pandemic and who was familiar with the ICT infrastructure at the hospital (eg, the information systems in use), and (3) the chief information officer (CIO)/person who was in charge of the planning and implementation/upgrade of the eHealth system. We have also set the inclusion criteria that participants must have worked at the hospital for a minimum of 1 year, and were full-time or part-time staff (contract workers were not included).

Three interviews were piloted with a representative from each participant group of interest. The purpose was to evaluate the interview guide, for its readability, relevance, and difficulty of interpreting and answering the questions asked. The guide was modified accordingly.

Figure 1. NSW Ministry of Health structure in 2011.



Data Collection

A hospital located in Sydney, NSW was selected for the case study. It is a major teaching hospital with almost 3000 staff and has 440 beds–an average occupancy rate (number of beds occupied) of well over 90%, ensuring that the hospital has a relatively small but highly complex caseload. Each year more than 30,000 acute patients are admitted and about 40,000 patients are treated in the emergency department. The hospital also attends to around 900,000 non-admitted patient occasions of service each year through community health, outpatients, and rural outreach services.

Potential participants from the hosptial site were nominated and contacted by email. The information services division (ISD) was a seperate entity which provided the hospital with IT-related services. Through snowball sampling, 12 eligible participants were gradually recruited from April to August 2011. The recruitment process ended once enough detailed insights were provided to reach a point of saturation with respect to the preparedness areas. Of the 12, 10 clinicians (6 medical doctors, 3 nurse managers, and 1 microbiologist) were from the infectious diseases department, emergency department, hospital epidemiology center, department of microbiology at the hospital, while the remaining 2 (the chief information officer and an IT manager) were affiliated to the ISD.

Analysis

With the participants' agreement and consent, all the semi-structured interviews were conducted by 1 of the investigators and recorded with a digital voice recorder. After the interviews were transcribed verbatim, the analytic process was conducted manually. A quarter of the transcripts were

randomly selected and coded independently by two investigators. Subsequent discussions between them developed a list of themes. An agreed framework was then applied by 1 of the 2 investigators to code the remainder of the transcripts and the themes were further modified. Based on the themes finally identified, all of the transcripts were analyzed. The analysis results were then discussed with the other authors. Lastly, modifications were made according to those comments and feedback.

Results

The NSW Ministry of Health divided the state into eight separate regions, which were called Area Health Services (AHS). Across the AHS where the hospital was located, there were more than 300 information systems in use for various purposes including care delivery, logistics, and finance. One widely used system within the AHS was the electronic medical records (EMR). The plan for the EMR was commenced in February 2007 and the Cerner EMR suite was selected. The key features of the system included scheduling of patients for operating theatres and retrieval of radiology and pathology results. At the hospital selected for the case study, EMR implementation occurred from October 2009 to June 2011. Until July 2011, the EMR had been utilized in 955 clinics across 16 hospitals for clinical practices. According to respondents from the ISD, there was still the challenge of sharing patient health information across hospitals even at the state level due to the absence of a unique patient identifier.

> *You may go to two different hospitals, having different identifiers, so EMRs may be created for you in the Hospital A; separately you may go into Hospital B when you've gone back home and Hospital B may say we don't have access to that information at this point in time … you've got all this information being collected in different hospitals. What you need is a unique identifier, which sits at the state level. Once that becomes available, it becomes easier.*

Respondents from the ISD indicated that IT development had never stopped and the ISD was continuously looking for opportunities to facilitate better patient care outcomes. They pointed out that it was planned to implement a new IT system, which was referred to as the Antimicrobial Stewardship System (ASS). The project team had involved health professionals from the infectious disease department of a few hospitals, as well as pharmacists. According to the interviewed clinicians, the system would monitor antibiotic use, cut down patient care cost by reducing inappropriate antibiotic prescriptions, and also improve patient treatment outcomes.

One clinician indicated that there was a list of restricted antibiotics which clinicians had to apply to use. The system would provide clinicians with an automatic authorization to prescribe these antibiotics once they gave the right indication. There would be a way of tracking that information if somebody tried to "game" the system. A few interviewed clinicians pointed out that inappropriate antibiotic prescriptions could occur, for example, by unnecessary use of antibiotics for a patient's condition or expensive antibiotics, and prescribing a larger amount, for a longer duration or a wider spectrum of antibiotic use than was required. A couple of clinicians perceived and suggested that the ASS could improve antimicrobial prescriptions and specificity in a timely fashion for only the duration that was required.

The following sections report on a subset of the case study's results, which identified issues related to the hospital's preparedness for the implementation of eHealth as well as response to influenza pandemics.

Motivational Preparedness

Capturing Alerts Issued by Public Health Units

Health alerts during the 2009 pandemic were issued by both the state Department of Health and the Local Health District. Clinicians received these alerts through emails, facsimile, the hospital intranet website, and the state bulletins. The alerts contained information on H1N1 updates (eg, the number of cases, updated case definitions, current and accepted best practice). Clinicians pointed out that there was a lack of reliable information at the beginning of the pandemic, followed by an unmanageable number of emails with excessive, unfiltered, and repetitive information on H1N1 at the later stages of the pandemic. Clinicians suggested that summarized updated information placed at workstations would have been more effective than multiple alerts sent on the same topic to each clinician.

Documentation Efficiency

When commenting on the statement that the retrieval, update, and storage of patient health records were inefficient during the 2009 pandemic, the majority of the interviewed clinicians agreed at varying degrees. Some clinicians highlighted that EMR were implemented at the hospital towards the end of the pandemic in 2009, and that a lot of information (eg, previous clinical assessments, medication history, and previous prescriptions) required for clinical practices was not available in the EMR. Consequently, clinicians still had to retrieve information from the understaffed medical records department where paper medical records were stored. This was a time consuming process and not ideal especially during a pandemic.

Completeness and Accuracy of Patient Health Records

A couple of interviewed clinicians deemed that complete and accurate patient medical history was not important for clinical decision making in the circumstances of the 2009 pandemic. During the pandemic, the majority of the patients affected were younger and therefore had few medical complications. As such, a complete history of the patient was often unnecessary, and that the appropriate medical treatment could be given based on their clinical presentations at the time. One of the others argued against that, indicating that information on "whether the patient was pregnant or got diabetes or other kinds of diseases" was important for clinical decision making in the flu setting.

Most clinicians reported inaccurate and incomplete patient medical records even with EMR in use. Some explained that the patient's medical history was generated by various health professionals at multiple locations such as hospitals and GPs and the accuracy and completeness was dependent on "how

well the clinical record was made originally" and on how all that information could be shared from different locations. They needed to have "a universal aligned EMR travel between hospital and GPs". When asked what information was available in the EMR at the hospital, the interviewed clinicians mentioned emergency assessment documentation, operating theatre reports, pathology and radiology results, and some discharge information, but pointed out that medical officers' assessments (eg, past diagnoses, medication and other clinical notes) were excluded.

Patient Privacy and Information Security

Access to eHealth applications such as Laboratory Information System (LIS) and EMR, required a username and password. Nevertheless, clinicians across departments could look up and access anybody's medical record at any time, leaving patient information uncontrolled. Common access to patient records between clinicians with different designations could breach the patients' confidentiality. One clinician argued against that electronic trails of system access and data operations that had been in use – what a user retrieved from the systems was recorded with time stamps within the workstation or in a repository, and assessment could then be subsequently conducted on those records. A senior clinician from the infectious disease department suggested that "the rules around confidentiality" should be better specified and that the EMR usage should conform to those rules.

Correctness of Diagnoses

Although the majority thought that diagnosis could have been incorrectly given due to lack of patients' full history, a small number of the interviewed clinicians disagreed. They argued that the clinical diagnosis could provisionally be made even before confirmation from the laboratory test, if patients presented the signs and clinical symptoms of the H1N1 case definition (eg, high fever, cough, and sore throat).

Various reasons were given for what could cause incorrect diagnoses, including poor clinical history taken, language barriers resulting in misinformation between patients and doctors for clinical decision making, mislabelled samples, delays on processing and testing of lab specimens, false positive or false negative laboratory test results, data entry errors into the LIS, and inexperienced practice or irresponsible clinicians.

> *People were not diagnosed with the pandemic flu who probably had the flu, but on the basis of negative lab test, were not considered to have the disease. That was poor understanding of the performance of the test result.*
>
> *If a rapid test comes back negative, some staff initially would go 'no, they don't need isolation'. They obviously had an influenza-like illness, people go 'but that test is negative', don't take into consideration what's happening.*

Appropriateness of Prescriptions

Few clinicians interviewed believed that no errors in prescriptions took place at the hospital. They explained that access to Tamiflu for influenza A (H1N1) had to be approved by the infectious disease department and also that a very standard treatment dose was specified in the case definition. Others argued that "there is always an error margin" around prescriptions.

A number of reasons could have led to errors. During times of a pandemic, clinicians were often bombarded with an overwhelming amount of questions simultaneously, which could increase work pressure and distractions, leading to forgotten medication orders or transcribing errors. There could also be issues with inadequate knowledge of medication, which could result in prescribing the wrong drugs for the wrong diagnosis, incorrect doses, and inappropriately assigned treatment duration. Due to the high level of pressure during these times, clinicians can prescribe a drug without careful consideration of contraindications (eg, patient's allergy history). Finally, there could also be dispensing errors at the pharmacy due to overwhelming volumes of prescriptions, illegible prescriptions, and a lack of time and resources to check for these errors, which they would normally have done.

Engagement Preparedness

Clinicians' Concern About Reliability of IT

The majority of the interviewed clinicians disagreed with the statement that information technology is always reliable, indicating that technology glitches and downtime had resulted in interruption and inconvenience in their clinical practice. Paper trail records were supported in addition to IT measures in case of failure of the technology in use.

Apart from IT itself, information available through the technology could also be unreliable, as some clinicians added. They explained that the reliability was dependent on the information source, which could be outdated, or there could be typing errors.

Clinicians' Concern About IT's Impact on Professional Autonomy

The majority disagreed with the statement that their professional autonomy in health care systems was not their concern after the ASS was introduced to health care practice. One explained that the EMR provided a drop-down list of diagnoses when patients were discharged. If the diagnosis was not in the list, clinicians were forced to select an incorrect choice from the list instead of being allowed to type in the actual diagnosis. These imposed operations caused the loss of their professional autonomy. A couple of clinicians added that, with pervasive information technology at the health care facility, clinicians' practice became dependent on IT professionals' support and frequent interactions occurred between these professional groups. "That's culture change", commented by one of the others who agreed with the statement. When referring to the ASS, 1 clinician from the emergency department argued that medical doctors needed to have the freedom to prescribe what they thought was necessary in the event of an emergency–"anything that will restrict us would be opposed quite strongly". A few argued further that with a guideline or protocol, the proven standard indication might not be appropriate for a patient's specific condition.

Some others agreed that the implementation of the ASS could challenge medical doctors' autonomy. However, they argued that better patient care outcomes and patient safety should be the primary concern of professional autonomy without required relevant clinical knowledge. .

Technological Preparedness

Clinicians' Dissatisfaction with the Software in Use

According to the IT manager, about 190 IT applications were in use within the AHS. At the hospital, available clinical and non-clinical applications included, the EMR, LIS, picture archiving and communication system, community health information management system for health workers visiting patients at home, human resource management system, car booking system, Oracle financials (solutions to a wide range of long- and short-term accounting system issues), and payroll management system.

Most interviewed clinicians were dissatisfied with the hospital IT systems and pointed out problems such as: (1) integration issues–although pathology and radiology results had been integrated with the EMR, external paper medical histories still existed and had not been integrated, therefore, clinicians had to look at the EMR and also check paper records (eg, "drug orders in the clinical note") for clinical decision making, (2) poor response time–wait times during loading of EMR, log in, redundant pop-up questions confirming identity of user, accessing internal links, log out, and shut down of the system, (3) unfriendly user interface–many felt that the interface of the EMR system was not intuitive, and (4) inconvenient secondary use of available clinical data–although clinicians could efficiently share laboratory test results on a single patient basis through the LIS, it was difficult to extract and collect these data on a population basis to do overall audits from the infectious diseases' perspective as the LIS did not have that capability.

A few clinicians indicated that due to some of these limitations of the IT systems, the clinician-patient relationship could be interfered. Clinicians now had to spend more time on the computer to handle these systems rather than in face-to-face contact and conversation with the patient.

> *You walked in the department before, you would see a few people at the computers, a lot of people with patients. Now, it's the other way around: a few people with patients, a lot of people at the computers … imagine a patient is screaming in pain and wants your attention.*

Inefficient IT Support Perceived by Clinicians

The procedure in the event of an IT systems failure is to contact the Statewide Service Desk (SWSD) and log a call (ie, ticket). The SWSD is a centralised service desk for health facilities across NSW and is run by the Health Support Services. Based on the information provided by the caller at the first point of contact, the SWSD operator makes a brief analysis and electronically allocates a ticket through the SWSD system to 1 of 6 groups at the ISD: (1) project planning group who provide IT project consultation services and manage procurement, (2) communication group who ensure that the computer network is working, (3) application support group who look after the applications provided by the AHS, specifically dealing with interface problems between EMRs and the patient administration system, (4) desktop support group who installs the required software in person, (5) technical services group who takes care of hardware and the data center such as data backup and email accounts, and (6) client support service group who provide IT support such as troubleshooting to the users.

Subsequently, the group, on behalf of the health facility to which the caller is affiliated, manages the ticket. The IT manager indicated that the client support service group (27 people) alone managed over 2,000 calls a month. He commented that the ticket allocation could sometimes be inappropriate due to the misdiagnosis of problems by the operator. He explained that email access failure, for example, could be caused by a faulty network card or a dysfunctional port; the former should be taken care of by the technical services group whereas the latter should be by the communication group. If the operator did not ask the caller appropriate questions, the ticket could be allocated to the wrong group and consequently it would take longer than it should to solve the problem.

When asked whether IT support for troubleshooting was efficient, half of the interviewed clinicians agreed ("absolutely fantastic") while the other half gave completely opposite views ("absolutely pathetic, terrible"). Most the clinicians who disagreed argued that there was inefficient communication with the SWSD. A few explained that it took minutes before they could even talk to someone and had to enter a lot of information before they could proceed. Some further explained that the line could be busy and they needed to log a call, waiting for SWSD operators to ring them back. If they missed the call back from the SWSD, they had to call again and re-log the call. Once the call was allocated to the right person, the problem could be solved efficiently by either remote or on-site support.

Societal Preparedness

The CIO indicated that there was always cooperation between departments of interest. The ISD had been working closely with clinicians from different departments and involving them through the implementation process. When asked to comment on the statement that communication across the department was efficient in the form of consultation, the minority of clinicians indicated their agreement. Some argued that there could be delays in the consultation process, for example, due to delays in sending out requests–clinicians with insufficient medical specialty knowledge might not realize the need for a consultation with specialist staff early on, and as a result the request would be made later than it should. Delays in response to requests are also possible; if the request was sent through the paging system, for example, and the recipient's page was inaccessible at that point of time, a delayed response would occur.

Some pointed out that communication efficiency was also dependent on the professional relationship of clinicians across departments who are involved in the consultation and involvement of key people who have a cross-department role for cooperation.

When you need to actually draw departments together (eg, consultation), it's better to speak to key people. You might send out a group email; it's important that you have the right person signing it; otherwise it will take no notice. Depends on the author (smile). So that's really important.

Discussion

Based on discussions with the participants and interpretation of their responses, we have identified the areas of deficiency in the hospital's preparedness for the implementation of eHealth (eg, the ASS) and response to future influenza pandemics. Table 1 summarizes these areas of deficiency with possible solutions.

Table 1. Identified areas of deficiency and suggestions.

Areas of deficiency	Suggestions
Timeliness of issuing and capturing health alerts	(a) Electronic case reporting rather than by telephone or in writing.(b) Exploring the modality of alerts being issued (eg, a hotlink on the desktop of clinicians' workstations).
Documentation efficiency	Scanning, indexing, and integrating external paper-based documents into the EMR.
Sharing of patient health records and protection of patient privacy	(a) Applying a unique patient identifier to facilitate collaborative health care delivery across facilities.(b) Defining what information needed to be shared with whom and in which way.(c) Adopting the role-based access control (RBAC).
Correctness of diagnoses	(a) In the circumstances of a pandemic: providing clinicians with updated case definitions with check-box criteria.(b) For medical practice in general: using a set of logical if-then rules extracted from medical guidelines.
Appropriateness of prescriptions	(a) Using the ASS being implemented.(b) Exploring other options such as automatic check for contraindication with complete and updated patient information.
Clinicians' concerns about IT reliability and dissatisfaction with the software in use	Requiring a more strategic methodology for its design and service management, such as an Eight-Stage Service Design and Management Model and House of Quality Matrix.
Clinicians' concerns about IT's impact on autonomy versus having inefficient IT support	(a) Requiring more operators the SWSD for a particular time period (eg, at the early implementation phase of a new system).(b) Providing SWSD operators with more IT-related training and education to correctly diagnose and allocate technical problems to ISD groups.
Inefficient communication across departments in the form of consultation	Involving key people for cross-department cooperation.

Regarding motivational preparedness, identification of the challenges within present practices for pandemic responses indicates a need for change. Perceived needs by health care providers impact on their behavioral intention to adopt and use an eHealth system [24,25]. In response to the issues raised during the interviews, some of the broad requirements for IT development in order to improve response to a future pandemic are outlined below.

Timeliness of Capturing Alerts Issued by Public Health Units

To issue reliable health alerts in a timely fashion, public health units initially needed to collect case information from reporting sources (eg, clinicians). If alerts were issued by email, clinicians would not be able to access them in real time due to overloaded clinical practices on the floor. To improve the timeliness of capturing alerts, case reporting should be made electronic rather than by telephone or in writing–case notification could be made immediately after clinicians' diagnosis. Furthermore, there is a need to explore the modality of alerts being issued (eg, a hotlink on the desktop of clinicians' workstations or SMS messages, as suggested by the interviewed clinicians).

Documentation Efficiency

Although EMRs were available, external paper-based documents (eg, patient medical history stored at the medical records department before the EMR implementation) were still required for current clinical decision making. Clinicians pointed out that retrieval of these documents was inefficient. The paper documents should be scanned, indexed, and integrated into the EMR.

Sharing of Patient Health Records and Protection of Patient Privacy

The interviewed clinicians indicated that it was difficult to share patient health records particularly across the area health services and between states. A unique patient identifier (ie, National Health Identifier, NHI) should be applied to facilitate collaborative health care delivery within and across service settings in the country. A variety of clinicians from multiple service settings should be assured of access to patient medical history when required, but with the RBAC utilized to protect patient privacy and information security. Further exploration was required to define what information needed to be shared with whom and in which way.

Correctness of Diagnoses

The interviewed clinicians explained the reasons why incorrect diagnoses happened and named a few (eg, false positive and false negative laboratory test results, inexperienced practice or irresponsible clinicians). Clinicians suggested that the ICT application could reduce diagnostic errors by providing updated case definitions to clinicians with check-box criteria in the circumstances of a pandemic. For the medical practice in general, ICT options should be explored, for example, using a set of logical if-then rules extracted from medical guidelines.

Appropriateness of Prescriptions

The reasons varied for prescription errors, as clinicians explained (eg, inadequate knowledge of medication and absence of the consideration of contraindications). The ASS is an example of how ICT can be applied to reduce prescription errors, but it will only be used for antibiotics. Options to decrease prescription errors need to be further explored (eg, automatic check for contraindication with complete and updated patient information).

Other Areas of Deficiency

Deficiencies in the hospital's preparedness were also identified under other three main themes and needs to be addressed. Many clinicians correlated their doubt about IT reliability with their frustration from or dissatisfaction with the IT systems in use (eg, poor response time and unfriendly user interface). Negative IT experience can cause them technology phobia and thus inhibit their adoption intention of a new eHealth system [26,27]. Any IT system in the future will require a more strategic methodology for its design and service management, such as an Eight-Stage Service Design and Management Model [28] and House of Quality Matrix [29].

Some clinicians perceived and indicated that due to the increasing penetration of information technology into health care settings, clinical practices have become more and more dependent on IT support. Nevertheless, the support was perceived as inefficient as a result of inefficient communication with the SWSD. Efficient technical support particularly for troubleshooting takes a predominant role in smoothing clinicians' re-engineered job routine and overcoming their technology phobia, and consequently can facilitate their acceptance and use of a new eHealth system [30,31]. Clinicians indicated that if there was no operator available, which was often the case, they had to be on hold for variable lengths of time to report their problem. As a possible solution, more operators should be put on duty at a particular time period at the early implementation phase of a new system. The IT manager also pointed out that due to insufficient IT knowledge among some SWSD operators, misdiagnosis of problems took place, and consequently problems were misallocated to ISD groups. He remarked that the misallocation resulted in a decrease of IT support efficiency. To address this issue, more IT-related training and education should be provided to these operators.

Clinicians reported that there were delays both in sending out consultation requests due to senders' insufficient medical specialty knowledge and in responding to requests due to some other facts in relation to the recipients (eg, performing an operation). In the context of a pandemic response or eHealth implementation, cooperation, and communication is often required between medical departments and the IT team to share ideas, address concerns, alleviate fears and mediate tensions amongst involved clinicians and IT staff [32]. A senior nurse from the hospital epidemiology center suggested that it was necessary to involve key people at least for cross-department cooperation, which could facilitate two-way communication.

Limitations

This article examined organizational and health care providers' preparedness at a hospital in NSW for the implementation of the ASS in the context of the 2009 pandemic. The results of this case study may be limited due to participants' over-reporting or their recall bias. The three groups of participants may have over-reported their preparedness in order to avoid embarrassment or judgement. We attempted to minimize any bias in the interpretation of the interview data by having it reviewed by two investigators.

Contributions

eHealth preparedness assessment helps the decision maker at a health care organization to be well-informed of deficient areas in preparedness, and therefore serve as a guide for preventive action to combat the failure to innovate [13,14]. A few studies [33,34] have been found in the literature on the development of a framework for eHealth preparedness assessment. These frameworks were developed from different perspectives. Most studied components reflected health care providers' and organizational perspectives, but these components were different from one framework to another [23]. By integrating these components, a 5-dimension framework [23] provided a guideline for eHealth preparedness assessment in the context of a pandemic. This integrated framework has not yet been applied in real health care settings. Also, there is no study internationally on the evaluation of eHealth preparedness in an organizational context. Regarding theoretical contributions, our study has demonstrated the applicability of the integrated framework in a real health care setting and also provides the medical informatics audience with an example of how eHealth preparedness assessment can be conducted in an organizational context. We believe that these theoretical contributions will prompt further investigation among practitioners and academicians on organizational preparedness for the implementation of e Health systems.

In practice, our findings and discussions may assist decision makers in the organizations to take action to address deficient areas in their preparedness and, as a result, facilitate the eHealth implementation success. Pandemic preparedness planning is necessitated during the inter-pandemic period to enable countries to be prepared to recognize and manage an influenza pandemic [35]. These reported findings may also provide policymakers at national, state, and local levels with empirical evidence and insights in order to refine relevant public health policies for the planning and management of pandemics from the ICT perspective. For example, a deficient area was found in the protection of patient privacy. The national and state governments need to enact and implement policies to address this issue and clearly define what information needs to be shared with whom and in which way to control the access. The implementation of those eHealth solutions would more likely succeed if there is a RBAC control feature in compliance with these policies. Health care providers and patients' concern over the security of patient information and protection of patient privacy has been identified in the literature as one of the most significant factors influential to their acceptance of eHealth [36].

Future Work

In the future, similar studies can be conducted at various health care settings (eg, residential aged care centers and primary health care centers) to manage and plan the implementation of varied

and specific eHealth systems such as electronic health records, elearning, chronic illness management, telecardiology, teleradiology, and teledermatology.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview Guide.

[PDF File (Adobe PDF File), 229KB - ijmr_v1i2e20_app1.pdf]

References

1. Homeland Security Council US. georgewbush-whitehouse.archives.gov. National strategy for Pandemic Influenza. 2005 URL: http://georgewbush-whitehouse.archives.gov/homeland/pandemic-influenza.html [accessed 2012-11-29] [WebCite Cache ID 6CYDgiM7C]
2. Watkins RE, Cooke FC, Donovan RJ, Macintyre CR, Itzwerth R, Plant AJ. Tackle the problem when it gets here: pandemic preparedness among small and medium businesses. Qual Health Res 2008 Jul;18(7):902-912. [doi: 10.1177/1049732308318032] [Medline: 18552317]
3. World Health Organization (WHO). 2010. Pandemic (H1N1) 2009-update 105 URL: http://www.who.int/sorry/ [accessed 2012-11-29] [WebCite Cache ID 6CYDis7u5]
4. Fraser C, Donnelly CA, Cauchemez S, Hanage WP, Van Kerkhove MD, Hollingsworth TD, WHO Rapid Pandemic Assessment Collaboration. Pandemic potential of a strain of influenza A (H1N1): early findings. Science 2009 Jun 19;324(5934):1557-1561 [FREE Full text] [doi: 10.1126/science.1176062] [Medline: 19433588]
5. Silber D. European Commission, Information Society, eHealth Conference; Atlanta, Belgium. 2003. The case for ehealth URL: http://ec.europa.eu/information_society/eeurope/ehealth/conference/2003/doc/the_case_for_eHealth.pdf [accessed 2012-11-29] [WebCite Cache ID 6CYDnJq94]
6. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
7. Kaur K, Forducey PG, Glueckauf RL. Prototype database for telerehabilitation. Telemed J E Health 2004;10(2):213-222. [Medline: 15319051]
8. Australia Dept of Health and Ageing (DoHA). National Health partnership, Australia. Australian action plan for pandemic influenza URL: http://www.nphp.gov.au/publications/cdna/flu_pandemic_plan_oct03.pdf [accessed 2012-11-29] [WebCite Cache ID 6CYDxuq59]
9. Li J, Moore N, Akter S, Bleistein S, Ray P. mHealth for Influenza Pandemic Surveillance in Developing Countries. 2010 Jan Presented at: The 43 Hawaii International Conference on System Sciences (HICSS); 2010; Hawaii, United States p. 5-8.
10. Li J, Ray P. Applications of E-Health for Pandemic Management. 2010 Presented at: IEEE 12th International Conference on e-Health Networking, Application & Services; 2010; Lyon France p. 1-3.
11. Asangansi I, Braa K. The emergence of mobile-supported national health information systems in developing countries. Stud Health Technol Inform 2010;160(Pt 1):540-544. [Medline: 20841745]
12. Miscione G. Telemedicine in the upper amazon: Inetrplay with local health care practices. MIS Quarterly 2007;31(2):403-425.
13. Demiris G, Oliver DRP, Porock D, Courtney K. Home telehealth: The Missouri telehospice project: Background and next steps. Home Health Care Technology Report 2004;1(49):55-57.
14. Jennett P, Jackson A, Healy T, Ho K, Kazanjian A, Woollard R, et al. A study of a rural community's readiness for telehealth. J Telemed Telecare 2003;9(5):259-263. [doi: 10.1258/135763303769211265] [Medline: 14599328]
15. Ingersoll GL, Kirsch JC, Merk SE, Lightfoot J. Relationship of organizational culture and readiness for change to employee commitment to the organization. J Nurs Adm 2000 Jan;30(1):11-20. [Medline: 10650431]
16. Lehman WE, Greener JM, Simpson DD. Assessing organizational readiness for change. J Subst Abuse Treat 2002 Jun;22(4):197-209. [Medline: 12072164]
17. Zaltman G, Duncan R, Holbek J. Innovations and Organizations. In: Innovations and Organizations. New York: Wiley; 1973.
18. Yin RK. Case study research, design and methods (3rd ed.). Thousand Oaks, California: Sage Publications, Inc; 2003.
19. Cronbach LJ. Reviewed by Worthen B. R. Toward reform of program evaluation: Aims, methods, and institutional arrangements. San Francisco: Jossey-Bass; 1980.
20. Guba EG, Lincoln YS. Effective evaluation. San Francisco: Jossey-Bass; 1981.
21. Patton MQ. Qualitative evaluation and research methods (2nd ed.). Thousand Oaks, California: Sage Publications, Inc; 1990.

22. US General Accounting Office, Program Evaluation and Methodology Division. Case study evaluations. Washington, DC: Government Printing Office; 1990.
23. Li J, Ray P, Seale H, MacIntyre R. An E-Health Readiness Assessment Framework for Public Health Services - Pandemic Perspective. 2012 Jan Presented at: The 45 ational Conference on System Sciences (HICSS); 2012; Hawaii USA p. 4-7.
24. Gómez Reynoso JM, Tulu B. Electronic Medical Records Adoption Challenges in Mexico. In: AMIA 2007 Symposium Proceedings. 2007 Presented at: AMIA 2007 Symposium; 10/11/2007 - 14/11/2007; CHICAGO, USA p. 1093.
25. Hsiao SJ, Li YC, Chen YL, Ko HC. Critical Factors for the Adoption of Mobile Nursing Information Systems in Taiwan: the Nursing Department Administrators' Perspective. Journal of Medical Systems 2008;33:369-377. [doi: 10.1007/s10916-008-9199-8]
26. Halamka J, Aranow M, Ascenzo C, Bates DW, Berry K, Debor G, et al. E-Prescribing collaboration in Massachusetts: early experiences from regional prescribing projects. J Am Med Inform Assoc 2006;13(3):239-244 [FREE Full text] [doi: 10.1197/jamia.M2028] [Medline: 16501174]
27. Kemper AR, Uren RL, Clark SJ. Adoption of electronic health records in primary care pediatric practices. Pediatrics 2006 Jul;118(1):e20-e24 [FREE Full text] [doi: 10.1542/peds.2005-3000] [Medline: 16818534]
28. Ramaswamy R. Design and Management of Service Processes. In: ATT Engineering Process Improvement Series Edited by J.W.Wesley. United Kingdom: Addison-Wesley Pub. Co; 1996.
29. Ray P. Integrated management from E-Business Perspective. In: Integrated management from e-business perspective: concepts, architectures, and methodologies. New York: Kluwer Academic/Plenum Publishers; 2003.
30. Bhattacherjee A, Hikmet N. Reconceptualizing Organizational Support and its Effect on Information Technology Usage: Evidence from the Health Care Sector. Journal of Computer Information Systems 2008;48(4):69-76.
31. Kaushal R, Bates DW, Jenter CA, Mills SA, Volk LA, Burdick E, et al. Imminent adopters of electronic health records in ambulatory care. Inform Prim Care 2009;17(1):7-15. [Medline: 19490768]
32. Jennett PA, Gagnon MP, Brandstadt HK. Preparing for success: readiness models for rural telehealth. J Postgrad Med 2005;51(4):279-285 [FREE Full text] [Medline: 16388170]
33. Wickramasinghe NS, Fadlalla AMA, Geisler E, Schaffer JL. A framework for assessing e-health preparedness. Int J Electron Healthc 2005;1(3):316-334. [Medline: 18048213]
34. Khoja S, Scott R, Casebeer AL, Mohsin M, Ishaq AF, Gilani S. e-Health readiness assessment tools for healthcare institutions in developing countries. Telemed J E Health 2007 Aug;13(4):425-431. [doi: 10.1089/tmj.2006.0064] [Medline: 17848110]
35. WHO. 2005. WHO checklists for influenza pandemic preparedness planning URL: http://www.who.int/csr/resources/publications/influenza/FluCheck6web.pdf [accessed 2012-11-29] [WebCite Cache ID 6CYEvPK8y]
36. Boonstra A, Broekhuis M. Barriers to the acceptance of electronic medical records by physicians from systematic review to taxonomy and interventions. BMC Health Serv Res 2010;10:231 [FREE Full text] [doi: 10.1186/1472-6963-10-231] [Medline: 20691097]

Abbreviations

AHS: area health services
ASS: antimicrobial stewardship system
EMR: electronic medical records
ICT: information and communication technologies
ISD: information services division
IT: information technologies
LIS: laboratory information system
NHI: national health identifier
NSW: New South Wales, Australia
RBAC: role-based access control
SWSD: statewide service desk

Edited by G Eysenbach; submitted 17.09.12; peer-reviewed by PC Masella; comments to author 26.10.12; revised version received 08.11.12; accepted 14.11.12; published 06.12.12.

Please cite as:
Li J, Seale H, Ray P, Rawlinson W, Lewis L, MacIntyre CR
Issues Regarding the Implementation of eHealth: Preparing for Future Influenza Pandemics
Interact J Med Res 2012;1(2):e20
URL: http://www.i-jmr.org/2012/2/e20/
doi:10.2196/ijmr.2357
PMID:23611788

Original Paper

The Effectiveness of an Interactive 3-Dimensional Computer Graphics Model for Medical Education

Bayanmunkh Battulga[1], MD, MSc; Takeshi Konishi[1], MSc; Yoko Tamura[1], MD; Hiroki Moriguchi[1], MD, PhD

Department of Medical Informatics, Institute of Health Biosciences, University of Tokushima graduate school, Tokushima, Japan

Corresponding Author:
Bayanmunkh Battulga, MD, MSc
Department of Medical Informatics
Institute of Health Biosciences
University of Tokushima graduate school
Kuramoto 3-18-15
Tokushima, 7708503
Japan
Phone: 81 886339178
Fax: 81 886339411
Email: bayanmunkh.b@gmail.com

Abstract

Background: Medical students often have difficulty achieving a conceptual understanding of 3-dimensional (3D) anatomy, such as bone alignment, muscles, and complex movements, from 2-dimensional (2D) images. To this end, animated and interactive 3-dimensional computer graphics (3DCG) can provide better visual information to users. In medical fields, research on the advantages of 3DCG in medical education is relatively new.

Objective: To determine the educational effectiveness of interactive 3DCG.

Methods: We divided 100 participants (27 men, mean (SD) age 17.9 (0.6) years, and 73 women, mean (SD) age 18.1 (1.1) years) from the Health Sciences University of Mongolia (HSUM) into 3DCG (n = 50) and textbook-only (control) (n = 50) groups. The control group used a textbook and 2D images, while the 3DCG group was trained to use the interactive 3DCG shoulder model in addition to a textbook. We conducted a questionnaire survey via an encrypted satellite network between HSUM and Tokushima University. The questionnaire was scored on a 5-point Likert scale from strongly disagree (score 1) to strongly agree (score 5).

Results: Interactive 3DCG was effective in undergraduate medical education. Specifically, there was a significant difference in mean (SD) scores between the 3DCG and control groups in their response to questionnaire items regarding content (4.26 (0.69) vs 3.85 (0.68), $P = .001$) and teaching methods (4.33 (0.65) vs 3.74 (0.79), $P < .001$), but no significant difference in the Web category. Participants also provided meaningful comments on the advantages of interactive 3DCG.

Conclusions: Interactive 3DCG materials have positive effects on medical education when properly integrated into conventional education. In particular, our results suggest that interactive 3DCG is more efficient than textbooks alone in medical education and can motivate students to understand complex anatomical structures.

(Interact J Med Res 2012;1(2):e2) doi:10.2196/ijmr.2172

KEYWORDS

Medical education; electronic health information; interactive 3D CG; educational effectiveness; Web-based learning management system; satellite network

Introduction

The Internet has become a social platform where millions of health consumers access and share health information [1]. One such medium, eHealth [2], has brought about improvements in public health and the health care system. Medical information technology has influenced the medical profession during the last decade and will continue to make advances. For example, 3-dimensional (3D) presentation of information is being increasingly used in medical education and health care [3].

Modern human anatomy pedagogy includes cadaver dissection, multimedia presentations, practical procedures, surface and

clinical anatomy, and radiological imaging [4]. Cadaver dissection is the standard method of learning anatomy and allows for a haptic understanding of 3D anatomical structures [5,6] but is expensive and time consuming [7], and curriculum hours for anatomy decrease yearly [8,9]. The current trend in medical education is to achieve anatomical understanding in less time by using information and communication technology. Information and communication technology can approximate human anatomy and motion through 3-dimensional computer graphics models (3DCG), but achieving realistic anatomy and motion is more difficult.

Students often have difficulty achieving a spatial understanding of 3D anatomy from 2-dimensional (2D) images and text. This can increase cognitive load and hinder anatomy learning for students with poor spatial skills [10-13]. 3DCG models promise to overcome many of these educational challenges. Mayer's cognitive theory of multimedia learning states that students learn best by using both images and words in an electronic learning environment [14-16]. 3DCG visually provides semireal information to users, thus enabling them to understand the content easily, and the interactivity of 3DCG content improves comprehension. 3DCG animation and interactive 3DCG have been developed at several institutions [17,18]. Kobayashi et al reported that it was easier and more accurate to explain details of surgery using 3DCG animation than 2D illustrations [19]. In addition, methods have been developed that enable users to make highly specialized 3DCG content on the Web [20].

This study was conducted by researchers at the Health Sciences University of Mongolia (HSUM) and the University of Tokushima, Japan. We chose a high-speed satellite communication network because the Internet has not completely spread to rural areas of Mongolia, which is the fifth-largest country in Asia with 2.6 million people (as of 2007). Mongolia has clear skies and annual precipitation as low as 200 mm per year, creating ideal conditions for satellite communication. Our study was selected by the Ministry of Internal Affairs and Communications of Japan as an experimental application for data collection and as a developmental application for satellite communication authorized by the Association of Radio Industries and Businesses for Japan [21].

Improvements in personal computer performance have led to an increase in the development of 3DCG content. The use of 3DCG models has advantages over traditional anatomy instruction methods; however, their development and adoption are time consuming and costly. Thus, new educational information and communication technology instruction methods are needed. To this end, we aimed to determine the educational effectiveness of interactive 3DCG using an interactive 3DCG shoulder model.

Methods

Development of the Interactive 3DCG Model

We selected the shoulder for this experiment given its anatomical complexity and because it is considered one of the most difficult joints for medical students to understand in human anatomy. The Department of Anatomy and Developmental Neurobiology, University of Tokushima carefully examined anatomical accuracy, such as the relative spatial relationship of each structure, at every development stage to ensure that our interactive 3DCG models would be of high quality and accurate (see Multimedia Appendix 1).

The process of model creation is not trivial. We built the models in LightWave 3D (NewTek, Inc., San Antonio, TX, USA) and exported them as object files to Blender (an open source 3D program; blender.org, Amsterdam, the Netherlands). Blender was used to generate clean U and V space texture maps for the models. The models were then sent back to LightWave as object files. We created the textures in Photoshop (Adobe Systems Incorporated, San Jose, CA, USA) from photo references, applied them to U and V space maps in LightWave 3D, then exported them in Filmbox format and imported them into Unity3D [22]. Shoulder movements were added in LightWave and exported as Filmbox animation data. Once the model and animation data were imported into Unity3D, we wrote scripts to allow interaction with the models. Unity's workflow allows swapping of models if changes have been made in LightWave 3D. The initial animation of shoulder movement was quite rigid, so we used a motion capture system to achieve more natural motion.

Motion Capture System

We used the Vicon MX motion capture system (Vicon Motion Systems, Oxford, United Kingdom) at Tokushima University Hospital [23]. Motion capture data were collected at 150 Hz using a near-infrared (780 nm) passive 8-camera system (Vicon MX T20; Vicon Motion Systems). A 3D position sensor captured the reflected rays of 9.5 mm diameter reflective markers. Nexus 1.4.1 (Vicon Motion Systems) software recorded 3D positions of the markers and extracted vector data. We took screenshots of the motion capture vector data and then moved the shoulder to match those screenshots for each part of the motion. We correlated the skeleton to the visual data, taking screenshots of data over time. This process improved the smoothness of shoulder movements.

3DCG Model User Interface and Textbook

The menu for the 3DCG interactive manipulation tools on the left side of the screen has two functional components: one for anatomy and the other for shoulder movements, with labeling ability in English or Japanese. Figure 1 shows the 3DCG of the area surrounding the shoulder and the terminology of each part on the upper right corner. Figure 2 shows the movements of the shoulder bones and upper extremity. Several important movements are available, including elevation, depression, retraction, protraction, flexion, extension, vertical abduction, vertical adduction, horizontal abduction, and horizontal adduction. Both views allow the user to zoom in and out, and to focus on a region of interest. Users are also able to quickly rotate or move to a specific angle such as anterior, lateral, and posterior. To study specific anatomical regions, the tool enables users to select single or multiple objects and hide them or make them semitransparent. We also developed an original textbook for this study. To provide a suitable condition for comparison, we replaced the black and white figures traditionally used in lectures at HSUM, which were drawn from a standard

anatomical viewpoint, with a new textbook based on 3DCG images selected from the interactive 3DCG system. We reproduced 34 images (anterior, lateral, posterior, and other views with appropriate angle and magnitude for evaluation) from the 3DCG shoulder models in gif format for the textbook and added appropriate text to explain the images.

We used Wideband InterNetworking engineering test and Demonstration Satellite for the communication system, which was jointly developed by the Japan Aerospace Exploration Agency and the National Institute of Information and Communications Technology of Japan. A small antenna 1.2 m in diameter receives up to 155 Mbps of data and transmits up to 6 Mbps, while an antenna approximately 5 m in diameter enables 2-way communication up to 1.2 Gbps [24]. For our study, the Japan Aerospace Exploration Agency allocated reception bands from 1.31 Mbps to 1.38 Mbps for the uplink connection, and from 15.0 Mbps to 20.7 Mbps for the downlink connection. The link was encrypted using an Internet Protocol Security virtual private network based on a Cisco ASA 5505 router (Cisco Systems Inc., San Jose, CA, USA) provided by a joint research project with Mitsubishi Electric Information Network Corporation (Figure 3).

Figure 1. Anatomical view of the 3-dimensional computer graphic showing the shoulder area.

Figure 2. Movement view of the 3-dimensional computer graphic of the shoulder bones and upper extremity.

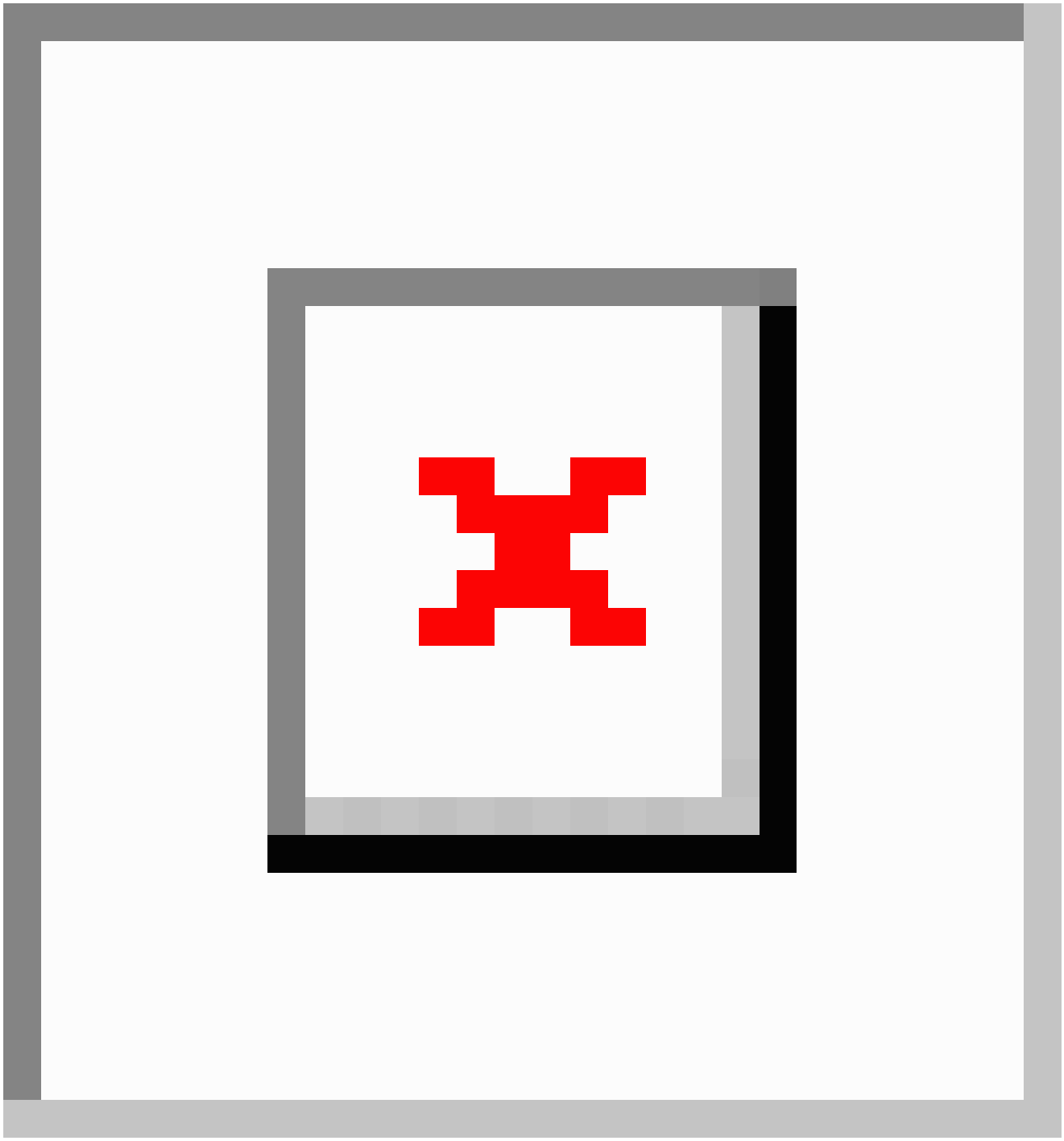
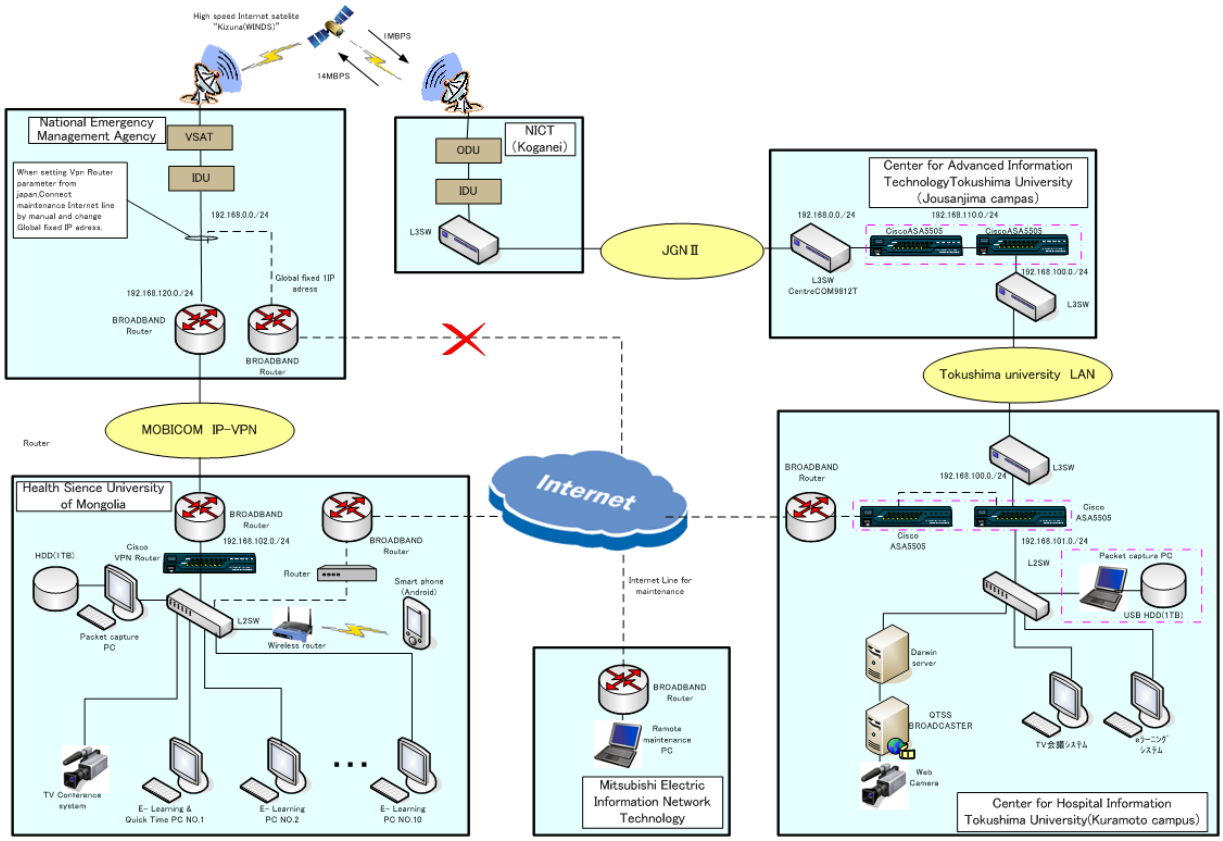


Figure 3. Network architecture.



Throughput of the Satellite Network

Round-trip time is the signal delay between the University of Tokushima and HSUM via the satellite network and the Internet. The round-trip time was between 1499.3 ms and 643.8 ms, with an average of 729.6 ms. The congestion window is a Transmission Control Protocol parameter that regulates the send window. The congestion window of Windows XP ranges from 16 KB in default to 64 KB. Therefore, the maximum throughput available in the Transmission Control Protocol is 0.795 Mbps, which is used for Hypertext Transfer Protocol. However, the actual throughput value was 0.384 Mbps or less.

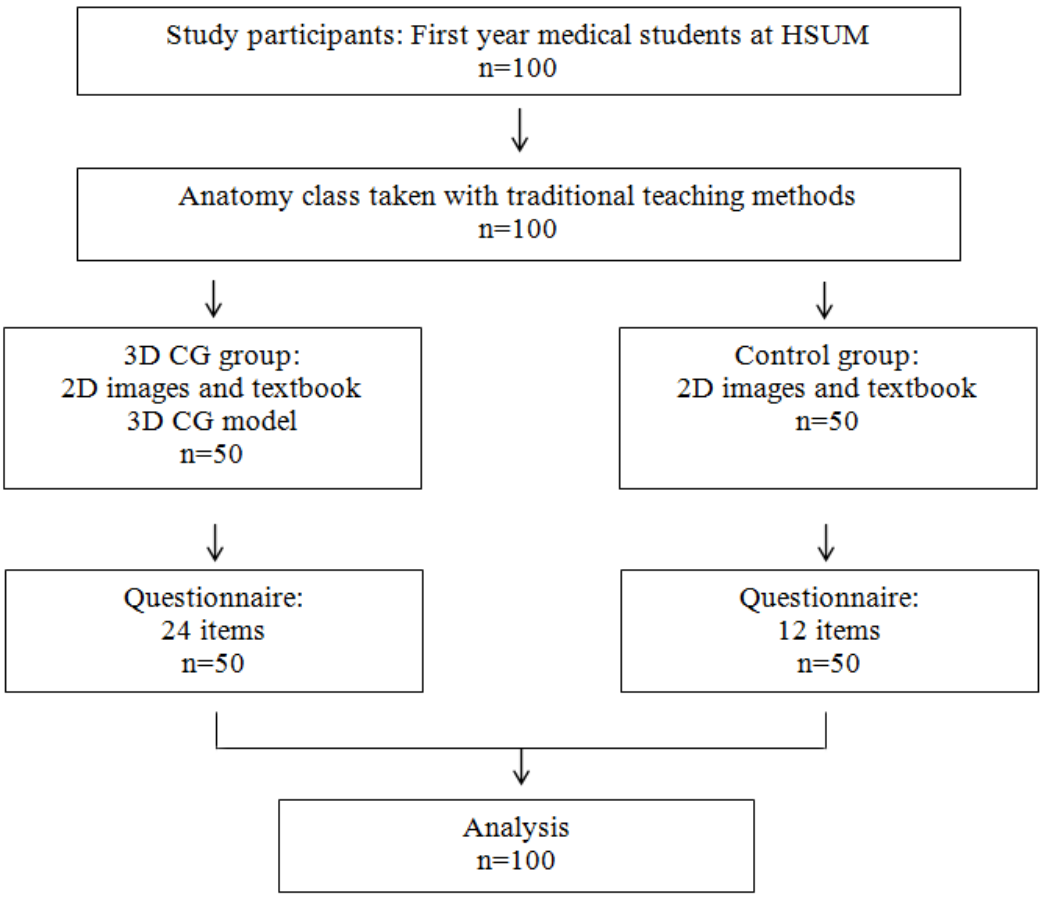
Web-Based Learning Management System

All activity associated with the course was hosted on a Linux-based server running the Apache Web server (Apache Software Foundation, Los Angeles, CA, USA), the PostgreSQL database server (PostgreSQL Global Development Group, http://www.postgresql.org/), and the CentOS operating system (CentOS Project, http://www.centos.org/). Our proprietary learning management system (LMS) allowed us to create a course website with a unique log-in password and ID for each student. We developed the interactive 3DCG model, textbook, and questionnaires, and embedded them into this LMS. These components were used for the related experimental section. We applied Java Web applications for the system and adopted Unity3D for the 3DCG container, which is an integrated 3D platform for 3D games and interactive content on the Web.

Study Design

Figure 4 shows the study design. All participants who had taken anatomy classes and finished cadaver dissection 3 months previously received a brief introduction to the LMS before the experiment. The study was conducted over a 1-week period. We divided participants into a 3DCG group and a textbook-only (control) group. The control and 3DCG groups were also given instructions on how to use the textbook, and the 3DCG group was given additional training to manipulate the interactive 3DCG model. Each participant received a 1-hour training session. Finally, all participants completed a questionnaire on the LMS.

Figure 4. Study design. 2D = 2-dimensional, 3DCG = 3-dimensional computer graphics, HSUM = Health Sciences University of Mongolia.



Statistical Analysis

Statistical analysis was performed using SPSS (version 16.0 for Windows; IBM Japan Inc., Tokyo, Japan). We conducted both the independent-samples *t* test and the Mann-Whitney *U* test to compare the two groups, with $P < .05$ defined as statistically significant. The internal consistency coefficient (Cronbach alpha) was calculated for both groups.

Results

Participants

A total of 100 first-year medical students (27 men and 73 women; Table 1) from HSUM volunteered, gave informed consent to participate in the study, and were randomly assigned to either the 3DCG group (n = 50) or the control group (n = 50). The mean age of participants was 18.1 (SD 1.1) years (men: 17.9 (SD 0.6) years, women: 18.1 (SD 1.1); range 16–25 years). Participants were freshmen at HSUM and were instructed in practical computer skills for half a semester (64 hours).

Table 1. Gender distribution in the two study groups (n = 100).

Gender	3DCG[a] group (n = 50)	Control group (n = 50)
Male	13	14
Female	37	36

[a] 3-dimensional computer graphics.

e-Learning Questionnaire

The questionnaire had 24 items grouped into four categories: content (3 items), teaching methods (6 items), Web (3 items), and 3DCG model interface (12 items). Each item was scored based on a 5-point Likert scale: strongly agree (score 5), agree

(score 4), neutral (score 3), disagree (score 2), and strongly disagree (score 1).

Table 2. Questionnaire scores for content, teaching methods, and Web items (n = 100).

Questionnaire item		Response					Mean score
		Strongly agree	Agree	Neutral	Disagree	Strongly disagree	
Q1	The content is useful.	30	51	19	0	0	4.11
Q2	The content is easy to read and understand.	19	55	22	4	0	3.89
Q3	The content is well formatted and well designed.	25	52	21	2	0	4.00
Q4	The support for my study is effective.	31	52	17	0	0	4.14
Q5	This teaching method can improve my knowledge.	27	59	13	1	0	4.12
Q6	This teaching method can help my learning.	27	53	19	1	0	4.06
Q7	This teaching method motivates me when I learn.	38	40	18	4	0	4.12
Q8	This teaching method gives me enough time in the lesson.	24	40	26	9	1	3.77
Q9	This lesson is appropriate for my learning demand.	31	43	25	1	0	4.04
Q10	The webpage is attractive.	25	58	15	1	1	4.05
Q11	The screen design is clear.	26	53	18	2	1	4.01
Q12	The menu is easy to use.	25	57	17	0	1	4.05

Table 2 shows that mean scores ranged from 3.77 to 4.14. For each item, 19%–38% of participants responded strongly agree, 40%–59% responded agree, 13%–26% responded neutral, 0%–9% responded disagree, and 0%–1% responded strongly disagree. As a result, 64%–86% of participants responded either strongly agree or agree to each item.

Table 3. Questionnaire scores for the 3-dimensional computer graphics (3DCG) model interface (n = 50).

Questionnaire item		Response					Mean score
		Strongly agree	Agree	Neutral	Disagree	Strongly disagree	
Q13	The interface for interacting with the 3D[a] content is accessible.	17	26	7	0	0	4.2
Q14	The volume of information in the 3D module is appropriate.	15	26	8	1	0	4.1
Q15	I am satisfied with the 360° rotation of the model.	32	17	1	0	0	4.62
Q16	I am satisfied with the selection menu.	13	35	2	0	0	4.22
Q17	I am satisfied with the show-and-hide function.	21	20	9	0	0	4.24
Q17	I am satisfied with the transparent and opaque function.	13	30	7	0	0	4.12
Q19	I am satisfied with the zoom function.	29	18	3	0	0	4.52
Q20	The screen size is appropriate.	11	26	13	0	0	3.96
Q21	I am satisfied with clicking the mouse to show anatomical terminology.	29	20	1	0	0	4.56
Q22	I am interested in 3DCG.	33	14	3	0	0	4.6
Q23	I am satisfied with the movement menu.	26	21	3	0	0	4.46
Q24	The 3DCG content is of high quality.	18	31	1	0	0	4.34

[a] 3-dimensional.

Table 3 shows that mean scores for the 3DCG module range from 3.96 to 4.62. For each item, 22%–66% responded strongly agree, 28%–70% responded agree, 2%–26% responded neutral, and 0%–1% responded disagree. None of the participants responded strongly disagree to the statements. Overall, 74%–98% of the participants responded either strongly agree or agree. The reliability of the entire questionnaire was acceptable (Cronbach alpha = .902).

Table 4. Comparison of mean (SD) questionnaire scores[a] between 3-dimensional computer graphics (3D G) and control groups.

Category	Questionnaire item	3DCG group	Control group	*P* value
Content	The content is useful.	4.34 (0.63)	3.88 (0.69)	.001[b]
	The content is easy to read and understand.	3.90 (0.70)	3.88 (0.80)	.97
	The content is well formatted and well designed.	4.18 (0.75)	3.82 (0.69)	.01[b]
Teaching methods	The support for my study is effective.	4.36 (0.63)	3.92 (0.66)	.001[b]
	This teaching method can improve my knowledge.	4.40 (0.61)	3.84 (0.58)	<.001[b]
	This teaching method can help my learning.	4.28 (0.64)	3.84 (0.71)	.002[b]
	This teaching method motivates me when I learn.	4.52 (0.58)	3.72 (0.88)	<.001[b]
	This teaching method gives me enough time in the lesson.	4.14 (0.70)	3.40 (1.03)	<.001[b]
	This lesson is appropriate for my learning demand.	4.32 (0.71)	3.76 (0.74)	<.001[b]
Web	The webpage is attractive.	4.26 (0.56)	3.84 (0.82)	.006[b]
	The screen design is clear.	4.10 (0.61)	3.92 (0.92)	.45
	The menu is easy to use.	4.14 (0.61)	3.96 (0.81)	.29

[a] 5-point Likert scale from strongly disagree (score 1) to strongly agree (score 5).

[b] Significant difference (Mann-Whitney *U* test).

Table 4 and Figure 5 present the mean (SD) scores. Table 4 compares scores between the 3DCG and control groups for each item. The 3DCG group scores ranged from 3.90 (SD 0.70) to 4.52 (SD 0.58), and control group scores ranged from 3.40 (SD 1.03) to 3.96 (SD 0.81). Differences in learning motivation scores were the largest, with the 3DCG group giving an average score 0.8 higher than the control group. Participants' comments on the advantages of using the interactive 3DCG model included "very interesting," "realistic," "saved time," "more understandable," "amazing movements," and "clarity." There were also some comments regarding disadvantages, such as the necessity for "muscle movement," "more detailed anatomy," and "larger screen size," which would require compilation to change the frame size.

Combined scores for the three categories are shown in Figure 5. There was a significant difference between the 3DCG and control groups for content (4.26 (SD 0.69) vs 3.85 (SD 0.68); $P = .001$) and teaching methods (4.33 (SD 0.65) vs 3.74 (SD 0.79); $P < .001$). No significant difference was found between the groups in the Web category. The mean score for male participants was higher than that of female participants, except for Q11 (Figure 6). There was no significant difference between men and women for Q1–Q8 and Q10–Q12, but the difference for Q9 was significant.

Figure 5. Mean scores of the 3-dimensional computer graphics (3DCG) and control groups for the three main categories. LMS = learning management system, n.s. = not significant.

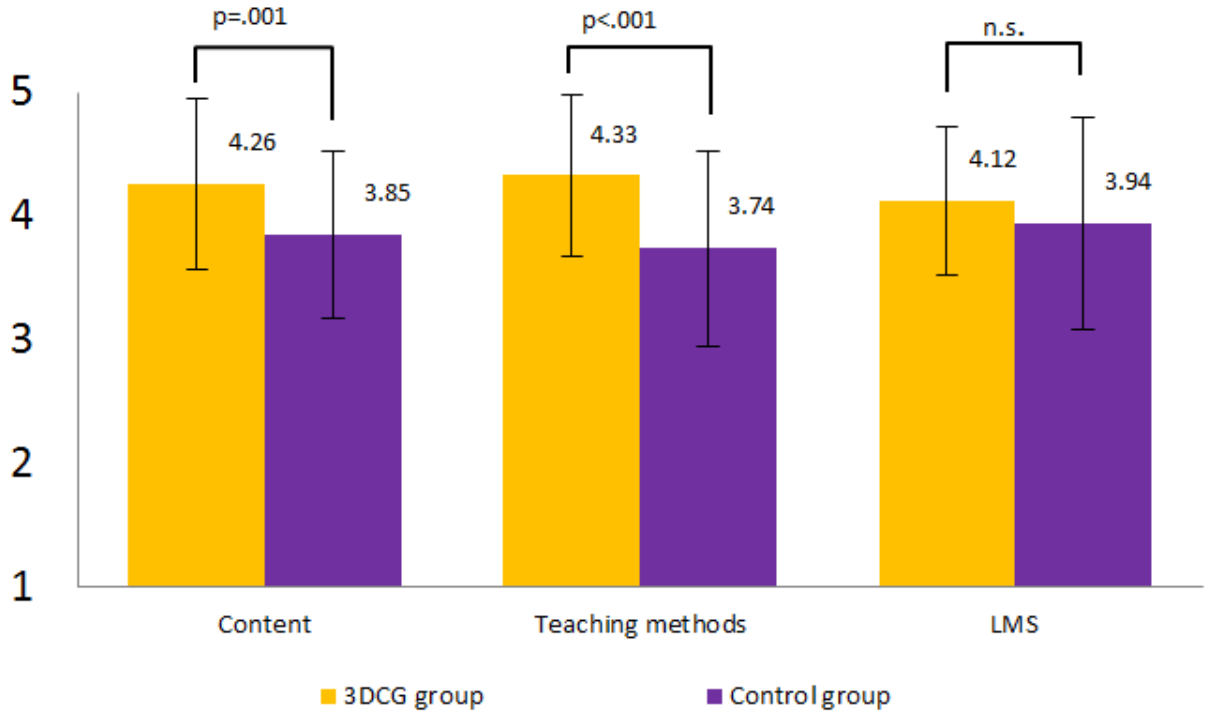
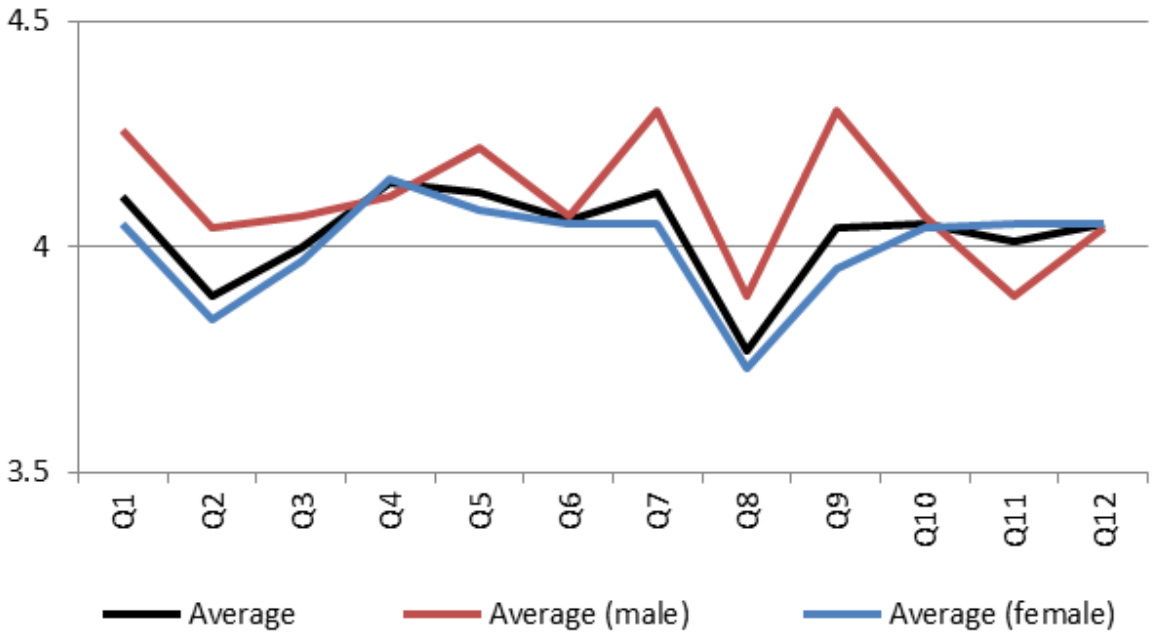


Figure 6. Questionnaire scores by gender.



Discussion

In this study, we investigated the educational effectiveness of using an interactive 3DCG model as a supplement to conventional learning methods. Our results show that the interactive 3DCG model is effective in undergraduate medical education and can enhance the motivation of medical students. We employed an LMS with two kinds of content. Content for the 3DCG group included Unity3D-based material, while content for the control group included only text and 2D images.

Responses of all participants (n = 100) on common questionnaire items in three categories are presented in Table 2. Items Q1, Q4, Q5, and Q7 had the highest scores, while scores for Q2 and Q8 were relatively low. Strongly agree responses were particularly high for items Q1, Q4, Q7, and Q9. Disagree and strongly disagree comprised less than 10% of all questionnaire

item responses. These results indicate that participants generally accepted the LMS, although they may require more time to acclimate to the new learning system.

Table 3 shows that 11 items out of 12 had mean scores higher than 4. The relatively low scores for screen size of the 3DCG may be due to the small size of the Unity framework, which has a fixed size of 950 × 534 pixels. Participants indicated a high interest in the 3DCG model and were satisfied with the 360° rotation, zooming function, ability to show terminology, and movement of the shoulder joint.

Table 4 shows significant differences in 9 out of 12 items. Scores are significantly different for 2 items in the content category. However, there was no difference for Q2, which was expected because textbook images were selected and copied from the 3DCG shoulder models. Q2 also has one of the worst scores in Table 2 and is particularly low compared with other scores in the 3DCG group (Table 4), which could be due to the fact that participants were not accustomed to operating the interactive 3DCG. We believe that participants in the 3DCG group accepted the interactive 3DCG because they found it to be effective and motivating, and it satisfied their learning demands. The significant difference for item Q10 indicates that evaluation of the LMS depends on content quality, such as interactive 3DCG, which attracts interest. The mean score of the 3DCG group (4.26), which is higher than that of the control group (3.84), indicates a strong interest in 3DCG and the necessity of the interactive component, including functions for scaling, changing perspectives, and movement. This may derive from the need to be appropriately positioned to view specific anatomical structures for better comprehension. The 3DCG group gave high scores for items in the teaching methods category, and particularly high scores on items addressing usefulness of content, demonstrating that students desired a better way to view specific anatomical structures. Garg et al [25] investigated the usefulness of computer-mediated anatomical 3D reconstructions in anatomical learning and found that learners with low visuospatial ability performed worse on an anatomical knowledge test following the multiple-view condition than following the key view. They also concluded that the key view is important for understanding 3DCG that has many dominants, which indicate the region of concern of users. On the basis of their key view theory, learners using an interactive 3DCG model that was made for a crucial anatomical site, such as the shoulder, could naturally select any important key views themselves for better understanding of the specific anatomical region. The spontaneous, easy, and unburdened way of searching in key view may promote student learning, while textbooks provide a restricted viewpoint. The significantly different scores in the teaching methods category suggest that participants in the 3DCG group felt the 3DCG model motivated their learning, improved their knowledge, and was effective for studying. Moreover, the model satisfied their learning demands and was helpful for self-study. There were no significant differences in scores for items regarding the design and menu in the Web category because the screens were the same. This might strengthen the validity of the responses provided by the two groups.

Comments from participants also suggested that interactive 3DCG increased the motivation to learn a large number of anatomical structures and clarified anatomy. Students generally dislike memorizing many names and learning the complexity of nerves and blood vessels and how joints move, in a short period of time. Scores of male participants tended to be higher than scores of female participants. Some research has revealed that women have higher computer anxiety [26] and less learning emphasis, strategy [27], computer self-efficacy, perceived usefulness, perceived ease of use, and behavioral intention to use e-learning [28]. Scores on Q9 showed that women had significantly lower learning demands. Men and women should have similar learning attitudes toward 3DCG anatomical content, as shown in these reports. They are also required to acquire terminology that must be remembered, although memorization is difficult in a limited time frame.

Some studies have attempted to evaluate the effectiveness of learning tools. Findings include that animated visual tools are more effective than static visual tools [29], the use of 3D animation leads to better topographical and theoretical understanding [30], 3D multimedia software has a positive impact on dental education [31], and 3D surgical simulators [32] and 3D larynx models [33] can enhance student learning through increased motivation. Our findings are consistent with these reports. Another study [34] showed that 2D visualization was superior to 3D visualization in improving the understanding of organic molecule structures, but scores for the 2D and 3D groups were similar. The report stressed that better results would be achieved under conditions of greater familiarity with 3D. A limitation of our study is that we assessed learning effectiveness using the 3DCG model only by questionnaire. We have not measured the outcome of anatomical understanding.

We conclude that interactive 3DCG materials have positive effects on medical education when properly integrated into conventional education. In particular, the interactive 3DCG motivated participants to understand a complex anatomical structure.

Acknowledgments

We would like to thank the Japan Manned Space Systems Corporation, Professor Yoshihiro Fukui and Hiromi Sakata, MD, at the University of Tokushima Graduate School Department of Anatomy and Developmental Neurobiology, Professor Natsuki Yasui and Shinjiro Takata at the Tokushima University Hospital Department of Orthopedics, and Professor Tserenkhuu Lkhagvasuren at the Health Sciences University of Mongolia.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Examples of screenshots and pictures of the web-based learning management system and development of the interactive 3D CG.

[PPTX File, 869KB - ijmr_v1i2e2_app1.pptx]

References

1. Eysenbach G. Medicine 2.0: social networking, collaboration, participation, apomediation, and openness. J Med Internet Res 2008;10(3):e22 [FREE Full text] [doi: 10.2196/jmir.1030] [Medline: 18725354]
2. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
3. Marks SC. The role of three-dimensional information in health care and medical education: the implications for anatomy and dissection. Clin Anat 2000;13(6):448-452. [doi: 10.1002/1098-2353(2000)13:6<448::AID-CA10>3.0.CO;2-U] [Medline: 11111898]
4. Sugand K, Abrahams P, Khurana A. The anatomy of anatomy: a review for its modernization. Anat Sci Educ 2010;3(2):83-93. [doi: 10.1002/ase.139] [Medline: 20205265]
5. Turney BW. Anatomy in a modern medical curriculum. Ann R Coll Surg Engl 2007 Mar;89(2):104-107. [doi: 10.1308/003588407X168244] [Medline: 17346399]
6. Collins JP. Modern approaches to teaching and learning anatomy. BMJ 2008;337:a1310. [Medline: 18782839]
7. McLachlan JC, Patten D. Anatomy teaching: ghosts of the past, present and future. Med Educ 2006 Mar;40(3):243-253. [doi: 10.1111/j.1365-2929.2006.02401.x] [Medline: 16483327]
8. Cottam WW. Adequacy of medical school gross anatomy education as perceived by certain postgraduate residency programs and anatomy course directors. Clinical Anatomy 1999;12(1):55-65. [doi: 10.1002/(SICI)1098-2353(1999)12:1<55::AID-CA8>3.0.CO;2-O] [Medline: 9890730]
9. Drake RL, McBride JM, Lachman N, Pawlina W. Medical education in the anatomical sciences: the winds of change continue to blow. Anat Sci Educ 2009;2(6):253-259. [doi: 10.1002/ase.117] [Medline: 19890982]
10. Garg AX, Norman G, Sperotable L. How medical students learn spatial anatomy. The Lancet 2001 Feb 3;357(9253):363-364. [doi: 10.1016/S0140-6736(00)03649-7] [Medline: 11211004]
11. McLachlan JC, Bligh J, Bradley P, Searle J. Teaching anatomy without cadavers. Med Educ 2004 Apr;38(4):418-424. [doi: 10.1046/j.1365-2923.2004.01795.x] [Medline: 15025643]
12. Levinson AJ, Weaver B, Garside S, McGinn H, Norman GR. Virtual reality and brain anatomy: a randomised trial of e-learning instructional designs. Med Educ 2007 May;41(5):495-501. [doi: 10.1111/j.1365-2929.2006.02694.x] [Medline: 17470079]
13. Huk T. Who benefits from learning with 3D models? The case of spatial ability. J Comput Assist Learn 2006;22:392-404. [doi: 10.1111/j.1365-2729.2006.00180.x]
14. Mayer RE, Heiser J, Lonn S. Cognitive constraints on multimedia learning: when presenting more material results in less understanding. J Educ Psychol 2001;93(1):187-198. [doi: 10.1037//0022-0663.93.1.187]
15. Mayer RE. The promise of multimedia learning: using the same instructional design methods across different media. Learn Instruct 2003;13(2):125-139. [doi: 10.1016/S0959-4752(02)00016-6]
16. Mayer RE. Multimedia Learning. 2nd edition. Cambridge: Cambridge University Press; 2009.
17. Silén C, Wirell S, Kvist J, Nylander E, Smedby O. Advanced 3D visualization in student-centred medical education. Med Teach 2008 Jun;30(5):e115-e124. [doi: 10.1080/01421590801932228] [Medline: 18576181]
18. Stull AT, Hegarty M, Mayer RE. Anatomy learning with virtual objects. In: Cognitive Shape Processing. 2010 Presented at: AAAI 2010 Spring Symposium; Mar 22-24, 2010; Stanford, CA. USA p. 39-44 URL: http://www.spatial-cognition.de/CSP/Stull.pdf
19. Kobayashi M, Nakajima T, Mori A, Tanaka D, Fujino T, Chiyokura H. Three-dimensional computer graphics for surgical procedure learning: Web three-dimensional application for cleft lip repair. Cleft Palate Craniofac J 2006 May;43(3):266-271. [doi: 10.1597/04-009.1] [Medline: 16681399]
20. Warren W, Brinkley JF. Knowledge-based, interactive, custom anatomical scene creation for medical education: the Biolucida system. AMIA Annu Symp Proc 2005:789-793. [Medline: 16779148]
21. Kizuna (WINDS) Experiment Promotion Page.: Japan Aerospace Exploration Agency; 2008. URL: http://winds-ets8.jaxa.jp/winds/en/index.html [accessed 2011-10-02] [WebCite Cache ID 628PrawZD]
22. Unity 3. Unity Technologies. 2011 URL: http://unity3d.com/unity/ [accessed 2011-10-02] [WebCite Cache ID 628PNTZZl]
23. Vicon MX System. Vicon Motion Systems. 2012 URL: http://www.vicon.com/products/viconmx.html [accessed 2011-10-02] [WebCite Cache ID 628PcwTlV]

24. Japan Aerospace Exploration Agency. WINDS: Wideband InterNetworking Engineering Test and Demonstration Satellite "KIZUNA" URL: http://www.jaxa.jp/pr/brochure/pdf/04/sat07.pdf [accessed 2011-10-03] [WebCite Cache ID 629jIXSxb]
25. Garg AX, Norman GR, Eva KW, Spero L, Sharan S. Is there any real virtue of virtual reality?: the minor role of multiple orientations in learning anatomy from computers. Acad Med 2002 Oct;77(10 Suppl):S97-S99. [Medline: 12377717]
26. Broos A. Gender and information and communication technologies (ICT) anxiety: male self-assurance and female hesitation. Cyberpsychol Behav 2005 Feb;8(1):21-31. [doi: 10.1089/cpb.2005.8.21] [Medline: 15738690]
27. Ausburn L. Gender and learning strategy differences in nontraditional adult students' design preferences in hybrid distance courses. J Interact Online Learn 2004;3(2):1-17 [FREE Full text]
28. Ong CS, Lai JY. Gender differences in perceptions and relationships among dominants of e-learning acceptance. Comput Human Behav 2006;22:816-829. [doi: 10.1016/j.chb.2004.03.006]
29. Lin H, Dwyer FM. The effect of static and animated visualization: a perspective of instructional effectiveness and efficiency. Educ Tech Res Dev 2009;58:155-174. [doi: 10.1007/s11423-009-9133-x]
30. Prinz A, Bolz M, Findl O. Advantage of three dimensional animated teaching over traditional surgical videos for teaching ophthalmic surgery: a randomised study. Br J Ophthalmol 2005 Nov;89(11):1495-1499 [FREE Full text] [doi: 10.1136/bjo.2005.075077] [Medline: 16234460]
31. Hu J, Yu H, Shao J, Li Z, Wang J, Wang Y. Effects of Dental 3D Multimedia System on the performance of junior dental students in preclinical practice: a report from China. Adv Health Sci Educ Theory Pract 2009 Mar;14(1):123-133. [doi: 10.1007/s10459-007-9096-9] [Medline: 18214704]
32. Hariri S, Rawn C, Srivastava S, Youngblood P, Ladd A. Evaluation of a surgical simulator for learning clinical anatomy. Med Educ 2004 Aug;38(8):896-902. [doi: 10.1111/j.1365-2929.2004.01897.x] [Medline: 15271051]
33. Hu A, Wilson T, Ladak H, Haase P, Fung K. Three-dimensional educational computer model of the larynx: voicing a new direction. Arch Otolaryngol Head Neck Surg 2009 Jul;135(7):677-681. [doi: 10.1001/archoto.2009.68] [Medline: 19620589]
34. Oke OK, Alam GM. Comparative evaluation of the effectiveness of 2 and 3D visualizations in students' understanding of structures of organic molecules. Int J Phys Sci 2010;5(5):605-611 [FREE Full text]

Abbreviations

2D: 2-dimensional
3D: 3-dimensional
3DCG: 3-dimensional computer graphics
HSUM: Health Sciences University of Mongolia
LMS: learning management system

Edited by G Eysenbach; submitted 14.05.12; peer-reviewed by R Farfan; comments to author 11.06.12; revised version received 22.06.12; accepted 25.06.12; published 09.07.12.

Please cite as:
Battulga B, Konishi T, Tamura Y, Moriguchi H
The Effectiveness of an Interactive 3-Dimensional Computer Graphics Model for Medical Education
Interact J Med Res 2012;1(2):e2
URL: http://www.i-jmr.org/2012/2/e2/
doi:10.2196/ijmr.2172
PMID:23611759

Original Paper

An Approach to Reducing Information Loss and Achieving Diversity of Sensitive Attributes in k-anonymity Methods

Sunyong Yoo[1]; Moonshik Shin[1]; Doheon Lee[1], PhD

Department of Bio and Brain Engineering, KAIST, Daejeon, Korea, Republic Of

Corresponding Author:
Doheon Lee, PhD
Department of Bio and Brain Engineering, KAIST
Daejeon, 305-701
Korea, Republic Of
Phone: 82 042 350 8681
Fax: 82 042 350 8680
Email: dhlee@kaist.ac.kr

Abstract

Electronic Health Records (EHRs) enable the sharing of patients' medical data. Since EHRs include patients' private data, access by researchers is restricted. Therefore k-anonymity is necessary to keep patients' private data safe without damaging useful medical information. However, k-anonymity cannot prevent sensitive attribute disclosure. An alternative, *l*-diversity, has been proposed as a solution to this problem and is defined as: each Q-block (ie, each set of rows corresponding to the same value for identifiers) contains at least *l* well-represented values for each sensitive attribute. While *l*-diversity protects against sensitive attribute disclosure, it is limited in that it focuses only on diversifying sensitive attributes. The aim of the study is to develop a k-anonymity method that not only minimizes information loss but also achieves diversity of the sensitive attribute. This paper proposes a new privacy protection method that uses conditional entropy and mutual information. This method considers both information loss as well as diversity of sensitive attributes. Conditional entropy can measure the information loss by generalization, and mutual information is used to achieve the diversity of sensitive attributes. This method can offer appropriate Q-blocks for generalization. We used the adult database from the UCI Machine Learning Repository and found that the proposed method can greatly reduce information loss compared with a recent *l*-diversity study. It can also achieve the diversity of sensitive attributes by counting the number of Q-blocks that have leaks of diversity. This study provides a privacy protection method that can improve data utility and protect against sensitive attribute disclosure. The method is viable and should be of interest for further privacy protection in EHR applications.

(Interact J Med Res 2012;1(2):e14) doi:10.2196/ijmr.2140

KEYWORDS

k-anonymity; l-diversity; Information loss; Conditional entropy; Mutual information

Introduction

Society is experiencing exponential growth in the amount of health information. However, this information is distributed across multiple sites, held in a variety of paper and electronic formats, and represented as mixtures of narrative and structured data. Electronic Health Records (EHRs) have been introduced as a method for improving communication between health care providers and improving access to patient data. This use of EHRs has now enabled large and complicated databases of health records to be used for medical and other research. However, as medical record data become more accessible, protecting patient privacy is an increasing concern that should not be overlooked or understated [1-4].

For patient health information to be de-identified, the Health Insurance Portability and Accountability Act (HIPAA) in the United States suggests the "Safe Harbor" technique, which requires 18 data elements to be removed [5,6]. Doing this can protect the confidentiality and privacy of research subjects. De-identification methods have been proposed for removal of identifiers and in general are performed by the following two steps. First, personal identifiers are located within a database. Second, these identifiers are masked, coded, and/or replaced with irreversible values to unauthorized personnel. However, de-identification methods have tended to be quite faulty as the possibility remains of re-identifying a patient by linking or matching the data to other data or by looking at unique characteristics found in the released data.

Avoiding re-identification requires the use of an anonymization method that prevents the data from being linked for identification of the patient. One popular anonymization method is k-anonymity, proposed by Samarati and Sweeny. A dataset satisfies k-anonymity if each record is indistinguishable from at least k-1 other records with respect to certain identifying attributes. This process is usually performed by suppressing or generalizing database entries [7-10].

While k-anonymity protects against identity disclosure, it is not sufficient for preventing sensitive attribute disclosure. To solve this problem, *l*-diversity has been proposed [11,12]. This method requires that each Q-block has at least *l* well-represented values for each sensitive attribute. While *l*-diversity protects against sensitive attribute disclosure, it has a limitation in that it focuses only on diversifying sensitive attributes. However, generalizing attributes leads to an information loss, so reducing the amount of information loss is also important [8,13].

In this paper, we propose a practical method that reduces information loss but still achieves diversity of sensitive attributes. This method is based on conditional entropy and mutual information. Conditional entropy can measure the information loss by generalization between the original database and a generalized database, while mutual information between the generalized database and sensitive attributes can be used to achieve the necessary diversity of sensitive attributes. By applying this method, we were able to offer appropriate Q-blocks for generalization. We used the adult database from the UCI Machine Learning Repository to evaluate the proposed method.

Related Work

Privacy has become an increasingly salient political issue and considerable progress has been made with de-identification. In general, de-identification methods aim to remove a patient's personal information and many other types of PHI (Protected Health Information). The de-identification process means that only explicit identifiers are hidden or removed. Despite using various measures to de-identify health records, it is possible to re-identify them in a large number of cases by using computerized network databases containing voter registration records, hospital discharge records, commercially available databases, and other sources. Indeed, it is likely that between 63% (Golle 2006) and 87% (Sweeney 2000) of the population of the United States could be uniquely identified by using only gender, ZIP code, and date of birth [14,15].

This kind of attack is called a linking attack. We assumed that an individual has a de-identified database containing some clinical data and that those databases also contain attributes (birth, gender, and zip code). If we could get access to an identification database or construct one from public data sources with the same attributes as the database, then it would be easier to link two databases and re-identify the individuals in the research database [16]. This linkage is performed with a set of quasi-identifier (QI) attributes that are in both datasets. In Table 1, *work* and *country* are QI attributes.

Table 1. An example of an original data table.

Index	Quasi-identifier (QI)		Sensitive
	Work	Country	Disease
1	Private	USA	Heart Disease
2	State-gov	Mexico	Cancer
3	Local-gov	Brazil	Cancer
4	Federal-gov	USA	Flu
5	Private	Canada	Heart Disease
6	Self-emp-not-inc	Canada	Heart Disease
7	Self-emp-inc	USA	Flu
8	Private	USA	Heart Disease
9	State-gov	Mexico	Flu

K-Anonymity

To protect data from a linking attack, Samarati and Sweeny proposed k-anonymity [7]. This method generalizes or suppresses the QI attributes so that each record is indistinguishable from at least k-1 other records within the dataset. The larger the value of k, the greater the implied privacy, since no individual can be identified with probability exceeding 1/k through linking attacks alone. For example, Table 1 is the original data table, and Table 2 is an anonymized version of it that satisfies 3-anonymity. 3-anonymity means that at least three instances are identical with respect to QI. We can find that Q-blocks are made by generalizing QI attributes to satisfy 3-anonymity.

Table 2. An example of a 3-anonymous data table after generalization.

Index	Quasi-identifier (QI)		Sensitive
	Work	Country	Disease
1	Private	North	Heart Disease
5	Private	North	Heart Disease
8	Private	North	Heart Disease
2	Government	South	Cancer
3	Government	South	Cancer
9	Government	South	Flu
4	Workclass	North	Flu
6	Workclass	North	Heart Disease
7	Workclass	North	Flu

Therefore, k-anonymity is defined as: Let D denote the original data table and D^* denote a release candidate of D produced by the generalization. Given a set of QI attributes $Q_1,\ldots,Q_d$, release candidate D^* is said to be k-anonymous with respect to $Q_1,\ldots,Q_d$ if each unique tuple in the projection of D^* on $Q_1,\ldots,Q_d$ occurs at least k times.

L-Diversity

While k-anonymity protects against linking attacks, it does not provide sufficient protection for sensitive attributes. This has been recognized by previous studies. The following two attacks are presented to show a homogeneity attack and a background knowledge attack [11].

Homogeneity Attack

In an anonymized table, if a Q-block exists in which all tuples share the same value of sensitive attributes, it will be exposed to a homogeneity attack because an adversary can easily infer an individual's sensitive value by linking an external table.

Background Knowledge Attack

An adversary can infer individuals' sensitive information from an anonymity table using his/her background knowledge. In order to guarantee privacy against such adversaries, Machanavajjhala et al proposed the *l*-diversity principle.

Machanavajjhala et al indicate that *l*-diversity can resist homogeneity and background knowledge attacks [11]. *l*-diversity is defined as: A Q-block is said to have *l*-diversity if it contains at least *l* "well-represented" values for sensitive attribute. A table is said to have *l*-diversity if every Q-block has *l*-diversity. Table 3 is an example of a 3-diverse data table. Machanavajjhala et al provide a number of interpretations of the term "well-represented."

Table 3. An example of a 3-diverse data table.

Index	Quasi-identifier (QI)		Sensitive
	Work	Country	Disease
1	Workclass	America	Heart Disease
3	Workclass	America	Cancer
7	Workclass	America	Flu
2	Workclass	America	Cancer
8	Workclass	America	Heart Disease
9	Workclass	America	Flu
4	Workclass	North	Flu
5	Workclass	North	Heart Disease
6	Workclass	North	Heart Disease

Distinct *l*- diversity

The simplest understanding of "well represented" would be to ensure that there are at least *l* distinct values for the sensitive attribute in each Q-block. It can guarantee that the sensitive value is predicted correctly by the adversary as equation (1), where Q is the number of rows in the Q-block (see Equation (1) in Figure 1). However, distinct *l*-diversity cannot provide a stronger privacy guarantee because when Q-block may have one value that appears much more frequently than other values, an adversary would be able to predict that an entity in the Q-block is most likely to have that value. This motivated the development of the following two stronger notions of *l*-diversity.

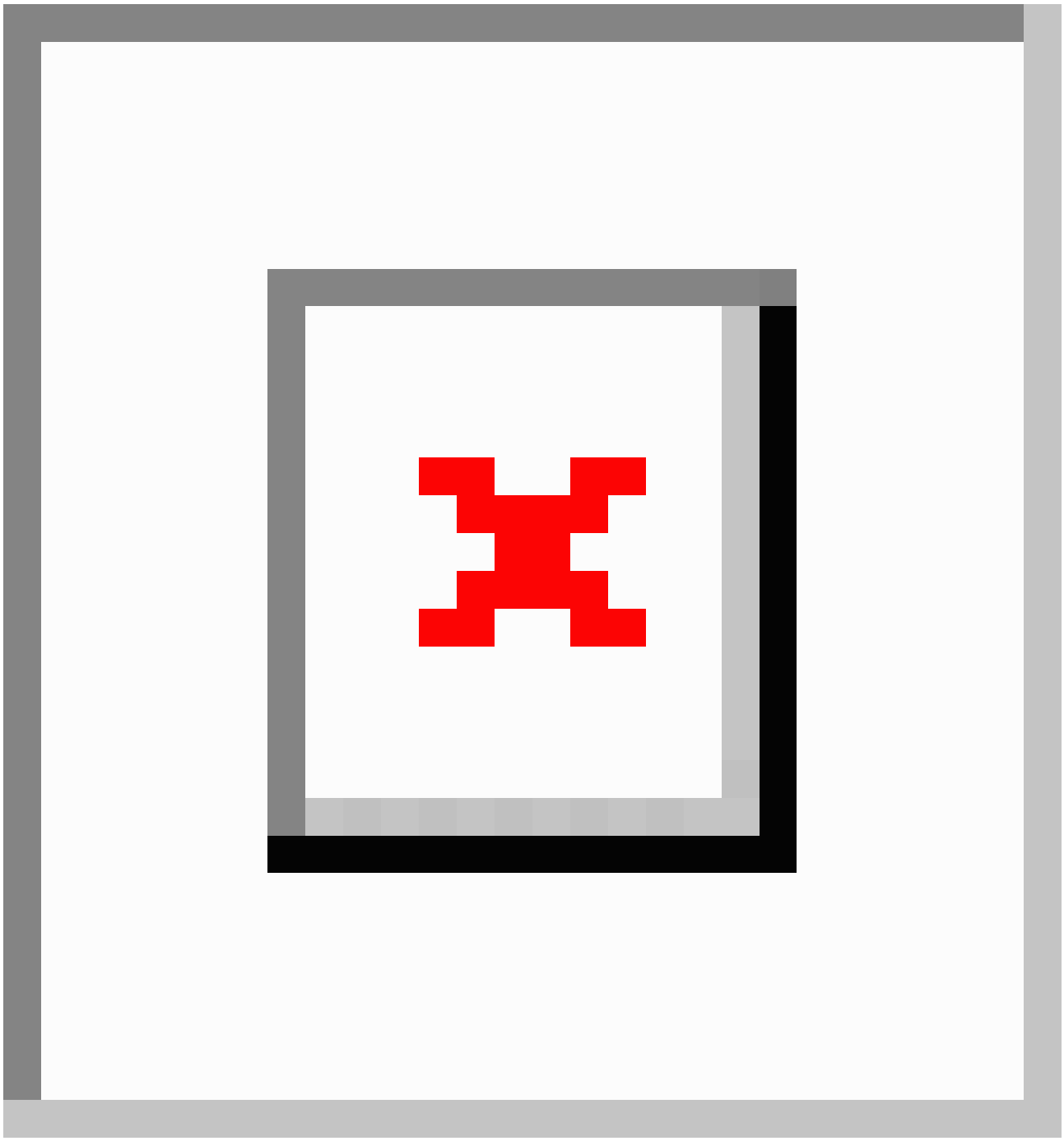
Entropy *l* -diversity

When s is the domain of the sensitive attribute, and $p(Q, s)$ is the fraction of instances in Q (Q-block) that have sensitive value s, Equation (2) represents the entropy of a Q-block (see Figure 1). A table is said to have entropy l-diversity if all Q-blocks satisfy "$Entropy(Q) \geq \log l$". Entropy l-diversity is stronger than distinct l-diversity. In order to have entropy l-diversity for each Q-block, the entropy of the entire table must be at least log l. Sometimes this may be too restrictive, as the entropy of the entire table may be low if a few values are very common. This leads to the following less conservative notion of l-diversity.

Recursive (c , l)-diversity

Recursive (c,l)-diversity ensures that the most frequent value does not appear too frequently, and the less frequent values do not appear too rarely. Let m be the number of values in a Q-block and $r_i (1 \leq i \leq m)$ be the number of times that the i_{th} most frequent sensitive value appears in a Q-block. Then Q-block is said to have recursive (c,l)-diversity if $r_1 < (r_l + r_{l+1} + \ldots + r_m)$. A table is said to have recursive (c,l)-diversity if all of its Q-blocks have recursive (c,l)-diversity.

Figure 1. Equations (1) to (8).



Limitations of Recent Studies

While the l-diversity principle represents an important step beyond k-anonymity for protecting against attribute disclosure, it has several shortcomings. We have already explained that the Q-block is made by generalizing database entries. Generalization of QI attributes leads to an information loss, so minimizing

information loss is a very important issue. However, most recent *l*-diversity studies focus only on diversifying sensitive attribute without accounting for information loss of QI attributes. It means that they consider k-anonymity and *l*-diversity independently.

Li et al proposed the *t*-closeness method, which protects against sensitive attributes disclosure by defining semantic instance among sensitive attributes [12]. This approach requires distance between the distribution of the sensitive attribute in the group and the distribution of the attribute in the whole dataset to be no more than a threshold *t*. However *t*-closeness would greatly damage the data utility when *t* is small because enforcing *t*-closeness destroys the correlations between quasi-identifier attributes and sensitive attributes [17,18]. Other recent studies proposed privacy protection methods, which handle k-anonymity and *l*-diversity [19,20]. These studies proposed an improved algorithm to reduce the complexity or efficient implementation. However, the methods they have proposed improve only the individual performance of k-anonymity and *l*-diversity.

Data utility and sensitivity disclosure have to be considered for actual EHR data release. Therefore, research that covers both characteristics of k-anonymity and *l*-diversity is necessary. As such, we have developed a method that considers both algorithms (k-anonymity and *l*-diversity).

Methods

We have indicated some of the limitations of k-anonymity and *l*-diversity in the previous section. In this paper, we propose a method to make a Q-block that minimizes information loss while achieving diversity of sensitive attributes. For this, we use two measurements: conditional entropy and mutual information. These two measurements are based on entropy characteristics. The use of conditional entropy to obtain minimum information loss has already been studied [13,21]. However, this method cannot guarantee the diversity of sensitive attributes. Therefore, we use conditional entropy as well as mutual information to calculate the distance between instances in order to offer an appropriate Q-block. Mutual information is a quantity that measures a relationship between generalized and sensitive attributes. Therefore, choosing a set that has a small value for mutual information can achieve the required diversity of sensitive attributes.

To calculate the conditional entropy and mutual information, we assume that a dataset holds information on an individual from a population $D = \{D_1,\ldots,D_n\}$. Each individual consists of a collection of QI attributes and sensitive attributes. In this paper, when i is the index of attribute, r is the total number of attributes, and j is the number of possible values, we will treat both attributes and define these as Equation (3) (see Figure 1). If A_i is *work* class, then $A_i = \{$*Private*, *Self-emp-not-inc*,…, *Never-Worked*$\}$.

Figure 2 shows the individual conditional entropies and mutual information. Entropy H equals the negative of the sum of category probabilities times the logarithms of category probabilities, where i is a particular value of attribute. See Equations (4) and (5) in Figure 1.

The value H lies between 0 and $\log_2 I$. It is zero only when the value of one of the p_is is one and all the others zero. Conditional entropy quantifies the remaining entropy of a random variable X, given that the value of another random variable Y is known [22,23]. Where p_{ij} is joint probability distribution, conditional entropy is referred to as the entropy of X conditional on Y (see Equation (6) in Figure 1).

To make a Q-block that satisfies 3-anonymity, we have to generalize the set that contains at least three instances. We chose these instances to calculate the distance between all possible pairs of instances. A small distance value means that they are close to each other. If attribute X in the original database is generalized into Y, then $H(X|Y)$ indicates the information loss by generalization. In order to minimize information loss, we use conditional entropy to calculate the distance.

For example, suppose we generalize Table 1. Assume that the Q-block is built with respect to the first instance. As a first step, we calculate the distance between the first instance and others. Second, we find instances that are close to the first instance using the results of distance. Table 4 shows an example of generalizing between first instance and second instance. In this case, *private* and *state-gov* of the *Work* attribute are generalized into *Workclass*. We calculate the conditional entropy between the original *Work* attribute and generalized *Work* attribute. Next, we perform the same process to the *Country* attribute. The sum of two conditional entropy values is the distance between the first instance and second instance and is expressed as $d_{1,2}$. We calculate distances $d_{1,2} \sim d_{1,9}$, which are all possible pairs of instances and then choose two instances that have minimum values. Generalizing these selected instances can make a Q-block with minimum information loss.

Table 4. Data table showing generalized QI attributes and sensitive attributes for first instance and second instance to explain conditional entropy and mutual information.

Index	Original quasi-identifier		Generalized quasi-identifier		Sensitive
	Work	Country	Work	Country	Disease
1	Private	USA	Workclass	America	Heart Disease
2	State-gov	Mexico	Workclass	America	Cancer
3	Local-gov	Brazil	Local-gov	Brazil	Cancer
4	Federal-gov	USA	Federal-gov	USA	Flu
5	Private	Canada	Private	Canada	Heart Disease
6	Self-emp-not-inc	Canada	Self-emp-not-inc	Canada	Heart Disease
7	Self-emp-inc	USA	Self-emp-inc	USA	Flu
8	Private	USA	Private	USA	Heart Disease
9	State-gov	Mexico	State-gov	Mexico	Flu

However, this method using only conditional entropy cannot prevent homogeneity attacks or background knowledge attacks. Therefore, the proposed method uses mutual information in addition to conditional entropy to achieve diversity of sensitive attributes. Mutual information is a general measure of dependence between two random variables [22,23]. It can be defined as Equation (7) (see Figure 1).

Mutual information is a useful concept for measuring the amount of information shared between a generalized database and sensitive attributes [24,25]. A low value of mutual information indicates that the generalized database and sensitive attributes are almost independent. In order to achieve diversity of a sensitive attribute, we use mutual information to calculate the distance.

We showed an example of calculating information loss by generalization between first instance and second instance using conditional entropy. Table 4 shows an example that calculates mutual information between generalized QI attributes and the sensitive attribute. We calculated the joint probability distribution of QI attributes and the probability distribution of sensitive attributes to achieve mutual information. Mutual information can measure the similarity of the probability distribution between QI attributes and the sensitive attribute. When the first instance and second instance are generalized, their QI values are changed to the same value. In this case, the mutual information {*Heart Disease, Heart Disease*} of the sensitive attribute is larger than {*Heart Disease, Cancer*}. Therefore, to achieve diversity of the sensitive attribute, we made a Q-block that has lower mutual information between the generalized database and the sensitive attribute.

We can now explain the concept in a more detailed manner. Figure 3 shows the set that minimizes conditional entropy between the original database and the generalized database and mutual information between the generalized database and the sensitive attribute. The distance function, defined as Equation (8) (see Figure 1), measures the information loss and diversity. We chose instances that have the smallest value of Equation (8) to make appropriate Q-blocks. The total information loss can be calculated by summing up the loss of all Q-blocks.

Algorithm 1 (see Appendix 1) shows the procedure of calculating distance. Let $S = \{s_i\}_{1 \le i \le N}$ be the set of instances, where N is number of instance. s_{ik} is the k^{th} attribute value of i^{th} instance of S. When i^{th} and j^{th} instances are generalized, the total conditional entropy is the addition of each attributes conditional entropy value. Next the mutual information between QI and sensitivity attribute is calculated. During this step, the combinational values of QI are considered as s single value for mutual information calculation. Total distance between two instances will be the sum of mutual information and total conditional entropy.

We have used a simple clustering method to construct a dataset that satisfies the k-anonymization (see Algorithm 2 in Appendix 1). First randomly select an instance as a seed, and then subsequently select and add k-1 records to build the Q-block. The distance is calculated based on Algorithm 1.

Figure 2. Individual conditional entropies and mutual information for a pair of correlated subsystems.

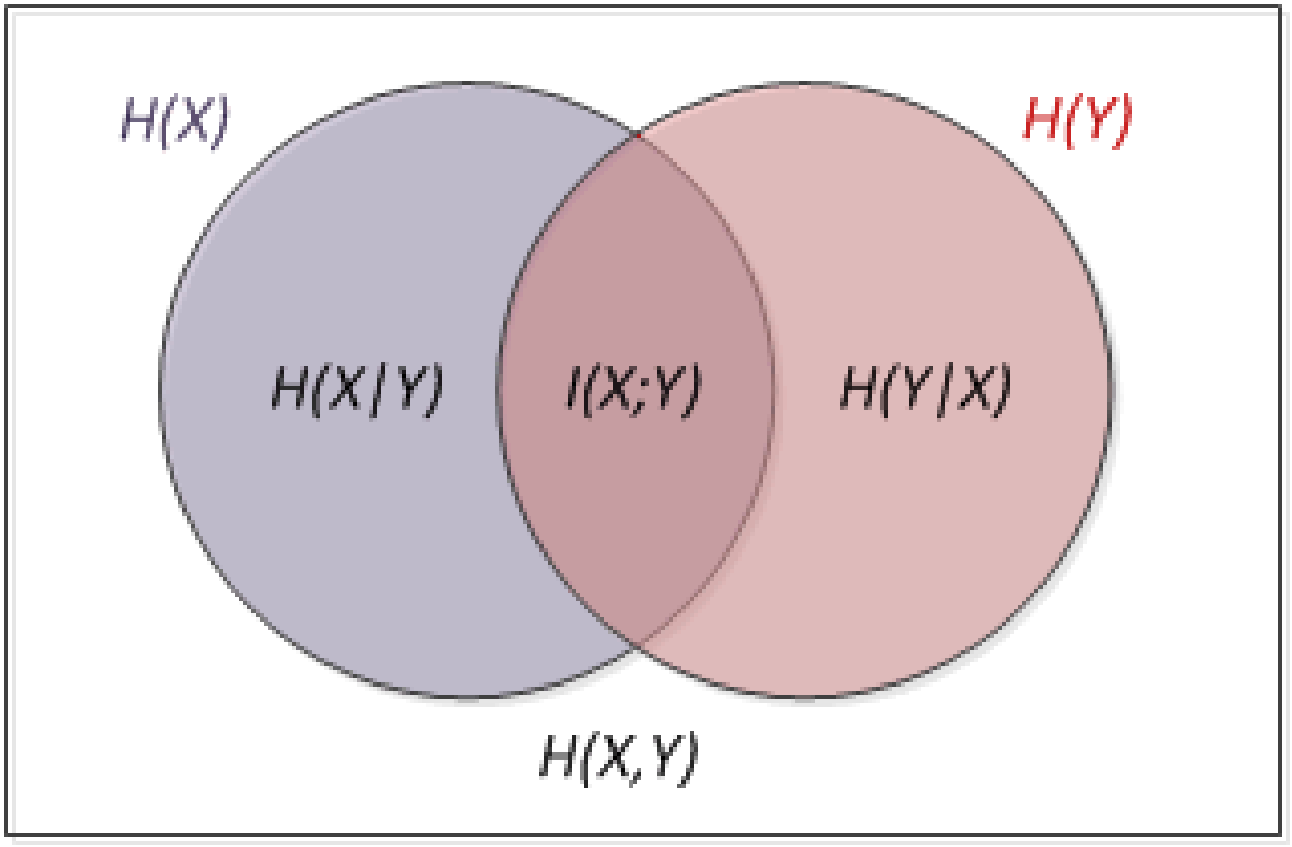
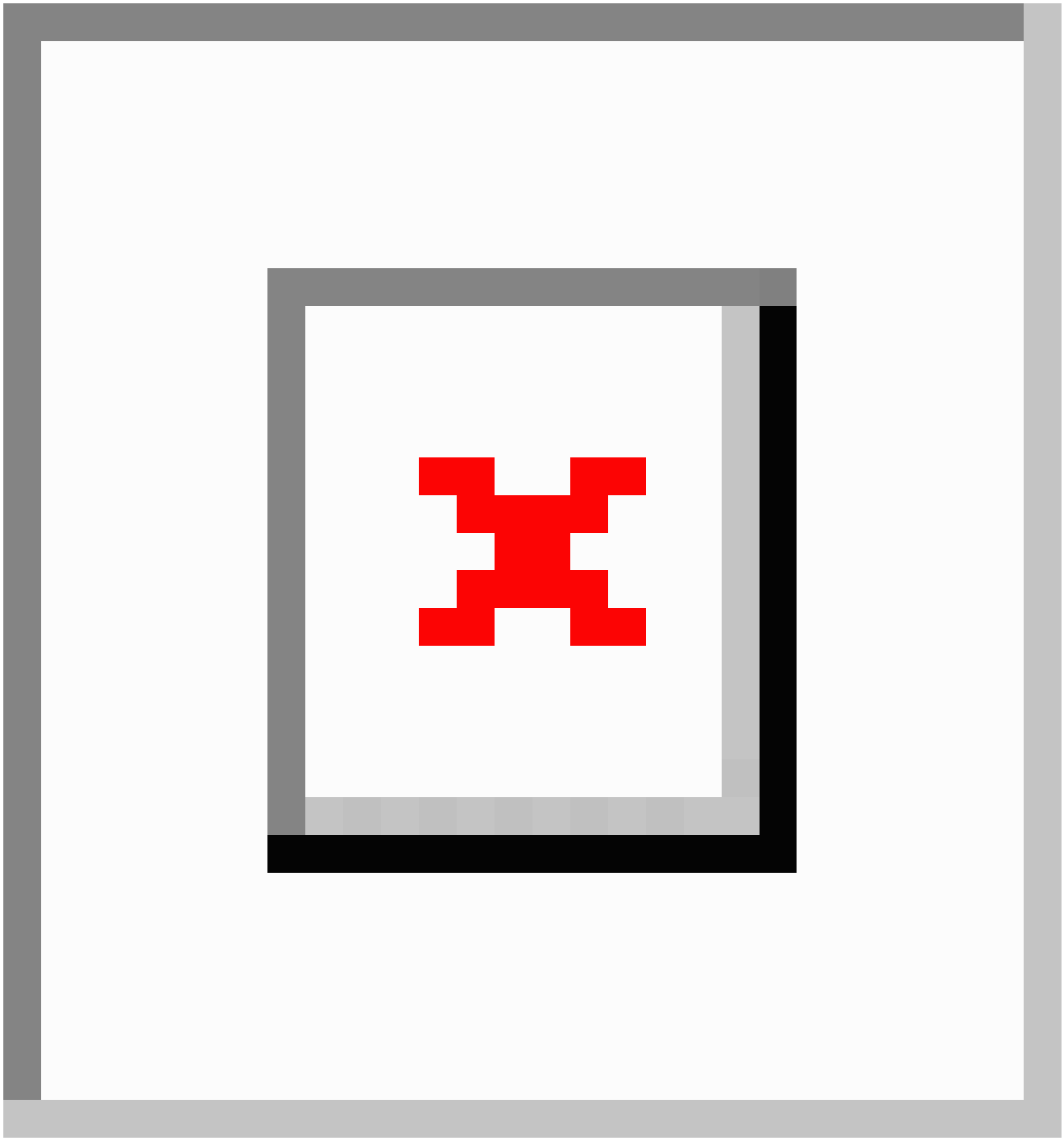


Figure 3. Simplified concept of the proposed method.



Results

We used the adult database from the UCI Machine Learning Repository in our experiments. This database contains census data and has become a commonly used benchmark for k-anonymity. This dataset consists of 15 attributes and 30,162 samples (patients), and we used 9 attributes where numeric attributes are not included. For the adult database, we used Occupation as the sensitive attribute and other attributes as the QI attributes. All methods are implemented in Java and run on a PC with Quad 2.4GHz processor and 4GB RAM under the Windows 7 operating system.

Figure 4 presents the performance of total information loss when the Q-block size is 3. The x-axis is the number of instances and the y-axis is the total information loss. We compared the proposed method with k-anonymity using conditional entropy (CE), entropy *l*-diversity, and *t*-closeness(t = 0.15). Total information loss of CE is decreased, associated with the number of instances. The large number of instances leads to a stochastic reduction in the average value of p_{ij} in Equation (6) (see Figure 1). In addition, more of the same attribute values can be obtained by increasing the number of instances, in which case conditional entropy is zero, so total information loss is not increased. However, total information loss of entropy *l*-diversity, *t*-closeness, and the proposed method is increased in response to the larger number of instances. Even though p_{ij} is reduced

with a large number of instances, entropy *l*-diversity, in particular, generalizes targets in proportion to the number of instances, so information loss is increased.

The proposed method shows some degradation of information loss when compared with CE. Even though the proposed method considers the information loss, it cannot surpass CE because the proposed method uses conditional entropy but also mutual information to make the Q-block. This means that the proposed method considers information loss to a lesser extent than does CE. However, the proposed method is more than four times better than entropy *l*-diversity. It also shows better (nearly three times better) performance compared to the *t*-closeness method ($t = 0.15$).

Figure 5 presents the number of Q-blocks for "$l = 1, 2, 3$". The x-axis is the number of instances, and the y-axis is the number of Q-blocks. We compare the proposed method with CE, entropy *l*-diversity, and *t*-closeness ($t = 0.15$). We have already explained that k-anonymity is susceptible to homogeneity attacks and background knowledge attacks. Assuming that the size of the Q-block is 3, a homogeneity attack will occur when *l* equals 1, and a background knowledge attack will occur when *l* equals 2. Therefore, we can confirm that reducing the number of Q-blocks for "$l = 1, 2$" represents a higher diversity of sensitive attributes. In Figure 5, the proposed method reduces the number of Q-blocks for "$l = 1, 2$" when compared with CE. We also found that the proposed method showed similar performance with *t*-closeness. However, the proposed method is somewhat inferior to entropy *l*-diversity in performance. From these results, we confirmed that the proposed method can reduce information loss while retaining diversity of sensitive attributes.

Figure 6 presents the execution time and compares the proposed method with CE, entropy *l*-diversity, and *t*-closeness ($t = 0.15$). The x-axis is the number of instances, and the y-axis is the execution time. We found that CE and *l*-diversity give almost the same performance, whereas the proposed method and *t*-closeness are slower than the other two methods (ie, CE and entropy *l*-diversity). The reason is that the proposed method calculates mutual information, and calculating the joint probability distribution is quite complex. This shows similar complexity level with KL-divergence calculated by *t*-closeness [12]. Therefore, the proposed method and *t*-closeness share a similar performance. Although our method is slower than others, the overhead is still acceptable in most cases considering its better performance with respect to the information loss and diversity.

Figure 4. Comparison of total information loss with respect to the number of instances.

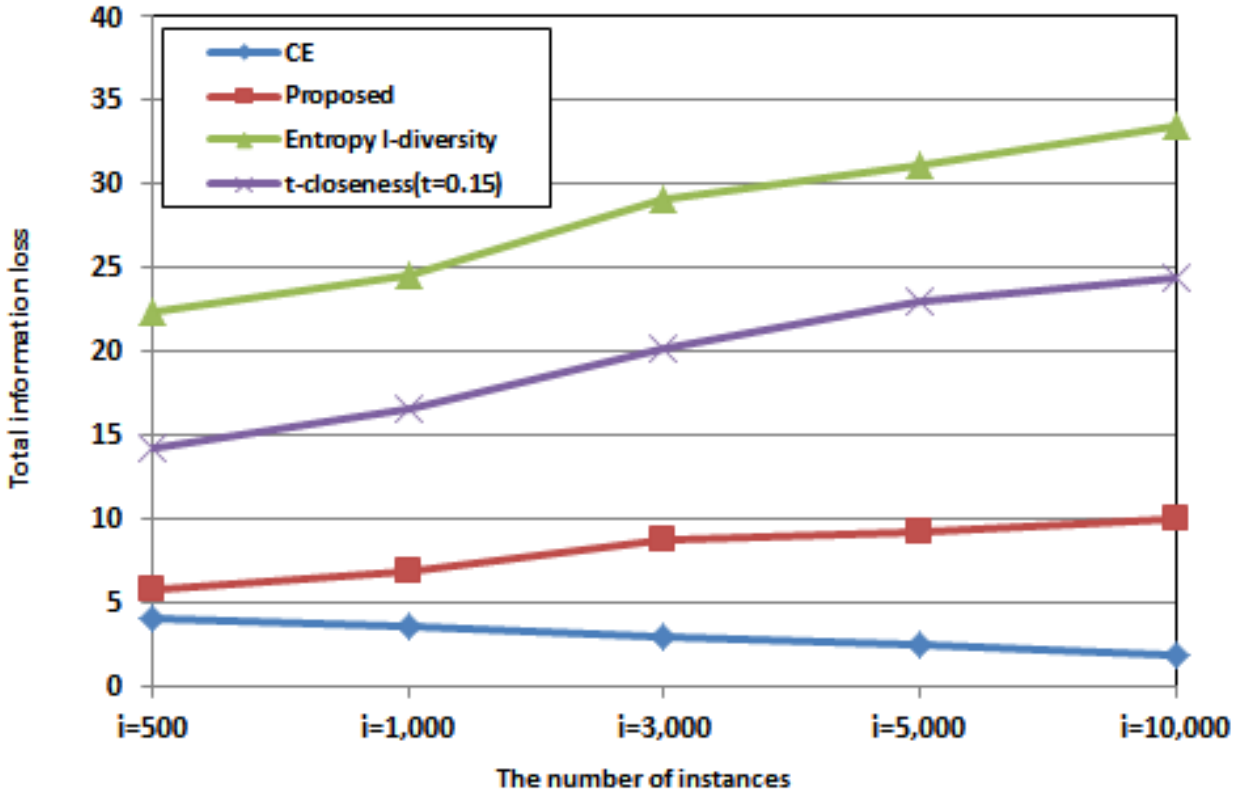


Figure 5. Comparison of the number of Q-blocks, which are l=1 (homogeneity attack), l=2 (background knowledge attack), and l=3 (safe), to measure the diversity (the size of Q-block is set to 3).

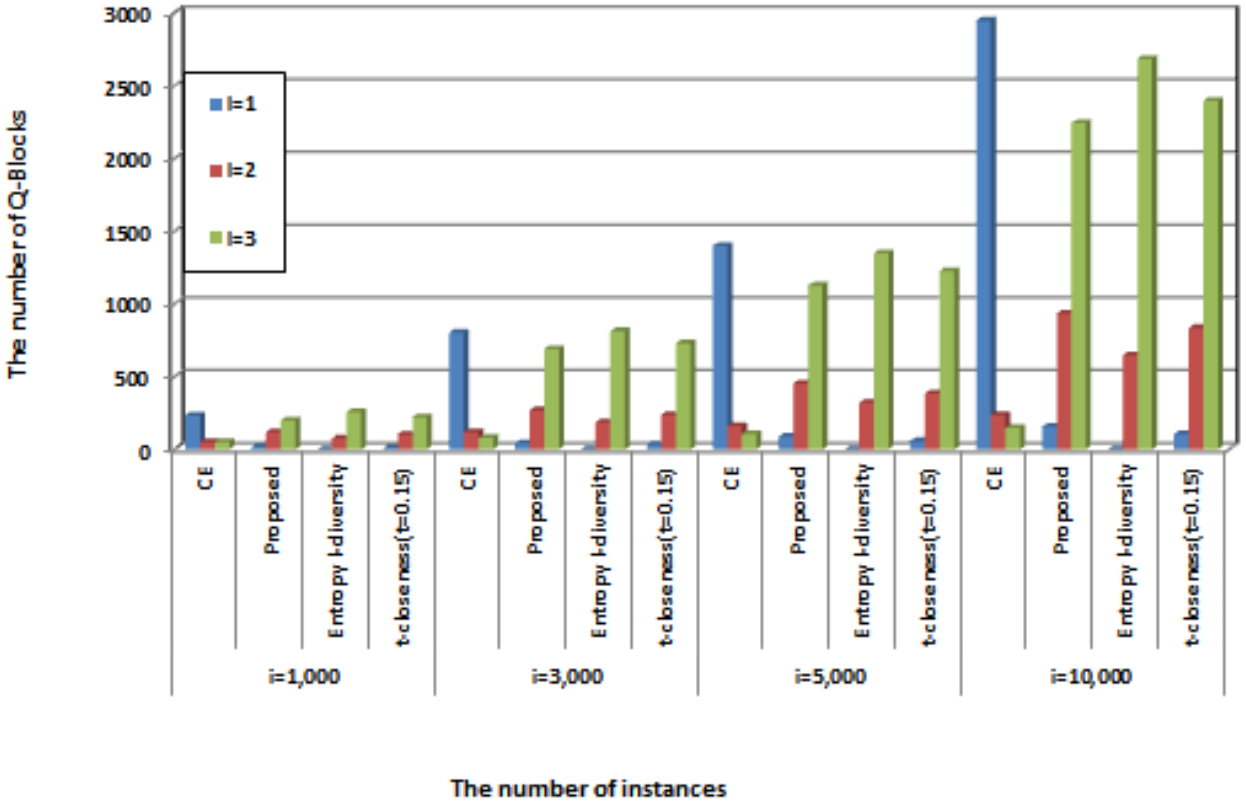
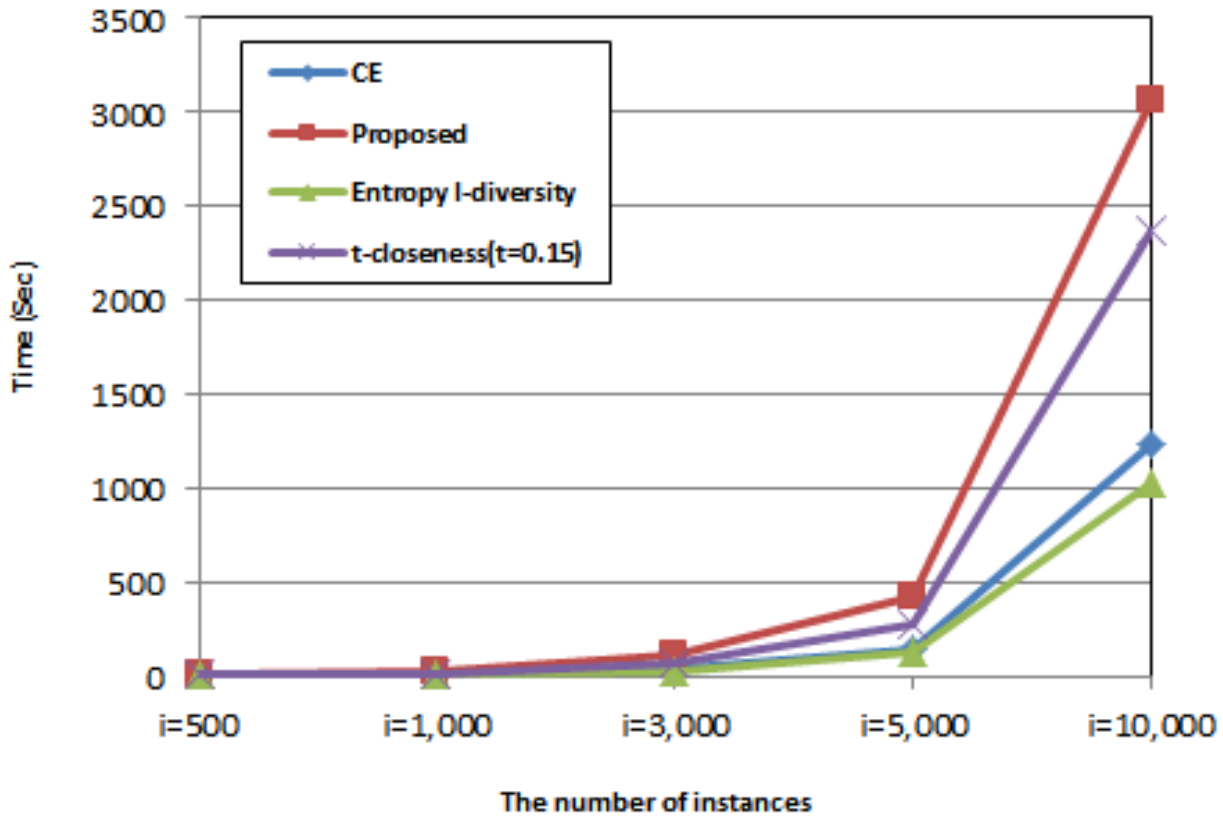


Figure 6. Comparison of execution time with respect to the number of instances.



Discussion

Limitations

We have used mutual information to achieve diversity of sensitive attributes. From the experimental results, we confirmed that the proposed method can reduce the probability of homogeneity and background knowledge attacks. However, there is still room for improvement. The proposed method can substantially increase the diversity, while metrics for calculating the increment of diversity is left for further study. Also, similarity attacks must still be considered. When the sensitive attribute values in a Q-block are distinct but semantically similar, an adversary can learn privacy information. We need to carry out further work to address these problems, and then we will be able to provide even better improvements in privacy protection in EHR applications.

Conclusions

This paper proposes a new privacy protection method that uses conditional entropy and mutual information. This method not only minimizes information loss but also achieves diversity of the sensitive attribute. This leads to increased data usability and prevents homogeneity attacks. This method was experimentally verified using an adult database from the UCI Machine Learning Repository. We compared the proposed method with previous *l*-diversity methods (ie, entropy *l*-diversity and *t*-closeness) to show that our method enables a reduction in information loss. It also can guarantee diversity of sensitive attributes when compared with CE. The method is viable and should be of interest for further utilization of privacy protection in various EHR data applications.

Acknowledgments

The study was supported by grants of the Korea Health Technology R&D Project, Ministry of Health and Welfare, Republic of Korea (A112022).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Algorithms 1 and 2.

[PDF File (Adobe PDF File), 267KB - ijmr_v1i2e14_app1.pdf]

References

1. Fienberg SE. Sharing statistical data in the biomedical and health sciences: ethical, institutional, legal, and professional dimensions. Annu Rev Public Health 1994;15:1-18. [doi: 10.1146/annurev.pu.15.050194.000245] [Medline: 8054076]
2. Hensler P. Electronic Medical Records (EMR).: Clendening, Johnson & Bohrer; 2012. URL: http://www.google.ca/url?sa=t&rct=j&q=p%20bohrer%20electronic%20medical%20records&source=web&cd=1&ved=0CDgQFjAA&url=http%3A%2F%2Flawcjb.com%2Fpublications%2Fwp-content%2Fuploads%2F2011%2F10%2FPH_Electronic-Health-Records-20111005.pptx&ei=ZTOVUO6pCMSO2AWI_oDYAw&usg=AFQjCNEeRpP5jZBXj4_2-xePAWiLj1NDdw [accessed 2012-11-02]
3. Kirwan JR. Making original data from clinical studies available for alternative analysis. J Rheumatol 1997 May;24(5):822-825. [Medline: 9150066]
4. Hrynaszkiewicz I, Altman DG. Towards agreement on best practice for publishing raw clinical trial data. Trials 2009;10:17 [FREE Full text] [doi: 10.1186/1745-6215-10-17] [Medline: 19296844]
5. GPO US. 2008. Part 46-Projection of human subjects URL: http://www.gpo.gov/fdsys/pkg/CFR-2008-title45-vol1/content-detail.html [accessed 2012-11-03] [WebCite Cache ID 6BtwCNYW2]
6. GPO US. 2008. Part 45-Security and Privacy URL: http://www.gpo.gov/fdsys/pkg/CFR-2008-title45-vol1/content-detail.html [accessed 2012-11-03] [WebCite Cache ID 6BtwDgkew]
7. Samarati P, Sweeney L. Protecting privacy when disclosing information: k-anonymity and its enforcement through generalization and suppression. Technical Report SRI-CSL-98-04. Menlo Park: SRI International; 1998. URL: http://epic.org/privacy/reidentification/Samarati_Sweeney_paper.pdf [accessed 2012-04-21] [WebCite Cache ID 675PWU0qp]
8. Sweeney L. Achieving k-anonymity privacy protection using generalization and suppression. International Journal of Uncertainty Fuzziness and Knowledge-Based Systems(5):571-88 2002;10(5):571-588. [doi: 10.1142/S021848850200165X]
9. Samarati P, Sweeney L. Generalizing data to provide anonymity when disclosing information. 1998 Jun 01 Presented at: PODS '98 Proceedings of the 17th ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems; 1998; Seattle, WA p. 188.
10. Sweeney L. k-anonymity: A model for protecting privacy. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems 2002;10(5):557-570. [doi: 10.1142/S0218488502001648]
11. Machanavajjhala A, Kifer D, Gehrke J, Venkitasubramaniam M. l-diversity: Privacy beyond k-anonymity. ACM Transactions on Knowledge Discovery from Data 2007;1(1):3 [FREE Full text] [doi: 10.1145/1217299.1217302]

12. Li N, Li T, Venkatasubramanian S. t-closeness: Privacy beyond k-anonymity and l-diversity. 2007 Presented at: Proceedings of the 23rd International Conference on Data Engineering: IEEE; 2007; Istanbul p. 106-115 URL: http://www.cs.purdue.edu/homes/ninghui/papers/t_closeness_icde07.pdf
13. Gionis A, Tassa T. k-Anonymization with minimal loss of information. IEEE Transactions on Knowledge and Data Engineering 2009;21(2):206-219. [doi: 10.1109/TKDE.2008.129]
14. Golle P. Revisiting the uniqueness of simple demographics in the US population. ACM 2006:77-80. [doi: 10.1145/1179601.1179615]
15. Sweeney L. Uniqueness of simple demographics in the US population. Carnegie Mellon University: Serial (Book, Monograph); 2000.
16. El Emam K, Jabbouri S, Sams S, Drouet Y, Power M. Evaluating common de-identification heuristics for personal health information. J Med Internet Res 2006;8(4):e28 [FREE Full text] [doi: 10.2196/jmir.8.4.e28] [Medline: 17213047]
17. Sun X. Towards Privacy Preserving Data Publishing URL: http://www.mcs.vuw.ac.nz/twiki/pub/Events/ACDC2009/XiaoxunSun.pdf [accessed 2012-04-21] [WebCite Cache ID 675StTSzv]
18. Domingo-Ferrer J, Torra V. A critique of k-anonymity and some of its enhancements. In: Proceedings of the 2008 Third International Conference on Availability, Reliability and Security. 2008 Presented at: 3rd International Conference on Availability, Reliability and Security; 2008; Barcelona p. 990-993. [doi: 10.1109/ARES.2008.97]
19. Ghinita G, Karra P, Kalnis P, Mamoulis N. Fast data anonymization with low information loss. In: VLDB`07 Proceedings of the 33rd International Conference on Very Large Data Bases. 2007 Presented at: 33rd International Conference on Very Large Data Bases; September 23-28, 2007; Vienna p. 758-769.
20. Xiao X, Tao Y. Anatomy: simple and effective privacy preservation. In: VLDB '06 Proceedings of the 32nd International Conference on Very Large Data Bases. 2006 Presented at: 32nd International Conference on Very Large Data Bases; September 12-15, 2006; Seoul p. 139-150.
21. Bezzi M. An entropy based method for measuring anonymity. In: Security and Privacy in Communications Networks and the Workshops. 2007 Presented at: 3rd International Conference on Security and Privacy in Communications Networks and the Workshops; September 18-20, 2007; Nice p. 28-32. [doi: 10.1109/SECCOM.2007.4550303]
22. Fano R. Transmission of Information: A Statistical Theory of Communications. Cambridge, MA: MIT Press; 1961.
23. Papoulis A. Probability, Random Variables, and Stochastic Processes. In: Probability, random variables, and stochastic processes. New York: McGraw-Hill; 1991.
24. Bezzi M. Protecting privacy of sensitive value distributions in data release. In: Proceedings of the 6th Workshop on Security and Trust Management (STM). 2010 Presented at: 6th Workshop on Security and Trust Management; September 23-24, 2010; Athens p. 255-270. [doi: 10.1007/978-3-642-22444-7_17]
25. Goldberger J, Tassa T. Efficient anonymizations with enhanced utility. Transactions on Data Privacy 2010;3:149-175.

Edited by K El Emam; submitted 27.04.12; peer-reviewed by B Malin, L Xiong; comments to author 06.07.12; revised version received 02.08.12; accepted 21.09.12; published 13.11.12.

Please cite as:
Yoo S, Shin M, Lee D
An Approach to Reducing Information Loss and Achieving Diversity of Sensitive Attributes in k-anonymity Methods
Interact J Med Res 2012;1(2):e14
URL: http://www.i-jmr.org/2012/2/e14/
doi:10.2196/ijmr.2140
PMID:23612074

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org