Interactive Journal of Medical Research

Contents

Improving Patient Preference Elicitation by Applying Concepts From the Consumer Research Field: Narrative Literature Review

Niki Ver Donck[1,2,3], MSc; Geert Vander Stichele[2,3], MSc, PharmD; Isabelle Huys[1], MSc, PhD

[1]Department of Pharmaceutical and Pharmacological Sciences, University of Leuven, Leuven, Belgium

[2]Health Economics Consultancy, ISMS, Turnhout, Belgium

[3]Digital Health Solution Development, MindBytes, Merksplas, Belgium

Corresponding Author:
Niki Ver Donck, MSc
Health Economics Consultancy
ISMS
Campus Blairon 477
Turnhout, 2300
Belgium
Phone: 32 14 701 930
Email: niki.verdonck@student.kuleuven.be

Abstract

Background: Although preference research finds its origins in consumer research, preference elicitation methods have increasingly attracted attention in different decision-making contexts in health care. Simulating real-life decision making is believed to be important during consumer preference elicitation.

Objective: The aims of this study were to compare the process of decision making between patients and consumers and to identify methods from the consumer research field that could be applied in patient preference elicitation.

Methods: A narrative literature review was performed to identify preference elicitation concepts from a consumer context that could offer improvements in health care.

Results: The process of decision making between patients and consumers was highly comparable. The following five concepts from the consumer research field that could effectively simulate a real-life decision-making process for applications in health care were identified: simulating alternatives, self-reflection, feedback-driven exploration, separated (adaptive) dual response, and arranging profiles in blocks.

Conclusions: Owing to similarities in the decision-making process, patients could be considered as a subgroup of consumers, suggesting that preference elicitation concepts from the consumer field may be relevant in health care. Five concepts that help to simulate real-life decision making have the potential to improve patient preference elicitation. However, the extent to which real decision-making contexts can be mimicked in health care remains unknown.

(Interact J Med Res 2020;9(1):e13684) doi:10.2196/13684

Introduction

Background

During the last decade, there has been growing interest in patient perspectives and experiences in health care decision making [1,2]. The idea of patient involvement has become increasingly accepted, as patients are in a unique position to share their day-to-day experiences in dealing with an illness and its treatment. Information about patients' perceptions and tradeoffs has the potential to inform decision making on different levels.

As patients are the end users of medical products, they are the utmost important stakeholder in the context of patient-centered health care and deserve to be involved in medical decision making [3,4]. At the individual level, patients can find themselves in a situation where multiple treatment options exist, without having one option that is clearly superior compared to the others [5]. In some cases, clinical evidence is scarce, resulting in high levels of uncertainty about treatment benefits. In other cases, there is abundant information on the benefits and risks of available options, but patients' views on the

XSL•FO
RenderX

desirability of these outcomes vary greatly, resulting in different opinions of "the best" option [6]. Patients should receive decision support when making these decisions, which are usually referred to as "preference-sensitive decisions" [5,6]. The treating physician can provide decision support to make an informed preference-based choice in the context of shared decision making (SDM) [7]. In this particular context, the process of forming preferences is often referred to as a "value clarification," which is followed by preference elicitation [6,8]; the combination of these two aspects is called a "value clarification exercise" (VCE) [8]. At the meta level, patient preference data can provide additional information for decisions on drug development, regulatory assessment, or reimbursement [9-13]. Patient preference information is defined by the Food and Drug Administration (FDA) as "Qualitative or quantitative assessments of the relative desirability or acceptability to patients of specified alternatives or choices among outcomes or other attributes that differ among alternative health interventions" [5]. From a societal perspective, the inclusion of the patient opinion could improve the transparency and acceptability of regulatory or reimbursement decisions [2,14]. Finally, the quality of decisions at both the individual and societal levels might increase when decision making is aligned with the patients' unmet needs [3].

Patient involvement can be realized in a variety of ways: by asking for input from patients via unstructured methods (eg, testimonials, comments in correspondence) or via structured methods (eg, conducting surveys, collecting patient-reported outcomes, or revealing patient preferences) [5,10]. As part of a structured process to reveal preferences, both qualitative and quantitative preference measurement methods can be used.

Experience of Preference Elicitation in Consumer Research

Quantitative methods for patient preference elicitation include discrete choice experiments (DCEs)/conjoint analysis (CA) or best-worst scaling [11,15]. The DCE technique was introduced by Louviere and Woodworth [16] in the context of marketing to forecast consumer choices. In 1990, CA and DCEs entered the health care setting and have since been increasingly used for patient preference elicitation [17]. Respondents are asked to choose between two or more alternatives, which are usually profiles consisting of different attributes (including product characteristics such as efficacy, adverse events, and mode of administration) and corresponding attribute levels (eg, oral, injection, and inhalation). By analyzing these results, researchers can derive the underlying utility of particular attributes or profiles [17-19]. Despite the application of DCEs in health care for several decades, the resulting data have not yet been systematically applied to societal decision making, and some uncertainties remain about the utility or validity of DCE results in particular decision-making contexts [20,21]. At present, consensus is lacking on how patient preferences can be optimally measured and incorporated into different health care community decision-making processes [2,22].

Since preference research has been conducted for decades in the context of consumers, experiences from this field might further inspire patient preference research [23,24]. Moreover, several innovative approaches to optimize preference elicitation (CA or other techniques) have been explored in the field of consumer research. Indeed, multiple industries offering innovative durable goods rely on preference elicitation methods to guide the development of new products [25]. However, the main difficulty in measuring consumer preferences for new products is the lack of knowledge and experience of respondents with the new product [26]. As these products typically do not yet exist, consumers have no basic understanding about how to assess the importance of new favorable and unfavorable characteristics or how to assess the tradeoffs between these characteristics [26,27]. Examples of such products are personal computers, smartphones, and electric cars [27,28].

Lack of understanding of the basic characteristics of new products resembles a major issue in patient preference elicitation. Considering that almost one in two Europeans have limited health literacy [29], weighing potential risks and benefits could therefore be a very difficult task for laypeople. This poses a challenge, especially in patient preference research, given the association of worse health states with lower levels of health literacy [29]. Furthermore, patients' medical states might influence their ability to understand the information and engage in a preference elicitation experiment.

Applying Consumer Research Experience to Inform Health Care Preference Studies

According to Louviere [26], the external validity of DCEs depends on the extent to which all key aspects of a real decision are simulated. Preference elicitation experiments that most closely resemble real choice situations (including framing of situations, relevant contexts, and consequences) should be able to provide real-life results. For this reason, simulations for informational purposes were introduced in consumer research many years ago so that all aspects, ranging from consumer reports, advertising, or even the whole store environment, can be simulated to resemble real-life decision processes as closely as possible [25,28].

Furthermore, when consumers need to construct their preferences while acquiring information, the tradeoffs they consider might be unstable and depend on context effects. Therefore, the results may not reflect true preferences [26,27]. Urban and colleagues offered a method to deal with forecasting problems with new products that they termed "information acceleration" [25,30]. Louviere [26] clearly described the use of information acceleration methodology as follows: "Acceleration of Information Methods rely on multimedia and other technologies to simulate the processes by which individuals become aware of new technologies/products, search for and acquire information about benefits and/or problem solutions, decide whether to consider them and whether they can take advantage of what they offer, decide if they want to buy a product now available, or wait to see how the product market develops and evolves over time."

In other words, when designing an experiment to elicit patient preferences, patients need to experience the same process as they would in real life. Their lack of knowledge or experience can be overcome by providing the necessary information in a natural way and showing them the results of various options.

Simple pictures or videos can be used; however, more interactive simulations allow for more user involvement while better stimulating learning and knowledge retention [31]. Further, Hoeffler [27] stated that consumers who are forced to construct their preferences during an experiment may be unable to provide enduring preferences. The need for deeper consideration of the decision problem is a natural process, which may cause preferences to change over time [6].

The decision process of consumers in the context of a tradeoff situation consists of the following stages: becoming aware of a specific need or a new product, deciding what information to acquire and how to acquire it, deciding which alternatives are available to attain the objectives, forming a utility function or decision rule, and ultimately deciding whether or not to purchase the product (depending on budget or other constraints). Finally, if they decide to purchase, consumers sometimes need to choose which option(s) to purchase [26,32]. Acquiring the right information and learning the different advantages and disadvantages of every option in order to make tradeoffs represents an important step of this process. The situation of naïve consumers might be comparable to that of patients being faced with certain treatment options or a specific disease for the first time. As with consumers, patients need to acquire and process information at a fast pace when confronted with a new product or treatment. CA techniques are well suited to analyze decision making in both cases, as they can either simulate already available alternatives (eg, to compare different therapies available to patients) or elicit preferences for goods that do not yet exist (eg, comparing therapies in the drug pipeline or before market authorization has been obtained). In both cases, using methods that simulate real-life choice situations, such as information acceleration, could potentially be useful in health care. However, a clear comparison between the decision-making process of consumers and patients is lacking, impacting the potential to transfer learnings from consumer to health research situations.

To fill this gap, the aim of this study was to compare the process of decision making between consumers and patients. Furthermore, the goal was to identify consumer research methods or concepts that may improve patient preference elicitation by simulating real-life decisions. Based on this analysis, the applicability of the identified methods or concepts in health care are assessed.

Methods

Comparative Description of the Decision-Making Process for Patients and Consumers

The decision-making process of consumers was compared to that of patients. First, the market evolution stages described by Louviere [26] were translated into analog examples for patients engaging in decision making in one of two possible contexts. On the one hand, the context of individual patients engaging in SDM was considered; on the other hand, gathering preference data from a group of patients to inform development, regulatory, or reimbursement decisions was evaluated [7,33].

Literature Review of Innovative Preference Elicitation Concepts in the Consumer Research Field

A literature search was conducted in the Scopus database to identify innovative concepts from the consumer research field that improve preference elicitation by simulating real-life decisions. Three key terms (Table 1) describing preference elicitation methods that resemble real-life decisions such as DCEs/CA were combined with several terms describing innovation, information methods, and the field of consumer research. Every combination was searched for independently and duplicates were removed during the first step of the process. Papers with a publication date >5 years old (ie, published before 2012) were excluded, as older ideas may have already been applied in the health care context. Finally, only articles in English were included. All identified papers were screened for exclusion based on the title. The exclusion criteria were the following: studies performed in a health care setting (as these papers describe techniques that have already been implemented in health care) and studies without sufficient description of the performed method or describing an actual stated preference experiment. In cases of doubt, papers were retained for a second selection stage. In this second stage, abstracts were reviewed for exclusion based on the above-described and two additional exclusion criteria: describing standard DCEs without any new elements (as described by the current standards for patient preference elicitation) and focusing solely on willingness to pay. The remaining articles were retrieved in full-text form and reviewed in a two-step process by the authors. In the first step, each concept was critically evaluated with respect to its capacity to simulate real-life decisions by one author (NVD). In the second step, another author (GVS) independently reviewed this analysis. Differences were resolved by discussion and, when no consensus could be reached, ties were settled by the third author (IH).

XSL•FO

RenderX

Table 1. Search strategy.

Key search term[a]	Combined with (AND)
Preference elicitation	Consumer – Innovative – Scenario based – Simulation – Virtual Reality – Simulation game – Market research – DCE OR conjoint analysis
DCE[b] or CA[c]	Innovative – Scenario based – Virtual Reality – Simulation game – Market research
measuring preferences OR measure preferences OR preference measurement	Consumer – Innovative – Scenario based – Simulation – Virtual Reality – Simulation game – Market research

[a]The key terms were combined using "AND" with each of the individual terms of column 2 in the same row.

[b]DCE: discrete choice experiment.

[c]CA: conjoint analysis.

Assessing the Applicability of Innovative Elicitation Concepts for Patient Preference Elicitation

The current standards to conduct CA or DCEs in health care were reviewed based on leading guidelines in the field issued by the International Society for Pharmacoeconomics and Outcomes Research (ISPOR), the US FDA, and the Medical Device Innovation Consortium (MDIC) [5,15,19]. These guidelines served as a baseline to assess the applicability of the identified methods and concepts from the consumer research field in a health care setting. For the applicability assessment, one author (NVD) evaluated each preference elicitation concept against every topic of the three guidelines by defining each concept as relevant or not relevant. The resulting findings were then reviewed by a second author (GVS). In case no consensus could be reached, ties were settled by the third author (IH). For every concept identified, the complementarity to current standards and the rationale for implementation were considered.

Results

Comparing the Decision-Making Process for Patients and Consumers

Table 2 presents the health care analogy in both individual and group decision-making contexts, alongside the steps defined in the consumer context [26].

Table 2. Different steps of a decision process: health care analogy for the different market evolution stages.

Market evolution stage; Consumer context [26]	Health care analogy	
	Individual context	Group context
Becoming aware of a need Becoming aware of a product	Receiving a diagnosis and becoming aware of (possible) therapies	The experiment is described: patients become aware of different alternatives (therapies)
Deciding what information to acquire and how to acquire it	Deciding what information (on possible treatments) to acquire and how to acquire it, deciding who (eg, family members, caretakers) needs to be involved in the decision-making process	Deciding what information to use that has been made available
Forming decision rules: deciding whether and which options to consider	Forming decision rules: deciding whether and which treatment options to consider	Forming decision rules: deciding whether and which treatment options to consider
Deciding whether to choose now, delay, or never choose	Deciding whether to choose a possible treatment, choose no treatment (eg, watchful waiting), choose to delay treatment, or choose not to be involved in the decision process	Deciding whether to choose a possible treatment or choose no treatment (eg, watchful waiting)
If choosing now, deciding which option to choose	If choosing now, deciding which treatment option (including the option of watchful waiting) to choose	If choosing now, deciding which treatment option (including the option of watchful waiting) to choose

Concepts from Consumer Research Methods

Article Selection and Retrieval

A total of 135 papers were identified using the described search strategy. After selection of titles, 40 papers remained and were screened further by abstract review using the aforementioned criteria. The full text of the resulting 12 papers was analyzed. Five concepts were judged to be potentially interesting for health care and are discussed below. Reasons for excluding the other seven papers were as follows. One paper was excluded as there was no description of a preference experiment, and another paper was judged not to present any innovative ideas, as these turned out to be already included in standard software [34,35]. Further, one paper focused on forecasting decision behavior instead of quantitative preference measurement, and another discussed a compositional approach to evaluate the attributes one by one, which is not complementary with the concept of real-life decision making [36,37]. The concepts of three papers were not applicable to health care: one method could only be applied on very similar products (in the example, different movies were used, whereas the attributes of health care options usually differ greatly); one method presented a framework

XSL•FO
RenderX

consisting of 39 engineering parameters that could not be easily translated to health care equivalents; and the last method was particularly useful for products with 70-100 attributes, whereas in health care typically 3-7 attributes are used [38-40]. The five concepts that are potentially interesting in health care are discussed below in turn.

Concept 1: Simulating Alternatives

By visualizing alternative land use scenarios, Vignola et al [41] provided a useful method to clarify different options and explore collaboration among stakeholders. The method promotes discussions between stakeholders by presenting the pros and cons of different alternatives and accounting for uncertainties. The scenarios describe possible consequences of different courses of action to improve users' understanding of causal processes associated with every decision. Synthesized images of land use patterns and their consequences on a given landscape are accompanied by a stylized narrative, explaining the key changes depending on the context. Using land development scenarios to represent possibilities in the future has been suggested as a mental exercise to improve planning [42]. Scenario use helps respondents to understand different alternatives and their consequences by improving the cognitive processes in which people collect and combine information and draw inferences [42]. Furthermore, it is recommended to involve all stakeholders as much as possible during the scenario creation phase through interviews, focus groups, and follow-up discussions to refine every aspect [41].

Concept 2: Self Reflection

Hauser et al [43] stated that consumers only learn their preferences as they make realistic decisions [43]. To simulate a realistic decision-making process, people need time to self-reflect upon their options. Without self-reflection, preference elicitation methods might not measure enduring (true) preferences, which is in line with Hoeffler's [27] findings on preferences for new products. In the study, respondents completed three tasks. First, they formed consideration sets of 30 realistic profiles chosen randomly from all available profiles, which means they had to decide whether they would consider buying the product or not for each profile. Next, they performed a structured preference-articulation procedure (Casemap) by selecting the best and worst level per attribute set. The final task was to state their consideration rules in an unstructured email to a friend. One week later, the respondents again formed consideration sets from a random set of 30 profiles. The predictive ability of the articulated preferences was measured with the relative Kullback-Leibler divergence and the predictions were compared with the consideration-set decisions 1 week later. The authors found that self-reflection was facilitated either by completing the 30-profile consideration set or a highly structured Casemap task (as a best-worst exercise). Self-reflection improved a respondent's capability to articulate preferences that predict consideration sets 1 week later [43]. Finally, the authors suggested that if consumers are asked to articulate their preferences before self-reflection, this articulation would interfere with their abilities to articulate preferences even after they have had a chance to self-reflect [43].

Concept 3: Separated (Adaptive) Dual Response

Some preference elicitation methods such as DCEs might encounter problems when "opt-out" options are provided, with respect to both context effects (ie, when a respondent chooses the opt-out option for a reason other than the lack of useful alternative products) and extreme response behavior (ie, respondents will always or never choose the opt-out option under some conditions). Schlereth et al [44] introduced the concepts of separated dual response (SDR) and separated adaptive dual response (SADR) to counter these problems. SDR implies separating forced- and free-choice questions, resulting in the respondents first choosing between two alternatives (forced choice) and then choosing whether or not they actually want the chosen option or would like to opt-out as a second step (free choice). This will overcome the context effects created by dominant alternatives (which decreases the likelihood of selecting the opt-out option) or the existence of very similar alternatives (not choosing is an "easy way out" in this case). SDR also eliminates extreme response behavior since the respondents do not have the opportunity to always or never go for the opt-out option. However, the authors noted that this method might introduce a new context effect of choice deferral, resulting in the respondents more frequently choosing the no-purchase option. They suggested solving this problem by separating the questions in time; that is, asking all forced-choice questions first and all free-choice questions later. SADR contains an extra adaptive mechanism that selects fewer, but more informative, free-choice questions.

Concept 4: Feedback-Driven Exploration

Boesch et al [4] proposed the implementation of feedback-driven exploration techniques to improve the validity and reliability when developing a stated-preference experiment. This involves implementing continuous feedback between researchers, respondents, and all other stakeholders throughout the process. The authors formulated the following steps to be included in the research design [45]:

i. Shape guiding research questions, concepts, theories, hypotheses.
ii. Collect and process data.
iii. Interpret and reflect on data (researcher, possibly with data providers).
iv. Report tentative research findings to data providers (e.g. survey respondents, interview participants) and broadly review, discuss and explore results with research stakeholders to arrive at overall conclusions,
v. Intermediate or preliminary results may indicate a need of getting back to earlier phases of the research process, or even of adjusting and starting the process anew.

The authors suggested that an iterative process (going through the different steps multiple times) might be necessary depending on the research question. Three aspects of a stated-preference experiment are specifically mentioned that may benefit from this approach. First, the validity and reliability of the results can be improved, which is particularly important when dealing with research questions for which no real-life data are available to validate the results. Second, the systematic approach of an overall framework will harmonize all of the different steps

required to conduct a preference elicitation experiment. Third, all relevant stakeholders can be involved in the process, whereas experts from outside academia are often overlooked during the research and development phase [45].

Concept 5: Arranging Profiles in Blocks to Improve Performance

Adaptive CA consists of two consequential approaches: a composition and a decomposition method. First, respondents evaluate independent attributes (composition method), and then the most preferred attributes are combined in profiles and presented in blocks of two randomly arranged profiles (decomposition method) [46]. This approach is particularly useful when tradeoffs need to be made between a high number of attributes in a user-tailored process. Huertas-Garcia et al [46] suggested a design strategy to improve the performance of the decomposition methodology in adaptive CA by arranging profiles, manually or automated by a computer algorithm, into subsets of two profiles. With this strategy, the respondents are asked to evaluate only a subset of profiles rather than the whole choice set. Dividing the profiles in different blocks has advantages both from behavior and statistical perspectives. Small choice sets are easier to handle and can be assessed faster by respondents. The statistical benefit is that both the variance and covariance of estimations are improved. The aim of this statistical design is to estimate the main factors and two-factor interactions in a quadratic equation with the lowest number of profiles. A limitation of their proposed design is that a maximum of four attributes can be analyzed at the individual level. They argue, however, that this is the average number of preferred attributes obtained after the first step in an adaptive CA.

Assessing the Applicability of Innovative Elicitation Concepts for Patient Preference Elicitation

Current Standards for Patient Preference Studies in Health Care

The ISPOR guideline (as published by Bridges et al [19]) consists of a checklist of 10 topics to be addressed when performing a CA in health care that aims at eliciting preferences at the meta level: Research question, Attributes and levels, Construction of tasks, Experimental design, Preference elicitation, Instrument design, Data-collection plan, Statistical analyses, Results and conclusions, and Study presentation. The MDIC framework focuses on patient preferences regarding benefit-risk assessments of medical device technologies in regulatory decision making [15]. They further provide several topics to consider when developing a preference study, which can be summarized as: defining the research question, the fit of a particular method to the research question, and resources available to undertake a patient preference study. The MDIC guideline discusses both qualitative and quantitative methods and when to use which [15]. The FDA guideline specifically refers to the ISPOR checklist and two other ISPOR guidelines related to good research practices when performing preference elicitation experiments [5,19,47,48]. The major complementarity of the FDA guideline to the other guidelines is its focus on how to inform or educate patients. This is equally important for preference elicitation at the individual or group level.

Applicability of Innovative Elicitation Concepts

The five identified concepts provide ideas on how to improve patient preference elicitation. Table 3 displays the assessment of which guideline items could potentially be improved by applying the five identified concepts [5,15,19]. Some concepts are process-oriented and could therefore potentially impact the entire development process. For example, the concept of feedback-driven exploration could have an impact on 9 out of the 10 steps described by the ISPOR guideline [19]. Other concepts focus on specific development steps, or even on more general challenges such as providing information to patients.

Table 3. Topics of health care guidelines that might benefit from implementing the identified concepts from the consumer research field.

Guideline and items	Simulating alternatives	Self-reflection	Separated adaptive dual response	Feedback-driven exploration	Arranging profiles in blocks
ISPOR[a] guideline					
Research question	—[b]	—	—	x[c]	—
Attributes and levels	—	—	—	x	—
Construction of tasks	—	x	x	x	x
Experimental design	—	x	x	x	—
Preference elicitation	—	x	—	x	—
Instrument design	x	—	—	x	—
Data collection	x	—	—	x	—
Statistical analysis	—	—	—	x	x
Results and conclusions	—	—	—	x	—
Study presentation	—	—	—	—	—
Total ISPOR guideline items that could be improved	2	3	2	9	2
FDA[d] guideline					
Patient centeredness	—	—	—	—	—
Representativeness of the sample and generalizability of results	—	—	—	—	—
Capturing heterogeneity of patients' preferences	—	—	—	—	—
Established good research practices by recognized professional organizations	—	—	—	—	—
Effective communication of benefit, harm, risk, and uncertainty	x	—	—	—	—
Minimal cognitive bias	—	—	x	—	—
Logical soundness	—	—	—	x	—
Relevance	—	—	—	x	—
Robustness of analysis of results	—	—	—	—	—
Study conduct	—	—	—	—	—
Comprehension by study participants	x	—	—	—	—
Total FDA guideline items that could be improved	2	0	1	2	0
MDIC[e] guideline: conjoint analysis and dual response experiments review					
Methodology criteria	x	x	x	x	x
Sample criteria	x	—	—	—	—
Analysis criteria	—	—	—	—	x
Output criteria	—	—	—	x	—
Total MDIC guideline items that could be improved	2	1	1	2	2

[a]ISPOR: International Society for Pharmacoeconomics and Outcomes Research.

[b]—: Guideline topic not impacted by concept implementation.

[c]x: Guideline topic might benefit from concept implementation.

[d]FDA: Food and Drug Administration.

eMDIC: Medical Device Innovation Consortium.

Discussion

Comparing the Decision-Making Process for Patients and Consumers

The decision-making process of consumers and patients is highly comparable. The main difference lies in the first step of the process, in which patients become aware of the decision context. More cognitive effort might be required to consider all relevant aspects of a health care context relative to that required for a consumer context. The remaining steps of the decision process are the same. For individual patients in the context of SDM, the process is equally comparable, with again only a few small differences. For example, upon receiving a diagnosis of breast cancer, a woman becomes aware of her need for therapy. The treating physician will provide information on the available options such as the possibility of breast-conserving surgery or mastectomy. The patient will then be advised to think about this for a few days and discuss her preferences with friends, family, or fellow sufferers. During a second consultation, the patient's preferences will be discussed, and a joint decision can be made. In case the patient is not ready to choose or does not want to participate in the decision-making process after all, the physician will propose their preferred option, which is likely to be carried out. The main difference here lies within the step of information gathering. High-quality information on diseases and potential therapy options is usually more difficult to obtain than information on consumption goods. Ideally, the patient receives all of the relevant information, or information sources, from the treating physician. As a second difference, the need for discussing the potential impact of available options with others might be higher in a health care setting than in a consumer setting. The other steps of the decision process are the same for both consumers and patients engaging in SDM. It should be noted, however, that these steps only apply when patients are offered the chance to actively engage in the decision-making process. According to the National Health Service in England, "SDM is relevant in any non-life threatening situation when a health or care decision needs to be made and a range of options (including doing nothing) is available" [49]. Although the process of SDM was introduced in health care decades ago, implementation is still lacking [7,50].

Applicability of Identified Concepts Within Current Standards for Patient Preference Elicitation

The identified concepts can be useful for one or more aspects of preference elicitation experiments as described by the guidelines. Some concepts can facilitate one or two specific items, whereas others can improve the entire development process. The latter is the case for process-oriented concepts such as feedback-driven exploration. By integrating all of the stakeholders' opinions in the development process, many aspects of preference experiments could be improved. For example, the attributes and levels could better reflect reality, as there is a smaller chance that relevant items will be left out. The construction of tasks, design, and data collection could be better adapted to patients' needs, resulting in clearer answers or higher performance rates. The same applies for developing a VCE as part of a decision aid for individual patients. Systematic development guidelines for decision aids already advise to work with a multidisciplinary team including patients and clinicians [51]. All relevant stakeholders should review multiple times and redesign as necessary. Owing to numerous initiatives, patients are now recognized as an important stakeholder in various aspects of health care [1,2,15,22,52]. As the European Patients Forum describes, there has been a transition from "doing things *to* the patient" to "doing things *with* the patient" [53]. Current standards for patient preference elicitation already suggest the use of interviews, and focus groups, among others, to guide the further development in (quantitative aspects of) preference instruments [5,15,19].

The concept of simulating alternatives will mainly improve informational aspects, as it will help people to fully understand the available choice alternatives. This can benefit patients by facilitating the entire process from becoming aware that a decision needs to be made to making that decision. The concept is equally applicable for designing VCEs in an individual context and for developing experiments with a group of patients. Defining the context and effectively communicating the benefits, harms, risks, and uncertainties is one of the first steps in both processes. The importance of this step has been highlighted by the FDA guideline [5]. Properly informing patients has been a longstanding challenge in health care; however, there are no satisfactory guidelines on *how* to do this. The PROTECT Benefit-Risk group compared visual representations to optimally provide information for a benefit-risk assessment [54]. They concluded that multiple formats (ranging from different graphs or plots to pictograms or risk scales) can be considered, and found no single visual type that was superior to others; however, the importance of considering the target audience when choosing a visual format was stressed. The authors further acknowledged the value of interactive/dynamic visuals, which enable active participation and improve understanding [54]. The use of simulating alternatives with photos or video materials seems to be a legitimate and feasible course of action to improve understanding and help create the necessary context to provide information [41]. For instance, researchers could show patients videos of how to use a medication with different modes of administration. This could help them to comprehend precisely what "an injection" entails, or how self-administration compares to administration by a nurse. This could be used by patients who recently underwent surgery and require anticoagulant therapy to prevent thrombosis, as these patients typically can choose between self-administering the injections or having a nurse administer the medication. If patients are shown a video of the complete procedure, including washing one's hands, disinfecting the skin area, using the right technique to pinch a fold of skin, injecting the syringe at the correct angle, and disposing the needle, they will be better equipped to make the decision of the administration method of the injections. When patients need to make an informed decision, it is important to adequately inform them on the different benefits and risks, but not influence their behavior [55]. That is, we want them to truly understand the benefits and risks, enabling them to make a fact-based decision depending on their values [55].

When all of the relevant information has been provided, respondents need time for self-reflection to let the acquired information sink in and decide which alternative(s) would be the most beneficial in their individual situation [43]. Current standards for patient preference elicitation do not explicitly state requirements concerning time needs to acquire and process information. However, both the FDA and the ISPOR task force warn against information overload or yea-saying, and suggest quizzing the respondents to verify comprehension [5,19]. The MDIC report also expressed the need to gain experience with preference studies and to learn how preferences that change over time can best be evaluated. Implementing this concept by conducting preference elicitation experiments over the course of a few days or weeks might be a good starting point. Researchers could alternatively provide respondents with the necessary information and a preparatory task to think about their preferences on the first day of the experiment. After a few days, the researchers would provide the same information and elicit their preferences during a final preference elicitation task. Of course, the downside to this approach is that the required time per respondent will almost double. Furthermore, the preferences of individuals with chronic diseases might change over time, along will their tradeoffs [15]. In the context of SDM, it is already considered to be good practice to provide patients with a decision aid in preparation for the consultation, as this will allow them to process the information, clarify their preferences, and prepare for a discussion [56].

After exploring the possible alternatives and taking the time to self-reflect, the next step in a decision process is deciding when to choose: now, later, or not at all (see Table 2) [26]. Including an opt-out option (opting not to make a choice or decision) in preference elicitation experiments can simulate the alternative of "choosing not to choose" [44]. In an individual context, this option may translate to watchful waiting or active surveillance. Another possibility is choosing to retain one's current course of action; for example, when a patient prefers their current treatment over all other options presented. As this situation is realistic, opt-out options should be included in patient preference elicitation experiments whenever relevant. This approach is already supported by the ISPOR health care guideline checklist [19]. An SADR can be used to overcome context effects or extreme response behavior in these cases [44]. However, in cases for which it would not be medically responsible to abstain treatment, this option should not be included in experiments that elicit group preferences [57].

Finally, arranging profiles in blocks of two has the advantage of imposing a low burden on respondents, as it requires less cognitive effort to consider two profiles multiple times rather than multiple profiles a few times [46]. In this way, respondents can repeat the process several times. There is also a statistical advantage, given that with a low number of tasks, doubling the tasks per respondent is equally effective in increasing precision as doubling the number of respondents [58]. The decompositional part of adaptive CA can also be completed with partial profiles, but the main benefit of evaluating two full profiles is that the respondents have the chance to evaluate complete products; this is more similar to the real-life situation by capturing all relevant aspects to consider [26]. Tailoring the

choice tasks for the user also fits within a natural decision-making process, as choosing must-have attributes can be a way of forming decision rules. For example, if a preference experiment comprises 10 different attributes, the respondents' answers could be used to gradually eliminate attributes that are considered less relevant by the respondent, resulting in fewer attributes that are used to form product profiles. This process can only be performed by a computer algorithm, implying the need for a computerized application. As this concept mainly provides statistical benefits, it is less relevant in a context of SDM where only the preferences of an individual patient have to be elicited.

Differences Between Consumers and Patients: Remaining Challenges

The extent to which we can apply consumer preference elicitation methods to simulate real-life decisions in a health care context is still unclear, both at the individual and meta level. In some respects such as when providing information, these methods could clearly offer improvements to enhance understanding. However, applying consumer preference elicitation concepts in health care will encounter limitations owing to some fundamental differences between health care products and consumables. First, health care is often a very complex matter relative to other consumer needs, making it difficult to fully understand the decision context such as a certain disease or the characteristics of the available options. Trying out different alternatives (eg, different smartphones, cars) is a useful approach in consumer research to obtain information on product characteristics or to determine the option that is most in line with personal needs. However, this solution is simply not possible in health care, as patients cannot test therapy options in the same way that consumers can test a new car. Simulations may be a very helpful alternative, although this will always require a high level of cognitive effort from respondents. Second, the impact of decisions in health care is relatively high, as the decisions are often irreversible. Third, health is an intrinsic part of a person, whereas consumable goods are interchangeable and can be used temporarily. This implies that preference elicitation methods in health care need to provide patients with more complex and more personal information to prepare them for decision making. Fourth, it is important to consider that consumer and health care products are often introduced differently in people's lives depending on the preference elicitation context. Buying consumables is usually a deliberate decision such as the decision to engage in a preference elicitation experiment for gathering data on market approval or reimbursement. This is different in the individual context, in which the need for health care products can be sudden and unexpected, as is the case upon receiving a diagnosis that is followed by the need to decide on therapy together with the treating physician. Additionally, buying consumables is usually more of an individual decision, whereas the decision-making environment in health care is very complex, often involving multiple stakeholders such as different health care providers, payers, regulatory agencies, and patient advocacy groups. When multiple stakeholder opinions must be taken into account, this impacts the choice of methodology. Finally, another challenge in health care is that preference data can be

useful for multiple purposes, ranging from individual to societal decisions. In addition to regulatory authorities, health technology assessment agencies or payers may also take patient preferences into account when making decisions regarding drug approval or reimbursement, respectively. Pharmaceutical companies might be equally interested in using this information to improve drug development. As each stakeholder evaluates preference data from its own perspective, it will be challenging to develop methods that fulfill all needs simultaneously. This versatile use of data is absent in consumer research, where the main goal is to align product development with consumer needs.

Limitations

The limitations of the study are the following. Only one search engine was used to perform the literature review, and although the list of search terms was quite extensive, it is possible that not all relevant papers were included. The publication date was limited to a maximum of 5 years ago, although older publications might also have concepts that have not yet been introduced in health care. Further, the identified papers were screened for exclusion by only one author, which could have resulted in selection bias.

Conclusions

The process of decision making is highly comparable between patients and consumers, although some small differences remain depending on the decision-making context. As a result, patients can be categorized as a subgroup of consumers. Therefore, learnings from the consumer research field might be valuable in health care. Five concepts from consumer preference elicitation that could help to simulate real-life decision making were identified in this study. Applying these concepts can result in structural improvements in the development process or improved execution of specific guideline items when eliciting patient preferences. However, the extent to which we can mimic real decision-making contexts in patient preference elicitation requires further research.

Conflicts of Interest

NVD is partially employed by Mindbytes, a company that creates serious games and educational tools for different applications, including preference elicitation experiments. GVS is the founder and CEO of ISMS, a pharmaceutical consultancy firm, and Mindbytes. IH declares no conflict of interest.

References

1. Ho M, Saha A, McCleary KK, Levitan B, Christopher S, Zandlo K, Medical Device Innovation Consortium's Patient Centered Benefit-Risk Steering Committee. A Framework for Incorporating Patient Preferences Regarding Benefits and Risks into Regulatory Assessment of Medical Technologies. Value Health 2016 Sep;19(6):746-750 [FREE Full text] [doi: 10.1016/j.jval.2016.02.019] [Medline: 27712701]

2. Mühlbacher AC, Juhnke C, Beyer AR, Garner S. Patient-Focused Benefit-Risk Analysis to Inform Regulatory Decisions: The European Union Perspective. Value Health 2016 Sep;19(6):734-740 [FREE Full text] [doi: 10.1016/j.jval.2016.04.006] [Medline: 27712699]

3. van Til JA, Ijzerman MJ. Why should regulators consider using patient preferences in benefit-risk assessment? Pharmacoeconomics 2014 Jan;32(1):1-4. [doi: 10.1007/s40273-013-0118-6] [Medline: 24288209]

4. Facey K, Hansen H, Single AV, editors. Patient Involvement In Health Technology Assessment 1st edition. Singapore: Springer Singapore; 2017.

5. US Food & Drug Administration. Patient Preference Information – Voluntary Submission, Review in Premarket Approval Applications, Humanitarian Device Exemption Applications, and De Novo Requests, and Inclusion in Decision Summaries and Device Labeling. 2016. URL: https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm446680.pdf [accessed 2019-01-10] [WebCite Cache ID 75K37o6uo]

6. Llewellyn-Thomas HA, Crump RT. Decision support for patients: values clarification and preference elicitation. Med Care Res Rev 2013 Feb;70(1 Suppl):50S-79S. [doi: 10.1177/1077558712461182] [Medline: 23124615]

7. Elwyn G, Frosch D, Thomson R, Joseph-Williams N, Lloyd A, Kinnersley P, et al. Shared decision making: a model for clinical practice. J Gen Intern Med 2012 Oct;27(10):1361-1367 [FREE Full text] [doi: 10.1007/s11606-012-2077-6] [Medline: 22618581]

8. Fagerlin A, Pignone M, Abhyankar P, Col N, Feldman-Stewart D, Gavaruzzi T, et al. Clarifying values: an updated review. BMC Med Inform Decis Mak 2013;13 Suppl 2:S8. [doi: 10.1186/1472-6947-13-S2-S8] [Medline: 24625261]

9. Klein AV, Hardy S, Lim R, Marshall DA. Regulatory Decision Making in Canada-Exploring New Frontiers in Patient Involvement. Value Health 2016;19(6):730-733 [FREE Full text] [doi: 10.1016/j.jval.2016.03.1855] [Medline: 27712698]

10.   Danner M, Hummel JM, Volz F, van Manen JG, Wiegard B, Dintsios C, et al. Integrating patients' views into health technology assessment: Analytic hierarchy process (AHP) as a method to elicit patient preferences. Int J Technol Assess Health Care 2011 Oct;27(4):369-375. [doi: 10.1017/S0266462311000523] [Medline: 22004779]

11.   Bewtra M, Johnson FR. Assessing patient preferences for treatment options and process of care in inflammatory bowel disease: a critical review of quantitative data. Patient 2013;6(4):241-255 [FREE Full text] [doi: 10.1007/s40271-013-0031-2] [Medline: 24127239]

12.   Mühlbacher AC, Zweifel P, Kaczynski A, Johnson FR. Experimental measurement of preferences in health care using best-worst scaling (BWS): theoretical and statistical issues. Health Econ Rev 2016 Dec;6(1):5 [FREE Full text] [doi: 10.1186/s13561-015-0077-z] [Medline: 26822869]

13.   European Patients' Academy. Guidance for patient involvement in industry-led medicines R&D.: EUPATI; 2016. URL: https://www.eupati.eu/patient-involvement/guidance-for-patient-involvement-in-industry-led-medicines-rd/ [accessed 2019-01-10] [WebCite Cache ID 75JzygMyz]

14.   Bernabe RDLC, van Thiel GJMW, van Delden J. Patient representatives' contributions to the benefit-risk assessment tasks of the European Medicines Agency scientific committees. Br J Clin Pharmacol 2014 Dec;78(6):1248-1256. [doi: 10.1111/bcp.12456] [Medline: 24995713]

15.   Medical Device Innovation Consortium. Medical Device Innovation Consortium (MDIC) Patient Centered Benefit-Risk Project Report: A Framework for Incorporating Information on Patient Preferences Regarding Benefit and Risk into Regulatory Assessments of New Medical Technology.: MDIC; 2015. URL: https://mdic.org/wp-content/uploads/2018/05/MDIC_PCBR_Framework_Web.pdf [accessed 2019-01-10] [WebCite Cache ID 75K3smD1m]

16.   Louviere J, Woodworth G. Design and Analysis of Simulated Consumer Choice or Allocation Experiments: An Approach Based on Aggregate Data. J Market Res 2018 Dec 13;20(4):350-367. [doi: 10.1177/002224378302000403]

17.   Lancsar E, Louviere J. Conducting discrete choice experiments to inform healthcare decision making: a user's guide. Pharmacoeconomics 2008;26(8):661-677. [doi: 10.2165/00019053-200826080-00004] [Medline: 18620460]

18.   De Bekker-Grob EW, Ryan M, Gerard K. Discrete choice experiments in health economics: a review of the literature. Health Econ 2012 Feb;21(2):145-172. [doi: 10.1002/hec.1697] [Medline: 22223558]

19.   Bridges JFP, Hauber AB, Marshall D, Lloyd A, Prosser LA, Regier DA, et al. Conjoint analysis applications in health--a checklist: a report of the ISPOR Good Research Practices for Conjoint Analysis Task Force. Value Health 2011 Jun;14(4):403-413 [FREE Full text] [doi: 10.1016/j.jval.2010.11.013] [Medline: 21669364]

20.   LaVela SL, Gallan AS. Evaluation and Measurement of Patient Experience. Patient Exp J 2014;1(1):28-36. [doi: 10.35680/2372-0247.1003]

21.   Dirksen CD, Utens CM, Joore MA, van Barneveld TA, Boer B, Dreesens DH, et al. Integrating evidence on patient preferences in healthcare policy decisions: protocol of the patient-VIP study. Implement Sci 2013 Jun 10;8:64 [FREE Full text] [doi: 10.1186/1748-5908-8-64] [Medline: 23758977]

22.   Chewning B, Bylund CL, Shah B, Arora NK, Gueguen JA, Makoul G. Patient preferences for shared decisions: a systematic review. Patient Educ Couns 2012 Jan;86(1):9-18 [FREE Full text] [doi: 10.1016/j.pec.2011.02.004] [Medline: 21474265]

23.   Johnson FR, Zhou M. Patient Preferences in Regulatory Benefit-Risk Assessments: A US Perspective. Value Health 2016 Sep;19(6):741-745 [FREE Full text] [doi: 10.1016/j.jval.2016.04.008] [Medline: 27712700]

24.   Cohen S, Neira L. Maximum Difference Scaling: Improved Measures of Importance and Preference for Segmentation. In: Sawtooth Software Research Conference Proceedings. Sequim, WA: Sawtooth Software Inc; 2003 Presented at: Sawtooth Software Conference; April 15-17, 2003; San Antonio, TX p. 61-74.

25.   Urban G, Hauser J, Qualls W, Weinberg B, Bohlmann J, Chicos R. Information Acceleration: Validation and Lessons from the Field. J Market Res 2018 Dec 20;34(1):143-153. [doi: 10.1177/002224379703400112]

26.   Louviere JJ. What You Don't Know Might Hurt You: Some Unresolved Issues in the Design and Analysis of Discrete Choice Experiments. Environ Resource Econ 2006 May;34(1):173-188. [doi: 10.1007/s10640-005-4817-0]

27.   Hoeffler S. Measuring Preferences for Really New Products. J Market Res 2018 Oct 10;40(4):406-420. [doi: 10.1509/jmkr.40.4.406.19394]

28.   Urban GL, Weinberg BD, Hauser JR. Premarket Forecasting of Really-New Products. J Marketing 1996 Jan;60(1):47-60. [doi: 10.2307/1251887]

29.   Sørensen K, Pelikan JM, Röthlin F, Ganahl K, Slonska Z, Doyle G, et al. Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). Eur J Public Health 2015 Dec;25(6):1053-1058 [FREE Full text] [doi: 10.1093/eurpub/ckv043] [Medline: 25843827]

30.   Urban GL, Hauser JR, Roberts JH. Prelaunch Forecasting of New Automobiles. Manage Sci 1990 Apr;36(4):401-421. [doi: 10.1287/mnsc.36.4.401]

31.   Kolb AY, Kolb DA. Experiential learning theory: A dynamic, holistic approach to management learning, education and development. In: Armstrong S, Fukami CV, editors. The Sage Handbook Of Management Learning, Education And Development. California: Sage Publications Ltd; 2009:42-68.

32.   Louviere J, Henser A, Swait J. Stated Choice Methods Analysis and Applications. New York: Cambridge University Press; 2010.

33. Ali S, Ronaldson S. Ordinal preference elicitation methods in health economics and health services research: using discrete choice experiments and ranking methods. Br Med Bull 2012 Sep 02;103(1):21-44. [doi: 10.1093/bmb/lds020] [Medline: 22859714]

34. Domigall Y, Albani A, Winter R. Identification of customer preferences for new service development in the electricity domain. 2014 Presented at: Conference on Business Informatics Identification; 14.07.2014; Geneva. [doi: 10.1109/CBI.2014.34]

35. Lancsar E, Swait J. Reconceptualising the external validity of discrete choice experiments. Pharmacoeconomics 2014 Oct;32(10):951-965. [doi: 10.1007/s40273-014-0181-7] [Medline: 24920196]

36. Holm S, Lemm R, Thees O, Hilty LM. Enhancing Agent-Based Models with Discrete Choice Experiments. J Artif Soc Soc Simul 2016;19(3):3. [doi: 10.18564/jasss.3121]

37. Eimecke J, Baier D. Preference Measurement in Complex Product Development: A Comparison of Two-Staged SEM Approaches. In: Lausen B, Krolak-Schwerdt S, Böhmer M, editors. Data Science, Learning by Latent Structures, and Knowledge Discovery. Berlin, Heidelberg: Springer; 2015:239-250.

38. Rossi S, Barile F, Di Martino S, Improta D. A comparison of two preference elicitation approaches for museum recommendations. Concurr Comp-Pract E 2017 Mar 21;29(11):e4100. [doi: 10.1002/cpe.4100]

39. Wang C. Using the theory of inventive problem solving to brainstorm innovative ideas for assessing varieties of phone-cameras. Comput Ind Eng 2015 Jul;85:227-234. [doi: 10.1016/j.cie.2015.04.003]

40. Huang D, Luo L. Consumer Preference Elicitation of Complex Products Using Fuzzy Support Vector Machine Active Learning. Market Sci 2016 May;35(3):445-464. [doi: 10.1287/mksc.2015.0946]

41. Vignola R, Gonzalez-Rodrigo B, Lane O, Marchamalo M, McDaniels T. A scenario approach to assess stakeholder preferences for ecosystem services in agricultural landscapes of Costa Rica. Reg Environ Change 2016 Sep 16;17(2):605-618. [doi: 10.1007/s10113-016-1051-y]

42. Xiang W, Clarke KC. The Use of Scenarios in Land-Use Planning. Environ Plann B Plann Des 2016 Nov 30;30(6):885-909. [doi: 10.1068/b2945]

43. Hauser JR, Dong S, Ding M. Self-Reflection and Articulated Consumer Preferences. J Prod Innov Manag 2013 Oct 04;31(1):17-32. [doi: 10.1111/jpim.12077]

44. Schlereth C, Skiera B. Two New Features in Discrete Choice Experiments to Improve Willingness-to-Pay Estimation That Result in SDR and SADR: Separated (Adaptive) Dual Response. Manage Sci 2017 Mar;63(3):829-842. [doi: 10.1287/mnsc.2015.2367]

45. Boesch I, Schwaninger M, Weber M, Scholz RW. Enhancing Validity and Reliability Through Feedback-Driven Exploration: A Study in the Context of Conjoint Analysis. Syst Pract Action Res 2012 Sep 18;26(3):217-238. [doi: 10.1007/s11213-012-9248-6]

46. Huertas-Garcia R, Gázquez-Abad JC, Forgas-Coll S. A design strategy for improving adaptive conjoint analysis. J Bus Ind Mark 2016 Apr 04;31(3):328-338. [doi: 10.1108/JBIM-02-2013-0043]

47. Hauber AB, González JM, Groothuis-Oudshoorn CGM, Prior T, Marshall DA, Cunningham C, et al. Statistical Methods for the Analysis of Discrete Choice Experiments: A Report of the ISPOR Conjoint Analysis Good Research Practices Task Force. Value Health 2016 Jun;19(4):300-315 [FREE Full text] [doi: 10.1016/j.jval.2016.04.004] [Medline: 27325321]

48. Reed Johnson F, Lancsar E, Marshall D, Kilambi V, Mühlbacher A, Regier DA, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. Value Health 2013;16(1):3-13 [FREE Full text] [doi: 10.1016/j.jval.2012.08.2223] [Medline: 23337210]

49. NHS England and NHS Improvement. Personalised Care. 2019. Shared Decision Making: Summary guide URL: https://www.england.nhs.uk/wp-content/uploads/2019/01/shared-decision-making-summary-guide-v1.pdf [accessed 2019-11-29]

50. Légaré F, Ratté S, Gravel K, Graham ID. Barriers and facilitators to implementing shared decision-making in clinical practice: update of a systematic review of health professionals' perceptions. Patient Educ Couns 2008 Dec;73(3):526-535. [doi: 10.1016/j.pec.2008.07.018] [Medline: 18752915]

51. Coulter A, Stilwell D, Kryworuchko J, Mullen PD, Ng CJ, van DWT. A systematic development process for patient decision aids. BMC Med Inform Decis Mak 2013;13 Suppl 2:S2. [doi: 10.1186/1472-6947-13-S2-S2] [Medline: 24625093]

52. European Patients' Academy on Therapeutic Innovation. EUPATI Project: Executive Summary. 2017. URL: https://3qksc436bu713cqimwcfglyj-wpengine.netdna-ssl.com/wp-content/uploads/2017/06/EUPATI-Final-Project-Executive-Summary-1.1.pdf [accessed 2019-01-10] [WebCite Cache ID 75K9yzhtQ]

53. European Patients Forum. EPF Background Brief: Patient Empowerment. Brussels: European Patients Forum; 2015. URL: http://www.eu-patient.eu/globalassets/campaign-patient-empowerment/briefing_paperpatient-empowerment_final_external.pdf [accessed 2019-01-10] [WebCite Cache ID 75KA8IU25]

54. Hallgreen CE, Mt-Isa S, Lieftucht A, Phillips LD, Hughes D, Talbot S, PROTECT Benefit-Risk group. Literature review of visual representation of the results of benefit-risk assessments of medicinal products. Pharmacoepidemiol Drug Saf 2016 Mar;25(3):238-250. [doi: 10.1002/pds.3880] [Medline: 26521865]

55. Fischhoff B, Brewer N, Downs J. Communicating Risks and Benefits: An Evidence-Based User's Guide. Silver Spring, MD: US Department of Health and Human Services, Food and Drug Administration; 2011. URL: https://www.fda.gov/

downloads/AboutFDA/ReportsManualsForms/REports/UCM268069.pdf [accessed 2019-01-10] [WebCite Cache ID 75K5KabXh]

56.   Stacey D, Légaré F, Col NF, Bennett CL, Barry MJ, Eden KB, et al. Decision aids for people facing health treatment or screening decisions. Cochrane Database Syst Rev 2014;1:CD001431. [doi: 10.1002/14651858.CD001431.pub4] [Medline: 24470076]

57.   Mühlbacher A, Bethge S. First and Foremost Battle the Virus: Eliciting Patient Preferences in Antiviral Therapy for Hepatitis C Using a Discrete Choice Experiment. Value Health 2016;19(6):776-787 [FREE Full text] [doi: 10.1016/j.jval.2016.04.007] [Medline: 27712705]

58.   Johnson R, Orme B. Sawtooth Softw Res Pap Ser. Sequim, WA: Sawtooth Software, Inc; 1996. How Many Questions Should You Ask in Choice-Based Conjoint Studies? URL: https://www.sawtoothsoftware.com/download/techpap/howmanyq.pdf [accessed 2019-01-10] [WebCite Cache ID 75KAJZ3ZG]

Abbreviations

CA: conjoint analysis
DCE: discrete choice experiment
FDA: Food and Drug Administration
FDE: feedback-driven exploration
ISPOR: International Society for Pharmacoeconomics and Outcomes Research
MDIC: Medical Device Innovation Consortium
SADR: separated adaptive dual response
SDM: shared decision making
SDR: separated dual response
VCE: value clarification exercise

XSL•FO
RenderX

Original Paper

Parental Knowledge, Attitudes, and Behaviors Toward Their Epileptic Children at King Abdulaziz University Hospital: Cross-Sectional Study

Abdulelah Kinkar[1], MBBS; Dalya Alqarni[1], MBBS; Abdulaziz Alghamdi[1], MBBS; Sahal Wali[1], MBBS; Nasser Alghamdi[1], MBBS; Saeed Saloom[1], MBBS; Mooataz Aashi[2], MBBS

[1]Faculty of Medicine, King Abdulaziz University, Jeddah, Saudi Arabia
[2]Department of Pediatrics, King Abdulaziz University Hospital, Jeddah, Saudi Arabia

Corresponding Author:
Abdulelah Kinkar, MBBS
Faculty of Medicine, King Abdulaziz University
Al Ehtifalat St
Jeddah, 2214
Saudi Arabia
Phone: 966 564832244
Email: alqarnidalya@gmail.com

Abstract

Background: Epilepsy is a chronic disease characterized by periodic seizures that result from abnormal integrated firing impulses in the brain. It is one of the most common neurological disorders. Over the past few years, there has been increasing awareness about the effect that having a child with epilepsy has on parents and the reciprocal impact of parental knowledge and attitudes regarding epilepsy on the affected child.

Objective: This study aimed to assess parental knowledge, attitudes, and behavior toward their epileptic children.

Methods: A cross-sectional study was conducted in 2018 by the Pediatric Neurology Department of King Abdulaziz University Hospital, Jeddah, the Kingdom of Saudi Arabia. A sample size of 115 of 332 parents who have a child diagnosed with epilepsy and aged 18 years or younger were recruited for this study. Statistical analysis was performed using SPSS version 21. Data analysis was performed using an independent *t* test, a chi-square test, one-way analysis of variance, and correlation analysis.

Results: A total of 115 participants answered the questionnaire; of these, 65 (56.5%) were men, with an average age of 40.3 years, and the mean age of the children was 9.0 years. Overall, 85 (85/115, 73.9%) children were taken care of by both of their parents. The mean parental knowledge score was 7.49 (SD 2.08) out of 12, and it was significantly related to the educational level of the parent (P=.004). The knowledge question that was most frequently answered incorrectly was "Diagnosis of epilepsy is usually made based on at least two unprovoked seizures." As only 28.7% (33/115) of participants chose the correct answer, mean parental attitude score was 26.51 (SD 4.284) out of 35, and there was no significant relation with the educational level of parents (P=.13); however, it was negatively correlated with the child's age (P=.045). Mean parental behavioral score was 23.35 (SD 4.121) out of 35, and there was no significant relation with the educational level of the parents (P=.24). The most negatively answered question for the behavior section was "I can leave my child without supervision," with a mean score of 2.25 (SD 1.09) out of 5. Gender did not play a significant role in parental knowledge, attitudes, or behavior (P=.44, P=.77, and P=.99, respectively).

Conclusions: Parental knowledge in our sample still needs improvement. Therefore, more awareness campaigns should be made for the community and for the parents of affected children to create a supportive environment for the children and help them thrive and develop.

XSL•FO
RenderX

Introduction

Epilepsy is a chronic disease characterized by periodic seizures that result from abnormal integrated firing impulses in the brain [1]. It is one of the most common neurological disorders [2]. The primary cause of epilepsy is unknown, but brain tumors and other diseases may cause epilepsy [3]. Most seizures in children are predisposed by disorders from outside the brain, such as high fever, infection, and trauma, and sometimes by genetic diseases [4,5].

Around 50 million people have epilepsy worldwide, which accounts for 1% of the global burden of disease [6]. Among Arab countries, it is demonstrated that around 724,500 persons had epilepsy. A survey was conducted in Saudi Arabia to determine the prevalence of epilepsy and other convulsive disorders, and it found that the rate of active epilepsy was 6.54 per 1000 persons (95% CI 5.48-7.60) [7], where 65% were patients whose epilepsy started before the age of 18 years [1]. Thus, in children, the dilemma of epilepsy is more severe than in adults; therefore, up to 50% of children with epilepsy might have psychiatric and behavioral comorbidities, such as learning disabilities, developmental delay, and autism spectrum disorders [8].

Public attitudes and knowledge toward epilepsy differ from one culture to another; for example, convulsive episodes at an unexpected moment in public may result in discrimination toward someone who is suffering from epilepsy [9-16]. It has been noted that traditional thoughts and poor knowledge strongly affect attitudes toward epilepsy [5,9]. Over the past few years, there has been increasing awareness about the effect of having a child with epilepsy has on parents and the reciprocal impact of parental knowledge and attitudes regarding epilepsy on the affected child [17,18]. Besides, parental attitudes toward epilepsy were as significantly associated with the child outcome as seizure history and epilepsy duration [19]. In addition, it is related to their psychosocial issue [15] because taking care of a child suffering from epilepsy results in a more significant amount of stress, which is associated with exaggerated fears and protective behaviors toward their children, that results in poor child behavioral outcome, compared with taking care of healthy child [19].

On reviewing multiple publications in Saudi Arabia, we concluded that there is little information related to this study; therefore, we aimed to study and evaluate parental knowledge, attitudes, and behavior toward their epileptic child and conclude whether further awareness programs are essential to improve the overall quality of life for epileptic children and their caregivers.

Methods

Study Design

The institutional review board of King Abdulaziz University (KAU) Hospital approved this study. A cross-sectional study assessing parental knowledge, attitudes, and behavior toward their epileptic children was conducted in 2018 by the Pediatric Neurology Department of KAU Hospital, Jeddah, the Kingdom of Saudi Arabia. A sample size of 115 of 332 parents who have a child diagnosed with epilepsy and aged 18 years or younger were recruited for this study. The data collected through telephone interview or Web-based Google form questionnaire, which was modified from international surveys that were previously used to assess parental knowledge and attitudes toward epilepsy, were translated into Arabic and then back translated to make sure the translation was accurate. Those who agreed to participate provided verbal consent.

Microsoft Excel version 2013 was used for data entry, and statistical analysis was performed using IBM SPSS Inc version 21. Data analyses were performed using an independent *t* test, a chi-square test, one-way analysis of variance, and correlation analysis.

The questionnaire contained two segments: sociodemographic characteristics of the parents (eg, gender, age, nationality, and educational degree) and demographic characteristics of their children with epilepsy (eg, age, educational level, child performance in school, the age of diagnosis, and who is the caregiver of the child).

Parental Knowledge Toward Epileptic Children

The knowledge section of the questionnaire contained 12 items, 10 of which were yes or no or I don't know questions. These questions included whether epilepsy is contagious, psychological, or genetic; whether or not all epileptic children lose consciousness during seizures; do all of the affected children show the same symptoms; is it possible for epileptic children to take vaccine according to a specific timetable; whether or not epileptic children had lower IQ than their peers; and whether or not epileptic children had difficulties in learning. The last 2 yes/no/I don't know questions were regarding the use of medications—whether or not the child is cured of seizures after using medicine and whether or not, after forgetting to take the prescribed drug, the patient must double the dose next time. The other 2 questions had multiple choices: age of onset of seizures (before 1 year of age/until 18 years of age/at any age) and what indicates the diagnosis of epilepsy (loss of consciousness during a seizure/one unprovoked seizure/at least two unprovoked seizures).

Each correct answer was given 1 point. Therefore, the total score of knowledge was graded between 0 (the lowest grade) and 12 (the highest grade) points.

Parental Attitudes Toward Epileptic Children

A total of 7 statements were scored using Likert scales (summative scales; strongly disagree, disagree, neither agree nor disagree, agree, and strongly agree), where the minimum score 1 represented "I strongly disagree" and the maximum score 5 represented "I strongly agree." The total score ranged from 7 to 35. Parents were asked about whether they feel comfortable if their family or friends knew that they have an epileptic child; if they would be worried about their child capability on making friends; if, perhaps, other people will treat their child differently; if they believe that their child will be able to achieve success in his/her career; if they believe their child will be as qualified as any other healthy children or, if there is more social support, the child will have more chances

to self-develop; and if they receive enough support from the society for their epileptic child.

Parental Behavior Toward Epileptic Children

A total of 7 statements were scored using Likert scales (strongly disagree, disagree, neither agree nor disagree, agree, and strongly agree), where the minimum score 1 represented "I strongly disagree" and the maximum score 5 represented "I strongly agree." The total score ranged from 7 to 35. Parents were asked about whether they are capable of taking care of their child during seizures, if they allow their child to participate in any activity he/she desires, if they would let their child to play video games or play a sport or to go on school trips, if they can leave their child without supervision, and if the parents have time to meet other people and participate in other activities for themselves.

Results

This study aimed to evaluate parental knowledge, attitudes, and behaviors toward their epileptic children.

Of the 115 study participants, 65 (56.5%) were men. The average parental age was 40.3 years (SD 8.4; age range: 24-63 years). Overall, 77 (67.0%) were Saudi parents. In addition, 54 (47.0%) parents received a college education, and 4 (3.5%) did not receive any education.

Almost three-fourth of the children were taken care of by both of their parents (85/115, 74.0%). The mean age of the children was 9.0 years (SD 5.3; age range: 1-18 years). Moreover, 51 (44.3%) children were diagnosed with epilepsy before the first year of life.

Other sociodemographic characteristics of parents and children are presented in Table 1.

The mean parental knowledge score was 7.49 (SD 2.08). Of 115 parents, 104 (90.4%) knew that epilepsy is not a contagious disease. In contrast, only 33 (28.6%) knew how epilepsy is diagnosed. There was no significant difference in the knowledge scores between men (mean score 7.35) and women (mean score 7.66; P=.44). Education had a statistically significant effect; parents with more than 12 years of schooling scored higher in knowledge (mean score 8.09) than those who spent less than 12 years (mean score 6.97; P=.004). There was no significant difference in the knowledge scores between Saudi (mean score 7.4) and non-Saudi parents (mean score 7.66; P=.54). Similarly, we found no significant relationship between knowledge scores and child performance in school (P=.59). All knowledge-related questions are presented in Table 2.

Mean parental attitude scores for each question are presented in Table 3. Parental attitudes were measured using the Likert scale, scored from 7 to 35, and the mean score was 26.51 (SD 4.284). The highest average positive attitude was obtained for the statement, "If there is more social support, my child will have more chances to self-develop" (4.3 out of 5). In contrast, the lowest average positive attitude was observed for the statement, "I do not feel that other people treat my child as different" (3.44 out of 5). There was no significant difference in attitude scores between men (mean 3.80) and women (mean 3.77; P=.77). Similarly, there was no difference in attitude according to the parent's years of studying (P=.13) or according to the parent's nationality (Saudi: mean 3.77 and non-Saudi: mean 3.82; P=.73). In addition, there was no significant association between the child's age of diagnosis and parental attitude score (P=.49).

Mean parental behavioral scores for each question are presented in Table 4. The mean score for the behavioral section was 23.35 (SD 4.121). The statement with the highest mean score was "I allow my child to play sport" (4.17 out of 5), whereas the statement with the lowest mean score was "I can leave my child without supervision" (2.25 out of 5). There was no significant difference in behavioral score between men (mean 3.34) and women (mean 3.33; P=.99). Similarly, there was no difference in behavior according to the parent's years of studying (P=.24) nor according to the parent's nationality (P=.88) or according to the child's performance in school (P=.96). The relationship between different sociodemographic characteristics and mean scores in knowledge, attitudes, and behavior are presented in the Tables 5 and 6.

Table 1. Demographic characteristics of parents and their children with epilepsy.

Characteristics	Values
Parents	
Gender, n (%)	
Male	65 (56.5)
Female	50 (43.5)
Age (years), mean (SD)	
24-34	28 (24.3)
35-44	55 (47.9)
45-54	26 (22.6)
≥55	6 (5.2)
Nationality, n (%)	
Saudi	77 (67.0)
Non-Saudi	38 (33.0)
Education level, n (%)	
Primary school	7 (6.1)
Intermediate school	16 (13.6)
High school	33 (28.7)
College	55 (47.8)
None	4 (3.5)
Children	
Age (years), mean (SD)	
≤6	41 (35.7)
7-14	50 (43.4)
15-18	24 (20.9)
Education level, n (%)	
Primary school	26 (22.6)
Intermediate school	9 (7.8)
High school	15 (13.0)
None	65 (56.5)
Performance, n (%)	
Excellent	11 (9.6)
Needs improvement	26 (22.6)
Good	13 (11.3)
None	65 (56.5)
Age of diagnosis (years), n (%)	
<1	51 (44.3)
1-4	26 (22.6)
5-9	25 (21.7)
10-14	6 (5.2)
15-18	7 (6.1)
Who is taking care, n (%)	
Both parents	85 (73.9)
Father	1 (0.9)

Characteristics	Values
Mother	28 (24.3)
Sister	1 (0.9)

Table 2. Percentage of correct answers to questions comprising the knowledge score of parents of children with epilepsy.

Item	Correct answers (%)
Epilepsy is not an infectious disease	90.4
Epilepsy is not a psychiatric disease	73.0
Epilepsy is not, for the most part, hereditary	33.0
Not all the affected children lose consciousness during seizures	33.0
Not all the affected children have the same symptoms	71.3
Children with epilepsy are able to get vaccinated according to the current immunization calendar	63.5
Children with epilepsy, for the most part, do not have a lower IQ	47.0
Withdrawal of seizures after medication use does not mean that the patient is cured	76.5
If you skip therapy, next time, you should not take a double dose of medications	82.6
Children with epilepsy may encounter difficulties in learning	64.3
Onset of seizures may occur at any age	85.2
Diagnosis of epilepsy is usually made based on at least two unprovoked seizures	28.7

Table 3. Mean scores and standard deviations of attitude-related questions.

Question	Score, mean (SD)
I want my family and friends to know my child is epileptic	3.84 (1.196)
I do not feel that other people treat my child as different	3.44 (1.265)
I believe my child will be able to achieve success in his/her career	4.12 (0.984)
I am not worried about my child's capability to make friends	3.75 (1.067)
I feel I receive enough support from my society regarding my epileptic child	3.49 (1.187)
I believe my child is as qualified as any other healthy children	3.57 (1.148)
If there is more social support, my child will have more chances to self-develop	4.30 (0.946)

Table 4. Mean scores and standard deviations of behavior-related questions.

Question	Score, mean (SD)
I am capable in taking care of my child during seizures	3.89 (1.145)
In general, I allow my child to participate in any activity he desires	3.64 (1.078)
I allow my child to play video games	3.06 (1.045)
I allow my child to play sport	4.17 (0.798)
Although my child is epileptic, I have time to meet other people and participate in other activities	4.06 (0.841)
I can leave my child without supervision	2.25 (1.091)
I allow my child to go on school trips	2.27 (1.126)

Table 5. Relationship between different sociodemographic characteristics and mean scores in knowledge, attitudes, and behavior, using independent samples *t* test and one-way analysis of variance.

Factor	Knowledge		Attitude		Behavior	
	Mean[a]	*P* value	Mean[b]	*P* value	Mean[b]	*P* value
Parents						
Gender		.44		.77		.99
Male	7.35		3.80		3.34	
Female	7.66		3.77		3.33	
Nationality		.54		.73		.88
Saudi	7.40		3.77		3.34	
Non-Saudi	7.66		3.82		3.32	
Education (years)		.004		.13		.24
<12	6.97		3.70		3.27	
>12	8.09		3.88		3.40	
Children						
Education		.51				.36
Primary school	7.58		3.96	.23	3.49	
Intermediate school	6.56		3.52	.23	3.33	
High school	7.27		3.69	.23	3.40	
Not studying	7.63		3.78	.23	3.26	
Age of diagnosis (years)		.35		.49		.32
<1	7.75		3.89		3.35	
1-4	7.50		3.66		3.29	
5-9	7.40		3.81		3.50	
10-14	7.33		3.60		3.02	
15-18	6.00		3.63		3.12	
Performance[c]		.59		.18		.96
Above average	7.13		3.90		3.44	
Below average	7.46		3.71		3.43	
Primary caregiver		.68		.76		.88
Father	6.00		3.29		3.29	
Mother	7.25		3.86		3.40	
Both parents	7.56		3.77		3.31	
Sister	9.00		3.71		3.57	

[a]Knowledge mean score out of 12.

[b]Attitudes and behavior mean score out of 5.

[c]Children who are not in school were excluded.

Table 6. Relationship between different sociodemographic characteristics and mean scores in knowledge, attitudes, and behavior, using Pearson correlation test.

Factor	Knowledge		Attitudes		Behavior	
	Pearson correlation	*P* value	Pearson correlation	*P* value	Pearson correlation	*P* value
Parental age	0.030	.75	−0.072	.44	0.004	.97
Child age	−0.068	.47	−0.188	.045	−0.016	.87

XSL•FO
RenderX

Discussion

Principal Findings

This study was conducted on parents to evaluate the knowledge, attitudes, and behaviors of parents toward their epileptic children; few parents answered all questions correctly (2/115, 2.0%). On the one hand, most of the participants (104/115, 90.4%) strongly think that epilepsy is not a communicable disease, and the mean score is less than 2 in previously conducted studies in Jordan (98.5%) and Serbia (99.5%) [20,21], yet slightly higher than that in a study conducted among Turkish parents (89.5%) and Iranian parents (83%) [22,23]. On the other hand, only 32% of parents in a study in Nigeria knew that epilepsy is not communicable [24]. This difference might be because of the variation in sociodemographic characteristics of other populations as well as a difference in educational quality. Most people still believe that the epilepsy is a mental disorder [25], but in our study, most parents knew it is not psychological (73%), compared with the Iranian study (60.2%) [23], Serbian study (68.1%) [21], and a study in Jordan (90.3%) [20].

Although most of our sample had a high level of education, there is still lack of knowledge and wrong beliefs about epilepsy, for example, 33% of parents think that epilepsy is mostly not hereditary, similar to the Serbian study in which 32% of parents answered correctly [21]. One possible cause of this misconception might be because of the high prevalence of consanguinity in our culture [26]. There is a significant relationship between the age of diagnosis and the knowledge about epilepsy; therefore, the earlier period of diagnosis might encourage the parent to read and know more about epilepsy to take care of their children properly and to make their life better.

In the attitudes section of the questionnaire, we found a positive response about the desire to inform friends and family members about their child who is suffering from epilepsy (mean score 3.84), suggesting that they already knew that the disease is not shameful. Serbian parents responded even more positively regarding that statement (mean score 4.3) [21]. Social support is essential to those children to help them defeat the frustrations and problems they might face and help enable them to be more productive in society. Moreover, our results showed that most of the parents believe that they have enough social support (mean score 3.49 out of 5); however, if they agreed that if they

have more social support, their child will have better chances for self-development (mean score 4.30 out of 5). However, the Serbian parents agreed more strongly that their children already received enough social support (mean score 3.8 out of 5). They also agreed, to a lesser extent, that more social support will help their children grow and improve (mean score 3.2 out of 5) [21].

Behaviors of parents toward their children differ according to educational level, economic level, and cultural habits [27]. In our sample, we found that parents are supportive of activities that would help their children be happy and healthy, and this is a good result compared with Turkish mothers who were found to be less supportive [28]. We received a profoundly negative result about the possibility of leaving an epileptic child without supervision (mean 2.25), which could indicate greater parental responsibilities toward their children and an understanding about the disease and its sudden seizures so that it will give those children a good chance to avoid injuries during an unprovoked seizure.

One of the limitations of this study was the small sample size because of the high number of parents not answering the telephone as well as the poor cooperation and linguistic abilities of some parents. Furthermore, patients who did not answer the call or refused to participate may have a different set of knowledge, attitudes, and behaviors.

Conclusions

In conclusion, the mean knowledge score of our sample was less than our expectation, and only 2 of 115 participants correctly answered all knowledge-related questions. As expected, we found a significant relationship between parental educational degree and knowledge score. Parental knowledge in our sample still needs improvement. Therefore, more awareness campaigns should be conducted for the community and for the parents of affected children to create a supportive environment for the children and help them thrive and develop and to help parents gain the skills to control their children's epilepsy to minimize the negative outcomes. Thankfully, multiple epilepsy-related educational programs around the world showed significant efficacy to correct misconceptions and improve parental knowledge toward their children's condition [18,29-31], and the implementation of such programs should be considered in the future plans for improving knowledge in Saudi Arabia.

Conflicts of Interest

None declared.

References

1. Baulac M, de Boer H, Elger C, Glynn M, Kälviäinen R, Little A, et al. Epilepsy priorities in Europe: a report of the ILAE-IBE Epilepsy Advocacy Europe Task Force. Epilepsia 2015 Nov;56(11):1687-1695 [FREE Full text] [doi: 10.1111/epi.13201] [Medline: 26415919]
2. Scott RA, Lhatoo SD, Sander JW. The treatment of epilepsy in developing countries: where do we go from here? Bull World Health Organ 2001;79(4):344-351 [FREE Full text] [Medline: 11357214]
3. Kumar P, Clark ML. Clinical Medicine. Sixth Edition. Edinburgh: Saunders; 2008.
4. Fukuyama Y. Proceedings of the first national meeting of the study group on seizures in infancy and early childhood Tokyo, April 4, 1998. Brain Dev 1998 Oct;20(7):550-558. [doi: 10.1016/s0387-7604(98)00055-2] [Medline: 9840678]

5. Osuntokun BO, Adeuja AO, Nottidge VA, Bademosi O, Olumide A, Ige O, et al. Prevalence of the epilepsies in Nigerian Africans: a community-based study. Epilepsia 1987;28(3):272-279. [doi: 10.1111/j.1528-1157.1987.tb04218.x] [Medline: 3582291]

6. World Health Organization. Atlas: Epilepsy Care in the World 2005. Geneva: World Health Organization; 2005.

7. Benamer HT, Grosset DG. A systematic review of the epidemiology of epilepsy in Arab countries. Epilepsia 2009 Oct;50(10):2301-2304 [FREE Full text] [doi: 10.1111/j.1528-1167.2009.02058.x] [Medline: 19389149]

8. Baca CB, Vickrey BG, Caplan R, Vassar SD, Berg AT. Psychiatric and medical comorbidity and quality of life outcomes in childhood-onset epilepsy. Pediatrics 2011 Dec;128(6):e1532-e1543 [FREE Full text] [doi: 10.1542/peds.2011-0245] [Medline: 22123895]

9. Saengpattrachai M, Srinualta D, Lorlertratna N, Pradermduzzadeeporn E, Poonpol F. Public familiarity with, knowledge of, and predictors of negative attitudes toward epilepsy in Thailand. Epilepsy Behav 2010 Apr;17(4):497-505. [doi: 10.1016/j.yebeh.2010.01.164] [Medline: 20189885]

10. Shore CP, Buelow JM, Austin JK, Johnson CS. Continuing psychosocial care needs in children with new-onset epilepsy and their parents. J Neurosci Nurs 2009 Oct;41(5):244-250 [FREE Full text] [doi: 10.1097/jnn.0b013e3181b2c721] [Medline: 19835237]

11. Pupillo E, Vitelli E, Messina P, Beghi E. Knowledge and attitudes towards epilepsy in Zambia: a questionnaire survey. Epilepsy Behav 2014 May;34:42-46. [doi: 10.1016/j.yebeh.2014.02.025] [Medline: 24681384]

12. Maiga Y, Albakaye M, Diallo LL, Traoré B, Cissoko Y, Hassane S, et al. Current beliefs and attitudes regarding epilepsy in Mali. Epilepsy Behav 2014 Apr;33:115-121 [FREE Full text] [doi: 10.1016/j.yebeh.2014.02.031] [Medline: 24657502]

13. Bhesania NH, Rehman A, Savul IS, Zehra N. Knowledge, attitude and practices of school teachers towards epileptic school children in Karachi, Pakistan. Pak J Med Sci 2014 Jan;30(1):220-224 [FREE Full text] [doi: 10.12669/pjms.301.4307] [Medline: 24639865]

14. Dorota T, Witkowska M, Michalak M. Attitudes of employees of service and trading companies towards people with epilepsy and their professional activity in Poland. Seizure 2014 Mar;23(3):178-183 [FREE Full text] [doi: 10.1016/j.seizure.2013.11.008] [Medline: 24332504]

15. Gzirishvili N, Kasradze S, Lomidze G, Okujava N, Toidze O, de Boer HM, et al. Knowledge, attitudes, and stigma towards epilepsy in different walks of life: a study in Georgia. Epilepsy Behav 2013 May;27(2):315-318. [doi: 10.1016/j.yebeh.2013.02.011] [Medline: 23524468]

16. Otte WM, Meeuwesse E, Kafiluddin E, Peerdeman SM, Baaijen JC, Sander JW, et al. Knowledge and beliefs about epilepsy among people with and without epilepsy in urban Suriname. Epilepsy Behav 2013 Oct;29(1):128-132. [doi: 10.1016/j.yebeh.2013.07.001] [Medline: 23948656]

17. World Health Organization. Epilepsy URL: http://www.who.int/news-room/fact-sheets/detail/epilepsy [accessed 2019-12-11]

18. Hagemann A, Pfäfflin M, Nussbeck FW, May TW. The efficacy of an educational program for parents of children with epilepsy (FAMOSES): results of a controlled multicenter evaluation study. Epilepsy Behav 2016 Nov;64(Pt A):143-151 [FREE Full text] [doi: 10.1016/j.yebeh.2016.09.027] [Medline: 27744243]

19. Rodenburg R, Meijer AM, Scherphof C, Carpay JA, Augustijn P, Aldenkamp A, et al. Parenting and restrictions in childhood epilepsy. Epilepsy Behav 2013 Jun;27(3):497-503. [doi: 10.1016/j.yebeh.2013.01.026] [Medline: 23602224]

20. Masri A, Aburahma S, Khasawneh A, Al Qudah A, Nafi O, Al Momani M, et al. Parental knowledge and attitudes towards epilepsy - a study from Jordan. Seizure 2017 Dec;53:75-80 [FREE Full text] [doi: 10.1016/j.seizure.2017.11.006] [Medline: 29149668]

21. Gazibara T, Nikolovski J, Lakic A, Pekmezovic T, Kisic-Tepavcevic D. Parental knowledge, attitudes, and behaviors towards children with epilepsy in Belgrade (Serbia). Epilepsy Behav 2014 Dec;41:210-216. [doi: 10.1016/j.yebeh.2014.09.062] [Medline: 25461218]

22. Rodenburg R, Meijer AM, Deković M, Aldenkamp AP. Family factors and psychopathology in children with epilepsy: a literature review. Epilepsy Behav 2005 Jun;6(4):488-503. [doi: 10.1016/j.yebeh.2005.03.006] [Medline: 15907744]

23. Kolahi A, Abbasi-Kangevari M, Bakhshaei P, Mahvelati-Shamsabadi F, Tonekaboni S, Farsar A. Knowledge, attitudes, and practices among mothers of children with epilepsy: a study in a teaching hospital. Epilepsy Behav 2017 Apr;69:147-152. [doi: 10.1016/j.yebeh.2016.10.018] [Medline: 28285235]

24. Frank-Briggs A, Alikor EA. Knowledge and attitudes of parents toward children with epilepsy. Ann Afr Med 2011;10(3):238-242 [FREE Full text] [doi: 10.4103/1596-3519.84706] [Medline: 21912010]

25. Zainy LE, Atteyah DM, Aldisi WM, Abdulkarim HA, Alhelo RF, Alhelali HA, et al. Parents` knowledge and attitudes toward children with epilepsy. Neurosciences (Riyadh) 2013 Oct;18(4):345-348. [Medline: 24141457]

26. el-Hazmi MA, al-Swailem AR, Warsy AS, al-Swailem AM, Sulaimani R, al-Meshari AA. Consanguinity among the Saudi Arabian population. J Med Genet 1995 Aug;32(8):623-626 [FREE Full text] [doi: 10.1136/jmg.32.8.623] [Medline: 7473654]

27. Mbuba CK, Ngugi AK, Newton CR, Carter JA. The epilepsy treatment gap in developing countries: a systematic review of the magnitude, causes, and intervention strategies. Epilepsia 2008 Sep;49(9):1491-1503 [FREE Full text] [doi: 10.1111/j.1528-1167.2008.01693.x] [Medline: 18557778]

28. Akay AP, Hiz Kurul S, Ozek H, Cengizhan S, Emiroglu N, Ellidokuz H. Maternal reactions to a child with epilepsy: depression, anxiety, parental attitudes and family functions. Epilepsy Res 2011 Aug;95(3):213-220. [doi: 10.1016/j.eplepsyres.2011.03.020] [Medline: 21543187]
29. Turan F, Ythler A. The effect of a modular education program related to children with epilepsy and their parents on disease management. Pediatrics 2015;135(Supplement 1):S6-S7. [doi: 10.1542/peds.2014-3330j]
30. Joo B, Lee Y, Kim HD, Eom S. An analysis of a novel, short-term therapeutic psychoeducational program for children and adolescents with chronic neurological illness and their parents; feasibility and efficacy. Front Neurosci 2017;11:311 [FREE Full text] [doi: 10.3389/fnins.2017.00311] [Medline: 28620276]
31. Lua PL, Neni WS, Lee JK, Abd Aziz Z. The interactive animated epilepsy education programme (IAEEP): how feasible, acceptable and practical is the technology to children? Technol Health Care 2013;21(6):547-556. [doi: 10.3233/THC-130758] [Medline: 24284547]

Abbreviations

KAU: King Abdulaziz University

XSL•FO
RenderX

Original Paper

Web Comparison of Three Contingent Valuation Techniques in Women of Childbearing Age: The Case of Ovulation Induction in Quebec

Aissata Dieng[1], MSc; Jie He[1], PhD; Thomas G Poder[2,3], PhD

[1]University of Sherbrooke, Deptartment of Economics, Sherbrooke, QC, Canada

[2]University of Montréal, School of Public Health, Montréal, QC, Canada

[3]Centre de recherche de l'Institut universitaire en santé mentale de Montréal, Montréal, QC, Canada

Corresponding Author:
Thomas G Poder, PhD
University of Montréal
School of Public Health
7101 Avenue du Parc
Montréal, QC, H3N 1X9
Canada
Phone: 1 8198613680
Email: thomas.poder@umontreal.ca

Abstract

Background: In Canada, 11.5% to 15.7% of couples suffer from infertility. Anovulation, or failed ovulation, is one of the main causes of infertility in women. In Quebec, the treatment for ovulation induction and other services related to assisted reproductive technology (ART) have been partially reimbursed by the government since 2010.

Objective: This study aimed to compare the willingness to pay (WTP) of women of childbearing age to receive drug treatment in the event of failed ovulation according to 3 different contingent valuation methods.

Methods: The following elicitation techniques were used: simple bid price dichotomous choice (DC), followed by an open-ended question (DC-OE), and a simplified multiple-bounded discrete choice (MBDC). Each participant was randomly assigned to 1 of 3 elicitation techniques. Bid prices ranged from Can $200 to Can $5000. Of the 7 bid prices, 1 was randomly proposed to each participant in the DC and DC-OE groups. For the DC-OE group, if the answer to the DC bid price was *no*, respondents were asked what was the maximum amount they were willing to pay. For the MBDC group, each respondent was offered an initial bid price of Can $1500, and the subsequent bid price offer increased or decreased according to the answer provided. "*Do not know*" responses were considered as a "*no*", and each individual was questioned as to their certainty after each choice. WT*P* values were estimated using probit and bivariate models; the Welsh and Poe model was also performed for the MBDC group.

Results: The survey was conducted from 2009 to 2010 with a total sample of 680 women. Analyses were performed on 610 respondents (199 DC, 230 DC-OE, and 181 MBDC). Of the 70 respondents who were excluded, 6 did not meet the age criterion, 45 had an annual income less than Can $2500, and 19 did not respond to the WTP question. Mean WT*P* values were Can $4033.26, Can $1857.90, and Can $1630.63 for DC, DC-OE, and MBDC, respectively. The WTP for MBDC "*definitely yes*" and "*probably yes*" values were Can $1516.73 and Can $1871.22, respectively. The 3 elicitation techniques provided WT*P* value differences that were statistically significant (*P*<.01). The MBDC was the most accurate method, with a lower confidence interval (Can $557) and a lower (CI/mean) ratio (0.34).

Conclusions: A positive WTP for ovulation induction was found in Quebec. Adding a follow-up question resulted in more accurate WT*P* values. The MBDC technique provided a more accurate estimate of the WTP with a smaller and, therefore, more efficient confidence interval. To help decision making and improve the effectiveness of the fiscal policy related to the ART program, the WT*P* value elicited with the MBDC technique should be used.

XSL•FO
RenderX

Introduction

Background

According to the World Health Organization, infertility is defined as the inability to conceive after 12 months of unprotected sex [1]. In the epidemiology of infertile couples, Brzakowski et al [2] state that failed ovulation affects a large number of couples around the world—approximately 80 million people or 1 in 10 couples. In Canada, 11.5% to 15.7% of couples suffer from infertility, according to the Institut national d'excellence en santé et en services sociaux [3]. According to the Association of Gynecologists and Obstetricians of Quebec, there is a decrease in fertility in 84% of couples, including 10% of infertility in women (eg, fallopian tube blockage on both sides) and 6% infertility in men (eg, no spermatozoa) [4]. Anovulation or abnormality of ovulation is one of the main causes of infertility in women. To counteract this issue, a drug treatment that aims to induce ovulation is needed. In Quebec, this treatment and other services of assisted reproductive technology (ART) have been partially reimbursed by the government since 2010 [1].

The benefit of the drug treatment to induce ovulation in adult female patients with infertility is generally measured by the proportion of women who ovulate as a result of such treatment [5]. It is difficult to measure the monetary benefit of this treatment without exchange value, which implies the problem of comparing the costs of this treatment with its *benefits*. One way to obtain a monetary value of this benefit is to estimate women's willingness to pay (WTP) for this infertility treatment.

Several methods can be used to estimate this monetary value, including the contingent valuation method (CVM) [6,7]. This method is increasingly used in health economics, as it provides a monetary value for nonmarket goods and services using a fictitious market [8]. The CVM consists in asking a hypothetical question using a variety of survey techniques (eg, telephone, face to face, internet, and postal mail) to measure the maximum amount that individuals would be willing to pay for something (in this case, the fertility treatment) and its consequent effects. This method offers several elicitation techniques that correspond to different ways of formulating the WTP question.

Although different variants exist, the 4 main techniques reported in the literature are bidding game (BG), payment card (PC), open-ended (OE) questions, and dichotomous choice (DC) [9,10]. The BG is the oldest elicitation technique [6]. The respondent is randomly assigned a particular bid from a range of predetermined bids. The respondent is then asked to say *yes* or *no* to that particular bid, and the process continues until the highest positive response is recorded [11]. The PC consists of presenting the respondent with a series of offers in a table in which the individual circles the amount corresponding to his or her WTP [12]. The OE consists in asking the respondent directly what is the maximum amount he or she would be willing to pay

for a given good or service [12,13]. In the DC approach, the respondent only answers *yes* or *no* to a given amount.

Objective

In this study, 3 CVMs were compared: DC, DC followed by an OE question (DC-OE), and a simplified multiple-bounded discrete choice (MBDC), which is very similar to a BG. These 3 methods were chosen because of their simplicity and because they are widely used in the literature. The main objective of this study was to assess the WTP of women of childbearing age to receive a drug treatment in the event of failed ovulation according to the 3 different CVMs. More specifically, this study aimed to assess whether these 3 techniques generate statistically different WTPs and, if so, to determine which method is the most accurate.

Methods

Study Design and Population

The data used in this study were from a survey conducted in Quebec between January 2009 and February 2010. Inclusion criteria required participants to be a woman aged 18 to 45 years and to agree to complete the survey in French. Women were excluded if they had an annual income less than or equal to Can $2500 (this amount corresponded to the middle of the lowest bracket proposed for annual income and because it is unlikely that women can afford an infertility treatment with this income) or if they did not respond to the WTP question. No sample size was calculated, but 200 patients per elicitation method were targeted, which is the usual number for this type of study [14].

Data Collection

The data were collected through a Web survey, which were first distributed using an email listing from previous studies (ie, respondents from previous studies who accepted to be contacted for future research) and were then distributed by a Web survey company. Participants were randomly allocated to 1 of the 3 elicitation methods tested (DC, DC-OE, or MBDC).

Questionnaire

Each questionnaire had 3 main components: introduction, socioeconomic variables, and WTP questions. The introduction presented a definition of infertility, gave the prevalence of infertility (including infertility related to ovulation failure), the type of treatment associated, the probability of success, and the associated risks. The socioeconomic variables included age (years), weight (kilograms), height (centimeters), employment status, is the job stressful (yes or no), individual annual income (using brackets), educational level, civil status, number of children, smoking (yes or no), general health (5 levels), fertility problems (yes or no), actually pregnant (yes or no), desire for a child (yes or no), and a ranking of 10 items (eg, have good health, have children, be financially comfortable).

The third component was the WTP question about receiving ovulation failure treatment, along with another question about

XSL·FO
RenderX

the degree of certainty of the respondent's answer. For DC and DC-OE, 7 price levels were randomly assigned to different versions of the questionnaire (Can $200, Can $500, Can $1000, Can $1500, Can $2000, Can $3000, and Can $5000). For the DC-OE, if the answer was *no* or *do not know*, respondents were asked to report the maximum amount they were willing to pay for this service. For the MBDC, the respondent was offered an initial WTP amount of Can $1500, where the possible answers were *yes*, *no*, or *do not know*. If the first answer was positive, the price increased to Can $3000; if it was negative or *do not know*, the price decreased to Can $500. Unlike a traditional MBDC, which uses a random start price, our approach used a predetermined starting price of Can $1500, and only 3 bids were possible to more quickly end the round of questioning. For all CVM, the *do not know* response was considered as a *no*, and a question about the certainty of the answer was asked to individuals after each choice (not at all certain, not certain, more or less certain, certain, and quite certain).

Data Analysis

Overall, 2 comparison criteria were used to judge the accuracy of the estimates. The obtained estimated WTP and standard deviations were compared between the 3 subsamples. The efficiency of the estimates was measured with the ratio of confidence interval on mean WTP. The expected efficiency associated with a follow-up WTP question is based on the fact that the confidence intervals should be narrower and closer to the mean WTP value [15]. Multivariate probit models were computed to estimate WTP for each method in considering only the *yes* or *no* responses. For the DC-OE, a bivariate probit model was used with the *yes* or *no* response and the OE response as the dependent variable. For MBDC, the Welsh and Poe model [15] was computed for *definitely yes*, *probably yes*, and *do not know* responses.

To calculate the mean WTP value of the probit models, the coefficients of each variable were multiplied by the mean value of the total sample and divided by the bid coefficient and the bid mean (thus, WTP equals $\sum(\beta_j * \mu_j)$ divided by $(\beta_{bid} * \mu_{bid})$,

where β_j=coefficient of variable j and μ_j=mean of variable j). Considering the mean value of the total sample allowed to better consider differences in WTP values associated with the 3 elicitation techniques and to reduce the effect of socioeconomic differences observed in the subsamples. For the bivariate model, the WTP was directly obtained by multiplying the coefficients of the equation by the mean value of each variable, again with the mean value of the total sample. WTP was also estimated using the parametric bootstrap method developed by Krinsky and Robb [16], with 1000 repetitions. This method consists of making a large number of draws from a multivariate normal distribution with the means and the variance-covariance matrix of the estimated parameters. The different simulated WTP values were calculated from the joint distribution of the coefficients. This method gives precise confidence intervals. Details of the WTP estimation methods are presented in Multimedia Appendix 1 [17-20].

Results

Patient Characteristics

The total sample consisted of 680 women; of these women, 215 responded to the DC, 255 responded to the DC-OE, and 210 responded to the MBDC (Figure 1). A total of 70 respondents were excluded: 6 did not meet the age criterion, 45 had an annual income less than Can $2500, and 19 did not respond to the WTP question. Analyses were thus conducted on 610 respondents (199 DC, 230 DC-OE, and 181 MBDC). Although the CVM groups were randomly distributed, they differed significantly for a number of variables. The respondents answering the DC questionnaire were older, had higher annual income, and were in better health. Those answering the DC-OE questionnaire had fewer problems with infertility, but among those, the percentage of failed ovulation was higher. The people invited to answer the MBDC questionnaire had a lower educational level, were mostly smokers, had less stress, and were less often employed (Table 1).

Figure 1. Flowchart of respondents randomly distributed among the 3 elicited methods (dichotomous choice, dichotomous choice followed by an open-ended question, or multiple-bounded discrete choice).
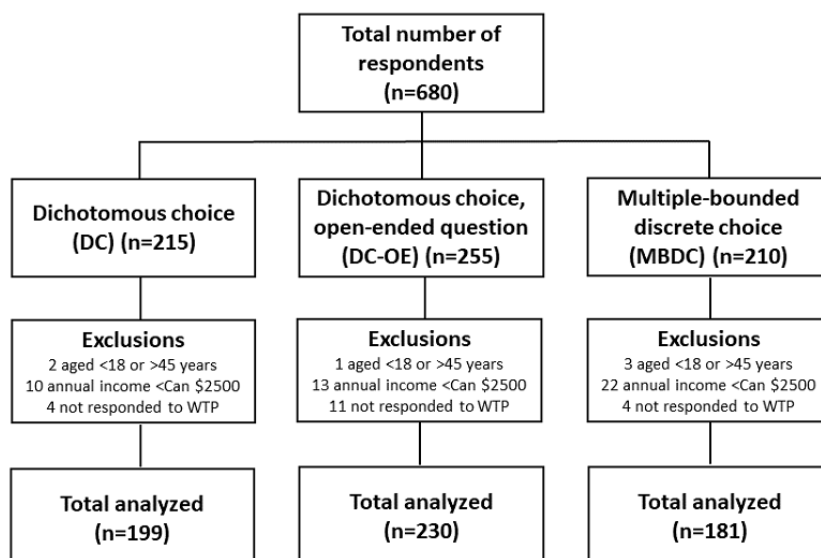
Table 1. Descriptive statistics.

Variable	Dichotomous choice (n=199)	Dichotomous choice followed by an open-ended question (n=230)	Multiple-bounded discrete choice (n=181)
Age (years), mean (range)	32 (18-45)	30 (18-45)[a]	30 (18-45)[a]
Annual income (Can $), mean (range)	42,622 (7500-130,000)	37,615 (7500-130,000)[b]	34,475 (7452-130,000)[a,c]
Schooling, n (%)			
Secondary school	31 (15.6)	41 (17.8)	42 (23.2)[b,d]
College	78 (39.2)	93 (40.4)	63 (34.8)
University	90 (45.2)	96 (41.7)	76 (42.0)
Very good health, n (%)	58 (29.2)	43 (18.7)[a]	36 (19.9)[a]
Current smoker, n (%)	39 (19.6)	45 (19.6)	51 (28.2)[a,e]
Stressful job, n (%)	174 (87.4)	198 (86.1)[a]	145 (80.1)[c]
Ovulation failure, n (%)	7 (3.5)	14 (6.1)[a]	5 (2.8)[e]
Infertility problem, n (%)	26 (13.1)	25 (10.9)[a]	22 (12.2)[e]
Having a child is important, n (%)	125 (62.8)	147 (63.9)	108 (59.7)
Employee, n (%)	158 (79.4)	183 (79.6)	130 (71.8)[b,e]

[a]$P<.01$ (comparison of dichotomous choice versus dichotomous choice followed by an open-ended question and dichotomous choice versus multiple-bounded discrete choice).

[b]$P<.05$ (comparison of dichotomous choice versus dichotomous choice followed by an open-ended question and dichotomous choice versus multiple-bounded discrete choice).

[c]$P<.05$ (comparison of dichotomous choice followed by an open-ended question versus multiple-bounded discrete choice).

[d]$P<.1$ (comparison of dichotomous choice followed by an open-ended question versus multiple-bounded discrete choice).

[e]$P<.01$ (comparison of dichotomous choice followed by an open-ended question versus multiple-bounded discrete choice).

Responses to Willingness to Pay Questions

As we expected, the higher the offered bid price, the lower the proportion of *yes* answers. The percentage of *yes* answers for the lowest value (Can $200) was 86% (12/14) for DC, 77% for DC-OE (26/34), and 100.0% (181/181) for MBDC. For the highest price (Can $5000), the percentage of *yes* answers was 26% (10/39) for DC, 53% (10/19) for DC-OE, and 12.7% (23/181) for MBDC. It should be noted that the decrease was gradual with DC and MBDC but more stable with DC-OE. When the cumulative decreasing frequencies of positive responses were analyzed, the DC and DC-OE had similar distributions, but the MBDC decreased more rapidly (Figure 2).

Figure 2. Cumulative decreasing frequencies of positive answers yes. DC: dichotomous choice; DC-OE: dichotomous choice followed by an open-ended question; MBDC: multiple-bounded discrete choice.
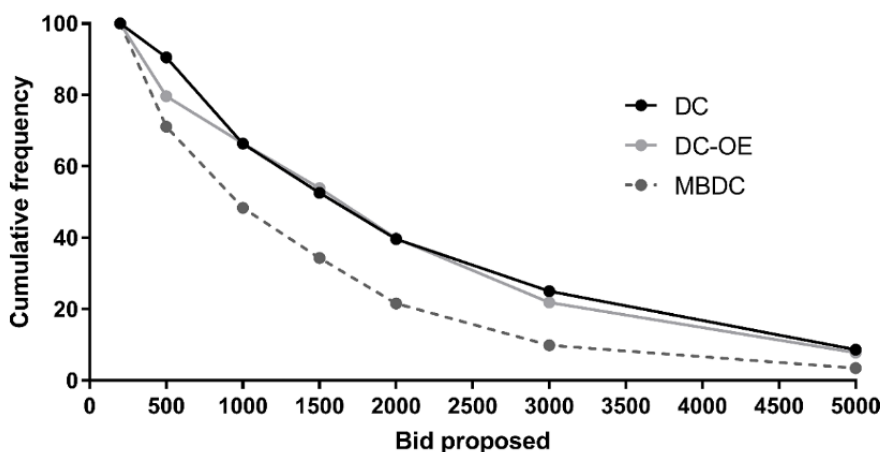
Willingness to Pay Estimated With Dichotomous Choice

The results of the probit analysis are presented in Table 2. The higher the offer (bid), the lower the probability to say yes to the WTP question ($P<.01$). The contribution of each explanatory variable to the WTP was calculated with the ratio of coefficients ($-\beta_{variable}/\beta_{bid}$). For example, women with a university education were willing to pay Can $2338 ($-.788/-.000337$) more than women without a university education ($P<.01$). Women with very good health, with a stressful job, or who considered having a child to be important were also willing to pay more (Can $1772, $P<.05$; Can $593, $P<.10$; and Can $1395, $P<.05$, respectively).

Table 2. Estimation results with probit and bivariate analysis.

Variables	Dichotomous choice		Dichotomous choice followed by an open-ended question				Multiple-bounded discrete choice	
	Probit (R^2=0.2375; N=199)		Probit (R^2=0.0863; N=230)		Bivariate (R^2=0.0633; N=230)		Probit (R^2=0.2807; N=1245)	
	Coefficient	T test (187)	Coefficient	T test (218)	Coefficient	T test (218)	Coefficient	T test (1233)
Bid	−0.000337[a]	−4.71[b]	−0.000111	−1.67[c]	—[d]	—	−0.000573	−17.17[b]
Age (years)	−0.0264	−1.54	−0.0171	−1.15	−11.50	−0.52	−0.0279	−4.43[b]
Annual income	0.00000310	0.69	0.00000509	1.00	0.00749	0.97	0.00000395	1.77[c]
University	0.788	3.34[b]	0.423	2.08[e]	523.9	1.69[c]	0.343	3.61[b]
Very good health	0.597	2.42[e]	−0.0571	−0.24	−110.2	−0.31	0.0157	0.15
Infertility problem	−0.165	−0.43	−0.113	−1.07	−170.3	−1.08	−0.202	−1.36
Ovulation failure	0.728	1.03	0.0395	0.29	−34.94	−0.17	0.384	1.39
Current smoker	0.401	1.49	−0.388	−1.62	−602.2	−1.72[c]	0.422	4.53[b]
Stressful job	0.200	1.81[c]	−0.0770	−0.30	83.55	0.21	−0.0546	−0.49
Having a child is important	0.470	2.1[e]	0.312	1.63	209.1	0.73	0.554	6.40[b]
Employee	0.421	1.63	0.242	0.99	207.8	0.58	0.185	1.88[c]
Constant	0.149	0.25	0.280	0.48	2218.5	2.60[b]	0.458	1.74[c]
Sigma	—	—	—	—	1583.9	13.34[b]	—	—

[a]Italics indicate that the term is statistically significant.

[b]$P<.01$.

[c]$P<.1$.

[d]Not applicable.

[e]$P<.05$.

Willingness to Pay Estimated With Dichotomous Choice Followed by an Open-Ended Question

The coefficients of the explanatory variables of the bivariate model directly illustrate women's WTP. For DC-OE, when the offer (bid) was higher, the WTP was significantly lower ($P<.1$). Women with a university education were willing to pay Can $3811 more in the probit model ($P<.05$) and Can $524 more in the bivariate model ($P<.1$). Women who smoked were willing to pay Can $602 less than other women ($P<.1$).

Willingness to Pay Estimated With Multiple-Bounded Discrete Choice

The probit model results are presented in Table 2, and the results using the model by Welsh and Poe are presented in Table 3. As for the DC and DC-OE methods, if the offer (bid) was higher, the WTP was significantly lower ($P<.01$). The probit model for MBDC had more significant variables that explain the WTP. Older women were willing to pay Can $48.69 ($P<.01$) and Can $82.50 ($P<.1$) less than other women per additional year in the probit model and in the *probably yes* model by Welsh and Poe, respectively. Women with higher incomes were willing to pay more than other women ($P<.1$) in the probit model. Women with a university education were willing to pay Can $599 more in the probit model ($P<.01$) and Can $1058 more in the *probably yes* model ($P<.1$). Women who smoked were willing to pay Can $736 ($P<.01$) more in the probit model and Can $941 ($P<.01$) more in the *do not know* model. Women for whom having a child was a priority were willing to pay more than others at an amount of Can $967 ($P<.01$) and Can $581 ($P<.1$) in the probit and *do not know* models, respectively. Having a job increased women's WTP by Can $323 compared with other women in the probit model.

Table 3. Estimation results using the model by Welsh and Poe.

Variables	Definitely yes (N=181)[a]		Probably yes (N=181)[a]		Do not know (N=181)[a]	
	Estimate	T test (169)	Estimate	T test (169)	Estimate	T test (169)
Equation 1						
Constant	*−0.000626*[b]	*−14.08*[c]	*−0.000343*	*−9.68*[c]	*−0.000542*	*−12.56*[c]
Equation 2						
Age (years)	−0.00412	−0.31	*−0.0283*	*−2.13*[d]	−0.0133	−1.05
Annual income	0.00000230	0.48	0.00000126	0.27	0.00000130	0.28
University	0.271	1.42	*0.363*	*1.86*[d]	0.280	1.44
Very good health	0.0331	0.16	0.0661	0.30	−0.115	−0.54
Infertility problem	−0.175	−0.59	−0.192	−0.62	−0.130	−0.41
Ovulation failure	0.460	0.74	0.309	0.51	0.135	0.23
Current smoker	0.213	1.12	0.218	1.12	*0.510*	*2.66*[c]
Stressful job	−0.0288	−0.12	0.177	0.74	−0.175	−0.74
Having a child is important	0.128	0.72	0.279	1.54	*0.315*	*1.76*[e]
Employee	−0.156	−0.77	−0.0866	−0.42	0.0772	0.38
Constant	0.558	1.01	0.778	1.41	*1.155*	*2.16*[d]

[a]P value<.001.

[b]Italics indicate that the term is statistically significant.

[c]$P<.01$.

[d]$P<.05$.

[e]$P<.1$.

Mean Willingness to Pay Estimated

Table 4 reports the mean WTP for each subsample and their confident intervals obtained with Krinsky and Robb's [16] bootstrap method. As shown, women were, in general, willing to pay for an ovulation failure treatment an average of Can $4033.26 in the DC questionnaire, Can $1857.90 in the DC-OE questionnaire, and Can $1630.63 in the MBDC questionnaire. The mean WTPs for MBDC definitely yes and probably yes were Can $1516.73 and Can $1871.22, respectively. A Student t test revealed a statistically significant difference among the mean WTPs obtained from the DC, DC-OE, and MBDC subsamples (all $P<.01$). The MBDC method can be considered

the most accurate, with the lowest confidence interval (896.51) and the lowest (CI/mean) ratio (0.53). The DC-OE method had a confidence interval higher than MBDC and a CI/mean ratio of 1.04. The least accurate approach was the DC method.

Comparing DC with DC-OE, we can see that adding 1 more question after the DC WTP question improves the accuracy of the WTP estimates. However, our results also revealed the existence of an anchoring effect in the DC-OE approach, where the implicit WTP values of the respondents were influenced by the first proposed bid price. With the Herriges and Shogren's model [21], the gamma coefficient (SE) was 0.7402 (0.0166), and the 95% CI was from 0.7074 to 0.7729 ($P<.001$).

Table 4. Mean willingness to pay estimated with probit, bivariate, or Krinsky and Robb methods.

Method	Average WTP[a]	Mean difference (95% CI)	CI/WTP
Probit and bivariate methods			
DC[b] (N=199)	4033.26	4386.01 (1840.25 to 6226.26)	1.09
DC-OE[c] (N=230)	1857.90	1945.1 (885.35 to 2830.45)	1.05
MBDC[d]—probit (N=1245)	1630.63	556.64 (1352.31 to 1908.95)	0.34
MBDC—definitely yes (N=181)	1516.73	2796.86 (118.30 to 2915.16)	1.84
MBDC—probably yes (N=181)	1871.22	4926.29 (−591.92 to 4334.37)	2.63
MBDC—do not know (N=181)	2514.49	3032.39 (998.29 to 4030.68)	1.21
Krinsky and Robb's methods			
DC (N=199)	4750.18	6985.57 (3911.67 to 10,897.24)	1.47
DC-OE (N=230)	1857.90	1924.93 (895.44 to 2820.37)	1.04
MBDC—probit (N=1245)	1701.37	896.51 (1103.43 to 1999.94)	0.53

[a]WTP: willingness to pay.

[b]DC: dichotomous choice.

[c]DC-OE: dichotomous choice followed by an open-ended question.

[d]MBDC: multiple-bounded discrete choice.

Discussion

Principal Findings

A total of 3 elicitation techniques were used to assess women's WTP for an ovulation induction treatment in case of failed ovulation. One of the main objectives was to discover whether a significant difference exists between different WTP elicitation approaches.

The results show that the DC technique yielded higher estimated WTP than the other 2 techniques. The higher value for WTP with DC methods is consistent with the literature [22-26]. In Welsh and Poe's study [26], they concluded that the WTP obtained by the DC technique was higher than that obtained by the *not sure* model. In our study, we also find that the DC WTP was statistically larger than the Welsh and Poe *not sure* model, and the DC-OE WTP did not statistically differ from the *probably yes* model of Welsh and Poe. On the contrary, the WTP of the MBDC method was between the *definitely yes* and the *probably yes* models of Welsh and Poe. The comparison of DC and DC-OE was also consistent with the findings of Hanneman et al [15], who used a bivariate model to compare estimates of DC and double-bounded DC. They found that the double-bounded model reduced the variance of the estimated parameters and decreased the covariance terms. They concluded that the double-bounded DC model was more efficient after correcting for the anchoring effect.

The value added by a follow-up question is based on the fact that the confidence intervals are closer to the estimated WTP and that the latter is, therefore, more accurate [15]. This is the case in our study, where the WTP estimate with the DC-OE model was more accurate than the DC approach. By comparing the confidence intervals and standard deviations of the different techniques, our results show that the MBDC technique gave lower mean WTPs and smaller standard deviations than the

other 2 techniques. Therefore, based on efficiency as the criterion of comparison (ie, the ratio of the confidence interval to the mean WTP [16]), the MBDC technique is preferable. The DC-OE gave a confidence interval that was wider than that of MBDC but still lower than that of DC. Our results are similar to the study by Scarpa and Bateman [27], where the authors concluded that MBDC WTPs are more efficient and that including one additional question in a contingent valuation survey improves the effectiveness of the WTP, although biases caused by a potential anchoring effect are likely to occur.

The estimated WTP in our study shows dissimilar results to the study by Poder et al [5] about failed ovulation. In their study, they found that the mean WTP for a medical treatment for ovulation induction was Can $3400 CAD in the DC technique, where *do not know* answers were considered as a *no*. We found a higher WTP in our DC database (Can $4033). This difference may be because of the mode of collection they used (paper and Web), their higher number of observations (327 vs 215 subjects), or the sociodemographic characteristics of their sample. However, what is consistent in these studies is that women have a positive WTP for infertility treatment. In our specific study about ovulation induction, the standard treatment is to administer clomiphene citrate over a 6-month period. This specific treatment can be done at a very low cost (less than Can $500) when compared with the WT*P* value found. This indicates that the social value of infertility treatment is highly valued by women and that to invest in it is worth it.

Our study gave coefficients of the expected signs, although the positive coefficient of the variable *income* was not significant in the 3 techniques. This result suggests that women's responses were independent of their income. One explanation for this is that infertility is of major importance in their lives, regardless of income. A similar result was found in the study by Poder et al [5]. Moreover, the negative coefficient on the variable of age

implies that older women place less importance on care for ovulation failure, perhaps because women's fertility decreases with age. The results of our different regressions also concur with the predictions of economic empirical theories, which state that women's WTP decreases with age [28,29].

This study has a number of limitations, so the results should be interpreted with caution. As we used a convenience sample, we cannot say with certainty that our regression equations will give the same results if applied to a larger or different sample because of the lack of representativeness. Another limitation of this study is that our approach used a fixed predetermined starting price of Can $1500 in the MBDC technique. This choice may have led to an anchoring effect, as individuals focus on the first proposition (Can $1500), and thus, their answers to the second and third questions may be influenced by the first bid offered. Unlike other techniques (DC and DC-OE) that use random starting prices between Can $200 and Can $5000, this anchoring effect cannot be assessed in the MBDC.

Each of the 3 elicitation techniques has its disadvantages. The DC technique yielded higher estimated WTP with little WTP information (ie, only 1 WTP question, so we only know if their maximum WTP is higher or lower to the bid proposed). However, the DC technique is more similar to the real market situation of *take it or leave it* [22]. Although the DC technique followed by an OE question provides more information for those answering *no* or *do not know*, it does not add information for those answering *yes* to the first question (ie, we do not know the maximum offer that would be accepted); moreover, the responses to the second question can introduce the possibility of strategic behavior on the part of respondents. Respondents may feel that giving a positive WTP to the second question may allow the government to increase their claims but answering *zero* to the second question could be because of the impression that the quality of the service offered may be reduced. Furthermore, a high zero-value rate (ie, many zeros) and an anchoring effect occur in the DC-OE technique.

Conclusions

In their study on psychosocial services for couples in infertility treatment, Read et al [30] reported that infertility is associated with considerable distress, and treatment is often characterized by cycles of hope and disappointment. Regardless of age, failed ovulation is the most common cause of infertility in women; today, it can be treated with fertility drugs [5]. In this study, the goal was to test whether an elicitation technique may have an effect on the estimation of WTP for women of childbearing age for a failed ovulation treatment service. The data from the 3 techniques reveal that women with a higher level of education placed more importance on the treatment of failed ovulation than other women. We also note that in the MBDC technique, the lowest bid price offered (Can $200) was accepted by all respondents. Thus, infertility treatment is seen as having a positive value.

We also compared the mean WTPs of the different techniques and found significant differences among the estimated WTPs. Adding a follow-up question resulted in more accurate WTPs but created anchoring biases. Results also indicated that the simplified MBDC technique provided more accurate estimates of the WTP with a smaller and, therefore, more efficient confidence interval. Consequently, for the purpose of a more efficient fiscal policy, the simplified MBDC technique provided the most appropriate WT*P* value.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Estimates of willingness to pay.
[DOCX File , 19 KB - ijmr_v9i1e13355_app1.docx]

References

1. Zegers-Hochschild F, Adamson GD, de Mouzon J, Ishihara O, Mansour R, Nygren K, International Committee for Monitoring Assisted Reproductive Technology, World Health Organization. International Committee for Monitoring Assisted Reproductive Technology (ICMART) and the World Health Organization (WHO) revised glossary of ART terminology, 2009. Fertil Steril 2009 Nov;92(5):1520-1524. [doi: 10.1016/j.fertnstert.2009.09.009] [Medline: 19828144]
2. Brzakowski M, Lourdel E, Cabry R, Oliéric M, Claeys C, Devaux A, et al. Epidemiology of the infertile couple. J Obstet Gynecol Reprod Biol 2009;38:F3-F7. [doi: 10.1016/S0368-2315(09)70226-1]
3. Institut national d'excellence en santé et en services sociaux (INESSS). Assisted Human Reproduction - Medical Criteria for Eligibility for Publicly Funded Treatments and Safety of Repetition in In Vitro Fertilization Cycles. Québec, Canada: Bibliothèque et Archives; 2015.

XSL•FO
RenderX

4.   AOGQ: Association of Obstetricians and Gynecologists of Quebec. Infertility URL: http://www.gynecoquebec.com/sante-femme/infertilite/20-infertilite.html [accessed 2019-12-17]

5.   Poder TG, He J, Simard C, Pasquier JC. Willingness to pay for ovulation induction treatment in case of WHO II anovulation: a study using the contingent valuation method. Patient Prefer Adherence 2014;8:1337-1346 [FREE Full text] [doi: 10.2147/PPA.S67742] [Medline: 25328385]

6.   Mitchell RC, Carson RT. Using Surveys to Value Public Goods: The Contingent Valuation Method. Washington, DC: Resources for the Future; 1989.

7.   Hanemann WM. Valuing the environment through contingent valuation. J Econ Perspect 1994;8(4):19-43. [doi: 10.1257/jep.8.4.19]

8.   Ryan M, Scott DA, Donaldson C. Valuing health care using willingness to pay: a comparison of the payment card and dichotomous choice methods. J Health Econ 2004 Mar;23(2):237-258. [doi: 10.1016/j.jhealeco.2003.09.003] [Medline: 15019754]

9.   Boyle KJ, Johnson FR, McCollum DW, Desvousges WH, Dunford RW, Hudson SP. Valuing public goods: discrete versus continuous contingent-valuation responses. Land Econ 1996;72(3):381-396. [doi: 10.2307/3147204]

10. Venkatachalam L. The contingent valuation method: a review. Environ Impact Assess Rev 2004;24(1):89-124. [doi: 10.1016/S0195-9255(03)00138-0]

11. Randall A, Ives B, Eastman C. Bidding games for valuation of aesthetic environmental improvements. J Environ Econ Manage 1974;1(2):132-149. [doi: 10.1016/0095-0696(74)90010-2]

12. Reaves D, Kramer R, Holmes T. Does question format matter? Valuing an endangered species. Environ Resour Econ 1999;14(3):365-383. [doi: 10.1023/A:1008320621720]

13. Walsh RG, Loomis JB, Gillman RA. Valuing option, existence, and bequest demands for wilderness. Land Econ 1984;60(1):14-29. [doi: 10.2307/3146089]

14. Lin P, Cangelosi MJ, Lee DW, Neumann PJ. Willingness to pay for diagnostic technologies: a review of the contingent valuation literature. Value Health 2013;16(5):797-805 [FREE Full text] [doi: 10.1016/j.jval.2013.04.005] [Medline: 23947973]

15. Hanemann M, Loomis J, Kanninen B. Statistical efficiency of double-bounded dichotomous choice contingent valuation. Am J Agric Econ 1991;73(4):1255-1263. [doi: 10.2307/1242453]

16. Krinsky I, Robb AL. On approximating the statistical properties of elasticities. Rev Econ Stat 1986;68(4):715-719. [doi: 10.2307/1924536]

17. Alberini A, Kanninen B, Carson RT. Modeling response incentive effects in dichotomous choice contingent valuation data. Land Econ 1997;73(3):309. [doi: 10.2307/3147170]

18. Carson RT, Flores NE, Martin KM, Wright JL. Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods. Land Econ 1996;72(1):80-99. [doi: 10.2307/3147159]

19. Cameron TA, Huppert DD. OLS versus ML estimation of non-market resource values with payment card interval data. J Environ Econ Manage 1989;17(3):230-246. [doi: 10.1016/0095-0696(89)90018-1]

20. Wang H, He J. Estimating individual valuation distributions with multiple bounded discrete choice data. Appl Econ 2011;43(21):2641-2656. [doi: 10.1080/00036840903299789]

21. Herriges JA, Shogren JF. Starting point bias in dichotomous choice valuation with follow-up questioning. J Environ Econ Manage 1996;30(1):112-131. [doi: 10.1006/jeem.1996.0008]

22. Johannesson M, Jönsson B. Economic evaluation in health care: is there a role for cost-benefit analysis? Health Policy 1991;17(1):1-23. [doi: 10.1016/0168-8510(91)90114-D]

23. Seller C, Stoll JR, Chavas JP. Validation of empirical measures of welfare change: a comparison of nonmarket techniques. Land Econ 1985;61(2):156-175. [doi: 10.2307/3145808]

24. Johnson RL, Bregenzer NS, Shelby B. Contingent valuation question Foryats: dichotomous choice versus open-ended responses. In: Johnson RL, Johnson GV, editors. Economic Valuation Of Natural Resources: Issues, Theory, And Applications. Boulder, CO: Westview Press, Chapter 12; 1990.

25. Desvousges WH, Johnson FR, Dunford RW, Boyle KJ, Hudson SP, Wilson KN. Chapter III - measuring natural resource damages with contingent valuation: tests of validity and reliability. In: Hausman JA, editor. Contributions to Economic Analysis. New York: North Holland; 1993.

26. Welsh MP, Poe GL. Elicitation effects in contingent valuation: comparisons to a multiple bounded discrete choice approach. J Environ Econ Manage 1998;36(2):170-185. [doi: 10.1006/jeem.1998.1043]

27. Scarpa R, Bateman I. Efficiency gains afforded by improved bid design versus follow-up valuation questions in discrete-choice CV studies. Land Econ 2000;76(2):299-311. [doi: 10.2307/3147230]

28. Alberini A, Cropper M, Krupnick A, Simon NB. Does the value of a statistical life vary with age and health status? Evidence from the US and Canada. J Environ Econ Manage 2004;48(1):769-792. [doi: 10.1016/j.jeem.2003.10.005]

29. Carson RT, Flores NA. eScholarship. 2000. Contingent Valuation: Controversies and Evidence URL: https://escholarship.org/uc/item/75k752s7 [accessed 2019-01-07]

30. Read SC, Carrier ME, Boucher ME, Whitley R, Bond S, Zelkowitz P. Psychosocial services for couples in infertility treatment: what do couples really want? Patient Educ Couns 2014 Mar;94(3):390-395. [doi: 10.1016/j.pec.2013.10.025] [Medline: 24290241]

Abbreviations

ART: assisted reproductive technology
BG: bidding game
CVM: contingent valuation method
DC: dichotomous choice
DC-OE: dichotomous choice followed by an open-ended question
MBDC: multiple-bounded discrete choice
PC: payment card
OE: open-ended
WTP: willingness to pay

Original Paper

A Strategic Imperative for Promoting Hospital Branding: Analysis of Outcome Indicators

Gow-Jen Shieh[1], PhD; Shi-Liang Wu[1], MS; Che-Fu Tsai[1], MS; Chi-Sen Chang[1], MD, PhD; Tsung-Hung Chang[1], MD; Ping-Wing Lui[1], MD, PhD; Yuh Yao[1], MS; Wayne Huey-Herng Sheu[1,2,3,4], MD, PhD

[1]Department of Top Hospital Administration, Taichung Veterans General Hospital, Taichung, Taiwan

[2]Department of Medicine, School of Medicine, National Yang-Ming University, Taipei, Taiwan

[3]School of Medicine, National Defense Medical Center, Taipei, Taiwan

[4]Institute of Medical Technology, College of Life Science, National Chung-Hsing University, Taichung, Taiwan

Corresponding Author:
Wayne Huey-Herng Sheu, MD, PhD
Department of Top Hospital Administration
Taichung Veterans General Hospital
No 1650, Section 4, Taiwan Boulevard
Taichung,
Taiwan
Phone: 886 4 2359 2525 ext 2005
Fax: 886 4 2359 2728
Email: whhsheu@vghtc.gov.tw

Abstract

Background: Optimizing the use of social media to promote hospital branding is important in the present digital era. In Taiwan, only 51.1% of hospitals have official Facebook fan pages. The numbers of likes for these hospitals are also relatively low.

Objective: Our objective was to establish a special branding team for social media operation, led by top administrators of our hospital. Here we present our strategic imperative for promoting hospital branding as well as an analysis of its effectiveness.

Methods: Led by top administrators, the branding team was formed by 11 divisions to create branding strategies. From 2016 to 2018, the team implemented action plans. All information unique to the hospital was posted on Facebook, as well as on the hospital's official website. To determine the plans' efficiencies, we obtained reference data from Google Analytics, and we compared Facebook Insights reports for 2016 with those for 2017 and 2018.

Results: One of the branding team's main missions was to establish branding strategies and to integrate segmental branding messages. In each quarter we regularly monitored a total of 52 action plan indicators, including those for process and outcome, and discussed the results at team meetings. We selected 4 main performance outcome indicators to reflect the effectiveness of the branding efforts. Compared with 2016, the numbers of likes posted on the Facebook fan page increased by 61.2% in 2017 and 116.2% in 2018. Similarly, visits to the hospital website increased by 4.8% in 2017 and 33.1% in 2018. Most Facebook fan page and website viewers were in 2 age groups: 25 to 34 years, and 35 to 44 years. Women constituted 60.42% (14,160/23,436) of Facebook fans and 59.39% (778,992/1,311,605) of website viewers. According to the Facebook Insights reports, the number of likes and post sharing both increased in 2017 and 2018, relative to 2016. Comment messages also increased from 2016 to 2018 ($P=.02$ for the trend). The most common theme of posts varied over time, from media reports in 2016, to innovative services in both 2017 and 2018. Likes for innovative services posts increased from 2016 through 2018 ($P=.045$ for the trend). By the end of 2018, we recorded 23,436 cumulative likes for posts, the highest number among medical centers in Taiwan.

Conclusions: We achieved the largest number of Facebook fans among all medical centers in Taiwan. We would like to share our experience with other hospitals that might be interested in engaging in social media for future communications and interactions with their patients.

XSL•FO
RenderX

Introduction

Background

Branding plays an important role because a positive brand helps customers visualize and understand the products. A favorable brand image leads to positive outcomes in customer satisfaction, service quality, loyalty, and repurchasing intention [1-3]. Branding of health care is imperative in business marketing, especially in this digital era when most people search information daily on the internet. The International Telecommunication Union in its report on information and communication technologies for 2017 revealed that mobile broadband subscriptions grew by more than 20% during 2012 to 2017, with an expected global growth up to 4.3 billion users by the end of 2017 [4]. Social media apps are becoming increasingly popular because of the ease of sharing and disseminating information without barriers, both in time and in space [5]. Through social media, people can easily obtain information on health care services [6].

Consequently, the hospital corporate brand is increasingly focused on communications in social media. Digital marketing strategies need to be considered first. In an online world, corporate branding, and the brand experience, lives on the internet [7]. A Pew Research Center survey showed that 60% of US adults use the internet to search health care information, and 10% of them use social media to follow health care experiences shared by friends [8]. Another study reported that 34% of health scholars use social media, including online forums and message boards, to obtain information on health and wellness [9]. Therefore, the use of social media in health care will likely grow exponentially [10].

Increasingly more patients use social network sites to share their experiences with health care personnel and institutions [11]. People share their experiences and social support in particular with their families and friends via social media [12-16]. The main advantages of using social media in health care services, as perceived by patients, are the potential to improve doctor-patient communication, increase their understanding of health-related issues, and the ability to share experiences with other patients having similar health conditions [17]. Also, the ability to respond to patients' needs and display timely messages on social media platforms means that hospitals could harness existing networks [18].

According to a survey on broadband internet usage in Taiwan, 89.4% of the 3155 respondents (extrapolated to Taiwan's total of 18.81 million residents) had used the internet. Among 89.8% of social network or instant message service users, 75.6% used both of these services. Internet World Stats reported that Facebook has reached a penetration of 75.8% in Taiwan, the highest proportion of users in the world [19]. Another survey found that Facebook was also the most popular social medium (90.9%) among internet users, followed by LINE (a freeware app for instant communications on electronic devices; 87.1%), YouTube (60.4%), PTT (a terminal-based bulletin board system; 37.8%), and Instagram (32.7%) [20]. Therefore, we speculated that by deploying official Facebook fan pages, hospitals should be able to improve their exposure in the community, promote

their reputations, and foster a better impression of these institutions. All of these can help ensure patient loyalty and recruit more patients.

Health care in Taiwan is managed centrally by the Bureau of National Health Insurance. By the end of 2017, the total number of insured people was 23.88 million, and the national health insurance coverage rate hovered around 99.7%. Also, 21,326 (92.8%) of the medical institutions in Taiwan had signed contracts with the National Health Insurance Administration of the Ministry of Health and Welfare. As of 2017, accreditation had been granted to 423 hospitals and 131 teaching hospitals. Based on the levels of accreditation, medical institutions are classified into 3 major categories: medical centers, regional hospitals, and local hospitals [21].

Prior to 2016, most of the hospital branding in Taiwan was focused on press conferences and press releases. Most hospitals assigned a single department to handle press conferences. Only 51.1% (n=213) of the hospitals had official Facebook fan pages as of 2017. Among them, academic medical centers tended to receive more likes than regional and local community hospitals [22]. The public sector receives relatively fewer resources from the government despite bearing a heavier burden than the private sector. Personnel and purchasing systems in public institutions are also less flexible, making them less competitive than private hospitals in providing medical services. Thus, how to restore or even to promote public recognition becomes a major issue for public hospitals.

Objective

The aim of this study was to support the strategic imperative for promoting hospital branding by establishing a special branding team, led by top administrators of our hospital. Through constant monitoring of internet posts, we aimed to determine what audiences like and to continuously adjust the popular and useful types of posts on the hospital Facebook page and hospital website. The implementation of such cycles of monitoring and improvement, while successfully conducted in private businesses, has rarely been conducted in hospitals, and in particular in public hospitals, both in Taiwan and worldwide.

Here, we share our experience with other hospitals that might be interested in engaging in social media and using them for communications and interactions with their patients in the future.

Methods

Study Design

We conducted this study to support our strategic imperative for promoting hospital branding by establishing a special branding team, led by top administrators of the hospital. To ensure the effectiveness of this branding team, we obtained reference data from Google Analytics (Google LLC, Mountain View, CA, USA) and Insights reports from Facebook (Facebook, Inc, Menlo Park, CA, USA) for the year 2016 as the baseline for comparisons with data for 2017 and 2018.

Strengthening Functions of Social Media

Study Setting

Taichung Veterans General Hospital is one of 19 medical centers in Taiwan. It was established in 1982, and in central Taiwan it is the only government medical center providing medical services to the public. It has 1569 beds, with high daily volumes of inpatient and outpatient services. Reflecting its performance quality, it is the only medical center awarded by the Healthcare Quality Improvement Campaign run by the Joint Commission of Taiwan in 2018 and received the Government Service Award from The National Development Council in 2019.

In the past, our institute, like most hospitals in Taiwan, issued regular news releases as the only tool to disseminate health information to the public. About 60 to 70 press conferences were held every year. To cope with recent marketing trends using social media, a branding team was established to strengthen online media marketing that included both Facebook and an internet website (Multimedia Appendix 1). Our aim here is to share our experiences with other institutes worldwide.

Establishing the Branding Team

The branding team was first set up in 2016. Its main mission was to establish branding strategies and integrate segmental branding messages. Under the leadership of the hospital's superintendent, 11 divisions were recruited to form this special task force (Figure 1). All divisions were required to draw up their own action plans and performance indicators. Top administrative officers oversaw and monitored task progress periodically.

Figure 1. Taichung Veterans General Hospital branding team organization.
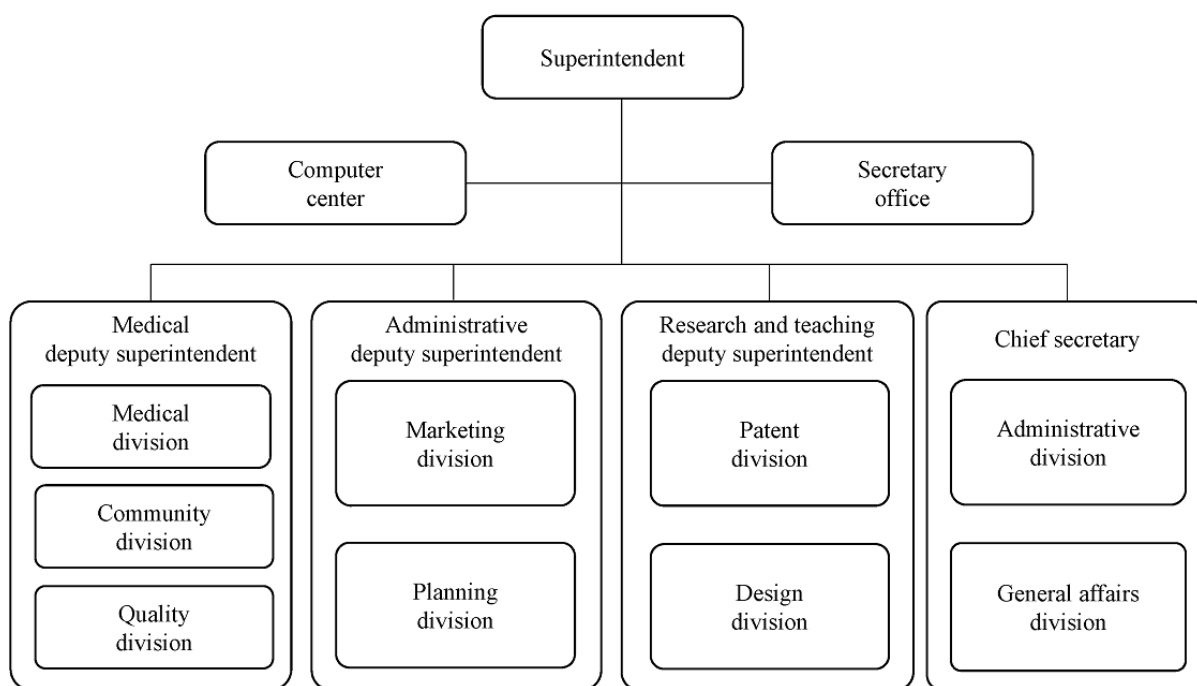
Data Collection and Analyses

We obtained reference data for 2016 for comparison with data for 2017 and 2018. We analyzed statistical differences between the means using IBM SPSS version 21 (IBM Corporation). We applied simple linear regression models (2-sided) to test for the trend in the numbers of likes on the Facebook fan page, posts, and hospital website as obtained during the 3 years from 2016 to 2018.

Results

Outcome Performance Indicators

Each division of the branding team provided action plans together with the corresponding indicators, all of which we periodically monitored. We monitored a total of 52 action plans and indicators in each quarter and discussed the results at team meetings (Multimedia Appendix 2). We selected 4 main outcome performance indicators to reflect the effectiveness and outcome of the branding efforts (Table 1). Compared with 2016, the growth rate in Facebook fans was 61.2% (n=17,474) in 2017 and 116.2% (n=23,436) in 2018 (Multimedia Appendix 3). Similarly, website visits increased by 4.8% (n=9,389,164) in 2017 and 33.1% (n=11,930,020) in 2018 (Multimedia Appendix 4).

The branding team also formed a LINE (LINE Corporation, Tokyo, Japan) group as a communication channel for patient referrals. There were 489 (in 2017) and 719 (in 2018) primary physicians participating in the LINE group with a growth rate of 100.4% in 2017 and 195.1% in 2018 (Multimedia Appendix 5). To obtain real-time information from the hospital, we invited all primary physicians to join our Facebook fan page. In addition, the branding team asked each medical group to set up peer support groups and to organize at least two activities each year. Reports of the activities held by 46 peer support groups in 2018 (Multimedia Appendix 6) were posted on Facebook. Patients and their families were also invited to join the hospital's fan page.

Table 1. Outcome performance indicators monitored by the branding team.

Indicators	Year			Growth rate (%)[a]		P value
	2016	2017	2018	2017	2018	
Number of likes on Facebook fan page	10,841	17,474	23,436	61.2	116.2	.02
Number of visits to the official website	8,963,097	9,389,164	11,930,020	4.8	33.1	.25
Number of clinic physicians joining LINE[b]	244	489	719	100.4	195.1	.01
Number of peer support groups	31	45	46	45	48	.30

[a]Growth rate is defined as the value of 2017 or 2018 minus that of 2016, and the result expressed as a percentage of 2016.

[b]A freeware app for instant communications on electronic devices.

General Public and Patient Engagement

According to the Facebook Insights report for 2018, 54.54% (227,989/417,988) of target audiences and 60.42% (14,160/23,436) of Facebook fans were women. Most were between 25 and 44 years of age (199,906/417,988, 47.83%). Based on Google Analytics, more women (778,992/1,311,605, 59.39%) than men visited the hospital website. Website viewers were mostly between 25 and 44 years of age (778,668/1,289,988, 60.36%) (Table 2).

Table 2. Sex and age group distributions on the Facebook fan page and hospital website in 2018, according to Facebook Insights reports Google Analytics.

Characteristics	Facebook fan page, n (%)		Website visits, n (%)
	Target audiences	Fans	
Sex			
Female	227,989 (54.54)	14,160 (60.50)	778,992 (59.39)
Male	189,999 (45.46)	9276 (39.50)	532,613 (40.61)
Total	417,988 (100.00)	23,436 (100.0)	1,311,605 (100.00)
Age group (years)			
13-17	1790 (0.43)	166 (0.71)	___a
18-24	41,699 (9.98)	2062 (8.80)	170,292 (13.20)
25-34	100,582 (24.06)	5479 (23.38)	424,779 (32.93)
35-44	99,324 (23.76)	6441 (27.48)	353,889 (27.43)
45-54	70,072 (16.76)	4675 (19.95)	202,188 (15.67)
55-64	71,287 (17.05)	3218 (13.73)	131,884 (10.22)
≥65	33,234 (7.95)	1395 (5.95)	6956 (0.54)
Total	417,988 (100.00)	23,436 (100.0)	1,289,988 (100.0)

[a]Not available.

Table 3 shows that the number of posts increased from 461 in 2016 to 595 in 2017 (growth rate of 29.1%) and 566 in 2018 (growth rate 22.8%). The number of videos posted was 16 in 2016, increasing to 49 in 2018 and to 90 in 2018 (P=.04 for the trend). The number of likes almost doubled in 2017 (n=99,262) over 2016 and stayed at approximately the same level in 2018 (n=91,337; P=.42, for the trend). Increments were similarly marked for comment messages, which increased nearly threefold in 2017 (n=2931) and fourfold in 2018 (n=4559; P=.02 for the trend). Post sharing also doubled in both years (n=1755 in 2016, n=4783 in 2017, n=4629 in 2018; P=.36 for the trend).

Table 3. Facebook Insights reports for posts made from 2016 to 2018.

Items	Year, n			Growth rate, %[a]		P value
	2016	2017	2018	2017	2018	
Posts	461	595	566	29.1	22.8	.47
Videos	16	49	90	206	463	.04
Post likes	46,014	99,262	91,337	115.7	98.5	.42
Comment messages	1127	2931	4559	160.1	305.2	.02
Post sharing	1755	4783	4629	91.9	85.6	.36

[a]Growth rate is defined as the value of 2017 or 2018 minus that of 2016, and the result expressed as a percentage of 2016.

Post Categories and the Analysis of Likes on Facebook

In total, 1622 posts were provided by each division of the branding team from 2016 to 2018. We divided posts according to their content characteristics into 14 unique post themes and hashtags in the text (Multimedia Appendix 7).

To identify what types of posts audiences liked, we classified the 14 themes into 5 groups for analysis: innovative service, media reports, activity information, patient gratitude letter, and health education information. Table 4 shows that the type of post with the highest number of likes was media reports in 2016, being replaced by innovative service in both 2017 and 2018. The mean number of likes for innovative service posts rose significantly from 2016 through 2018 (P=.045 for the trend). The numbers of likes for health education information posts remained the lowest for all 3 years.

**Table 4.** Likes per post category on the hospital Facebook fan page from 2016 to 2018.

Type of post	2016				2017				2018				P value
	Likes, n	Posts, n	Mean[a]	Rank	Likes, n	Posts, n	Mean[a]	Rank	Likes, n	Posts, n	Mean[a]	Rank	
Innovative service	4299	39	110	2	9377	40	234	1	19,539	59	331	1	.045
Media reports	7074	57	124	1	23,796	163	146	4	33,374	151	221	2	.2
Activity information	20,972	198	106	3	42,599	212	201	2	16,882	136	124	4	.89
Patient gratitude letter	3862	42	92	4	7214	46	157	3	6307	46	137	3	.53
Health education information	9807	125	78	5	16,276	134	121	5	15,235	174	88	5	.86

[a]Average number of likes per post (number of likes divided by number of posts).

Figure 2 shows the cumulative number of likes since 2016. The growth rate from the first quarter to the second quarter was 10.1% in 2016, 17.0% in 2017, and 5.5% in 2018. The growth rate from the third quarter to the fourth quarter was 21.0% in 2016, 21.4% in 2017, and 7.6% in 2018. Some of the increases in the numbers of likes were apparently related to promotional activities during festivals, as well as interactions with fans. For example, fan interaction activities were intensified during the Chinese New Year and Christmas season.

Figure 2. Cumulative numbers and quarterly growth rates of likes on the Facebook fan page from 2016 to 2018. Q: quarter.
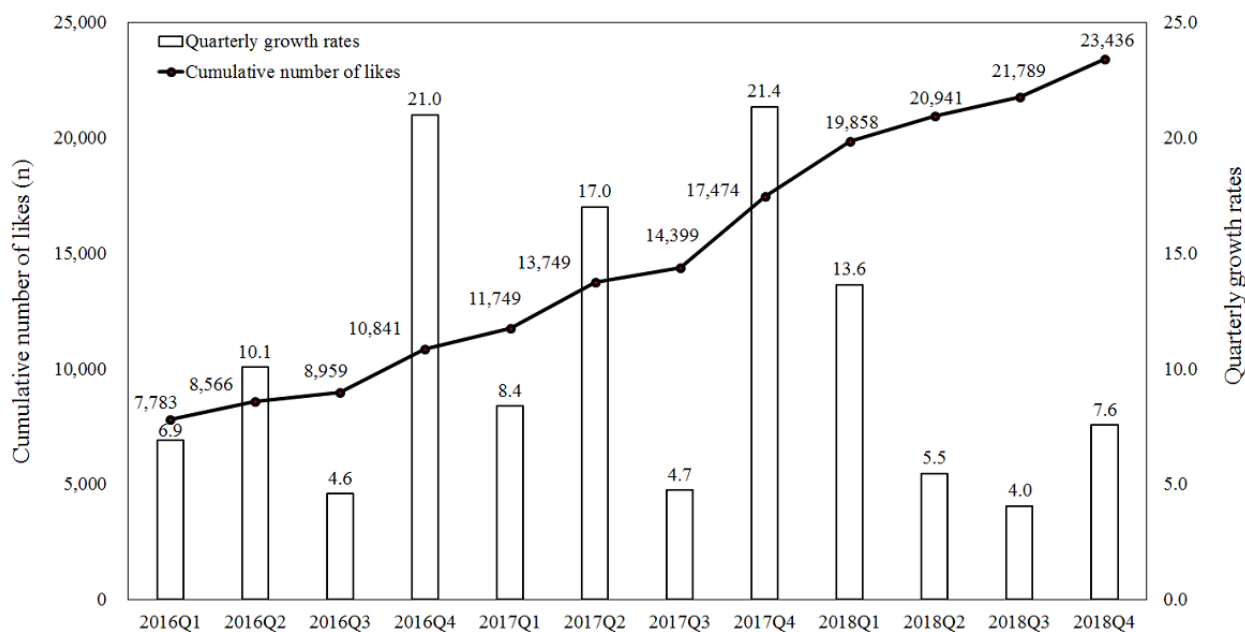
Discussion

Principal Findings

Our main finding is that a public medical center in Taiwan such as ours can, through the efforts of a branding team led by top administrators, gain Facebook fans. We attained the highest number of Facebook fans among medical centers in Taiwan, with a 116.2% increase over 2 years. We believe that the factors contributing to this success were (1) the interconnected Facebook and hospital webpages, (2) active participation of the branding team with timely provision of attractive posts and videos, (3) strengthened interactions with online visitors, (4) fast responses to users' queries and messages, and (5) increased online streaming.

Functions of the Branding Team

With strong support from the top administrators, the branding team was able to play a key role in implementing strategic marketing functions. We identified several distinguished medical services as strategic features through interdepartmental communications. Results suggested that, in this digital era, targeting social media is an effective approach for promoting health consumer education, health care group communication, and brand awareness. Our Facebook audience not only came from the general public, but also from strategically invited peer support groups and primary physicians. Given the importance of brand management and extraction of associated values, health service organizations should diligently attend to branding initiatives. Enhanced values can be derived by addressing nontraditional brand elements that provide unique opportunities to facilitate institutional viability and vitality [23].

Hospitals use the Facebook platform as an inexpensive way to educate people on topics of health and well-being, and to communicate different types of information and news to the general public. Based on more than 1700 Facebook posts from 17 hospitals in the United States, Kordzadeh and Young [24] identified 13 unique health social media post themes and classified them into 3 thematic groups that included announcing, sharing, and recognizing activities. The most frequently used theme was the sharing of health information, which appeared in 35.8% (424/1184) of the posts. Such posts provided health tips and advice to community members [24]. In our study, we classified 14 themes into 5 groups: innovative service, activity information, media reports, health education information, and patient gratitude letter. In 2016, the most popular post theme was media reports. The main reason is that this information was generally linked to television news reports rather than plain-text posts. Since 2017, for proposing innovative service, health tips and advice were strategically posted in various forms, such as online streaming or videos. Due to the good strategy, innovative service became the most popular post theme. The number of likes for health education information remained the lowest for all 3 years. The main reason may be related to the plain-text form by which the theme was posted. At present, we are systematically replacing plain texts with more interesting online streams or videos.

In addition, to expand our brand and intensify fan interaction, we sent branded gifts such as management books we had published, as well as mugs, ties, purses, and pens carrying our hospital logo, during online streaming. Since many people nowadays prefer watching videos to reading texts, the branding team decided to improve video streaming in real time and to strengthen links to other online media. Probably as the consequence of the above measures, the number of likes rose to make ours the most popular Facebook fan page among all medical centers in Taiwan.

Special Features of Our Facebook Page

Using social media not only promotes marketing, but also upgrades care for patients and their families, enhances health consumer education, advances medical research, and expands

brand awareness [25]. Facebook pages also serve as a tool for patient empowerment and allow for intercommunication between physicians and patients. Given the high volume of posts, it is imperative that the information provided be accurate and in accordance with the medical advice of physicians [26]. Other studies used Facebook for disease surveillance [27] and health interventions [28]. Facebook likes can reflect users' preference, and thus can help predict health-related behaviors [29,30]. A previous study in 2014 reported that 99.4% (3351/3371) of hospitals in the United States had established Facebook pages. The use of social media varied according to the different characteristics of hospitals such as their size, urban location, and whether they were private nonprofit or teaching hospitals. All these factors affected their levels of activity on Facebook [18]. Another study of 12 Western European countries in 2012 found that 67.0% (585/873) of their hospitals had Facebook fan pages [31].

The penetration rate of Facebook in Taiwan is 82%, which is the highest in the world [32]. Despite this, only 51.1% of the hospitals have official Facebook fan pages. Furthermore, in comparison with other commercial Facebook fan pages, the numbers of likes are relatively low on hospital Facebook fan pages in Taiwan [22]. Social media is cheaper than conventional marketing, but its effects are enormous. For public hospitals with a restricted marketing budget, promotion through social media is a good investment approach. To strengthen hospital branding and to synchronize with the latest news, our results showed that Facebook and the hospital website are better interconnected. In addition, for better results, information should be posted quickly, within hours of events, and query comment messages should be answered as soon as possible.

Limitations

Our study provided insights into the ways our hospital had established a special team to promote branding by recruiting fans to our hospital Facebook page. However, our study had the following limitations. First, we reported and analyzed the methods of only our own hospital, and did not collect data from other hospitals for comparison. Second, social media and channels are changing rapidly. The use rate of different social media in 2019 could have been different from 2017. We have already expedited the construction of a LINE group in our hospital because of its increasing popularity over Facebook. Third, our Facebook fan page was established in 2012, while we obtained the reference data only in 2016 when we first started our Facebook drive. Fourth, since in Taiwan Facebook and mobile phones had been gaining in popularity from 2016 onward, we implemented these strategies for only 3 years. Details on longitudinal trends remain to be studied.

Conclusions

Our branding team, led by the hospital's top officers, successfully implemented several strategies that achieved the most popular Facebook fan page among Taiwan hospitals. Strategies we used were powerful in providing information on time and in promoting better medical services. Our unique experience in Facebook management may lay the groundwork for hospitals' use of social media platforms to improve patient interactions and health care outcomes.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Facebook fan page and hospital website of Taichung Veterans General Hospital.
[PDF File (Adobe PDF File), 361 KB - ijmr_v9i1e14546_app1.pdf]

Multimedia Appendix 2
The 52 performance indicators monitored, by branding team.
[PDF File (Adobe PDF File), 115 KB - ijmr_v9i1e14546_app2.pdf]

Multimedia Appendix 3
Counts of visitors to the Facebook fan page per month from 2016 to 2018.
[XLSX File (Microsoft Excel File), 33 KB - ijmr_v9i1e14546_app3.xlsx]

Multimedia Appendix 4
Counts of visitors to the hospital website per month from 2016 to 2018.
[XLSX File (Microsoft Excel File), 11 KB - ijmr_v9i1e14546_app4.xlsx]

Multimedia Appendix 5
LINE groups for primary physicians.
[PDF File (Adobe PDF File), 99 KB - ijmr_v9i1e14546_app5.pdf]

XSL•FO
RenderX

Multimedia Appendix 6
Peer support group webpage.
[PDF File (Adobe PDF File), 161 KB - ijmr_v9i1e14546_app6.pdf]

Multimedia Appendix 7
Themes of Facebook posts from 2016 to 2018.
[XLSX File (Microsoft Excel File), 10 KB - ijmr_v9i1e14546_app7.xlsx]

References

1. Kim KH, Kim KS, Kim DY, Kim JH, Kang SH. Brand equity in hospital marketing. J Bus Res 2008 Jan;61(1):75-82. [doi: 10.1016/j.jbusres.2006.05.010]
2. Porter SS, Claycomb C. The influence of brand recognition on retail store image. J Prod Brand Manag 1997 Dec;6(6):373-387. [doi: 10.1108/10610429710190414]
3. Bloemer J, de Ruyter K. On the relationship between store image, store satisfaction and store loyalty. Eur J Market 1998 Jun;32(5/6):499-513. [doi: 10.1108/03090569810216118]
4. International Telecommunication Union. ICT facts and figures 2017. Geneva, Switzerland: ICT Data and Statistics Division, ITU; 2017 Jul. URL: https://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2017.pdf [accessed 2019-11-20]
5. Ye Q, Fang B, He W, Hsieh J. Can social capital be transferred cross the boundary of the real and virtual worlds? An empirical investigation of Twitter. J Electron Commerce Res 2012;13(2):145-156.
6. Tengilimoglu D, Sarp N, Yar CE, Bektaş M, Hidir MN, Korkmaz E. The consumers' social media use in choosing physicians and hospitals: the case study of the province of Izmir. Int J Health Plann Manage 2017 Jan;32(1):19-35. [doi: 10.1002/hpm.2296] [Medline: 26198772]
7. Esposito A. Hospital branding in Italy: a pilot study based on the case method. Health Mark Q 2017;34(1):35-47. [doi: 10.1080/07359683.2016.1275211] [Medline: 28350275]
8. Fox S. The social life of health information, 2011. Washington, DC: Pew Research Center; 2011 May 12. URL: http://www.pewinternet.org/2011/05/12/the-social-life-of-health-information-2011/ [accessed 2019-04-24]
9. Elkin N. How America searches: health and wellness. iCrossing 2008:1-17.
10. Warden C. 30 facts & statistics on social media and healthcare. Mt Pleasant, SC: ReferralMD; 2017. URL: https://getreferralmd.com/2017/01/30-facts-statistics-on-social-media-and-healthcare/ [accessed 2019-04-24]
11. Greaves F, Ramirez-Cano D, Millett C, Darzi A, Donaldson L. Harnessing the cloud of patient experience: using social media to detect poor quality healthcare. BMJ Qual Saf 2013 Mar;22(3):251-255. [doi: 10.1136/bmjqs-2012-001527] [Medline: 23349387]
12. Greene JA, Choudhry NK, Kilabuk E, Shrank WH. Online social networking by patients with diabetes: a qualitative evaluation of communication with Facebook. J Gen Intern Med 2011 Mar;26(3):287-292 [FREE Full text] [doi: 10.1007/s11606-010-1526-3] [Medline: 20945113]
13. Adams S. Post-panoptic surveillance through healthcare rating sites. Inf Commun Soc 2013 Mar;16(2):215-235. [doi: 10.1080/1369118x.2012.701657]
14. Antheunis ML, Tates K, Nieboer TE. Patients' and health professionals' use of social media in health care: motives, barriers and expectations. Patient Educ Couns 2013 Sep;92(3):426-431. [doi: 10.1016/j.pec.2013.06.020] [Medline: 23899831]
15. Smailhodzic E, Hooijsma W, Boonstra A, Langley DJ. Social media use in healthcare: a systematic review of effects on patients and on their relationship with healthcare professionals. BMC Health Serv Res 2016 Aug 26;16:442 [FREE Full text] [doi: 10.1186/s12913-016-1691-0] [Medline: 27562728]
16. Barber SK, Lam Y, Hodge TM, Pavitt S. Is social media the way to empower patients to share their experiences of dental care? J Am Dent Assoc 2018 Jun;149(6):451-459.e9. [doi: 10.1016/j.adaj.2018.01.007] [Medline: 29656806]
17. Al-Qahtani M, Alsaffar A, Alshammasi A, Alsanni G, Alyousef Z, Alhussaini M. Social media in healthcare: advantages and challenges perceived by patients at a teaching hospital in eastern province, Saudi Arabia. Saudi J Health Sci 2018;7(2):116. [doi: 10.4103/sjhs.sjhs_36_18]
18. Griffis HM, Kilaru AS, Werner RM, Asch DA, Hershey JC, Hill S, et al. Use of social media across US hospitals: descriptive analysis of adoption and utilization. J Med Internet Res 2014 Nov 27;16(11):e264 [FREE Full text] [doi: 10.2196/jmir.3758] [Medline: 25431831]
19. Internet World Stats: Usage and Population Statistics, Miniwatts Marketing Group. 2019 Nov 15. Internet usage in Asia: Asia internet use, population data and Facebook statistics – June 30, 2019. URL: https://www.internetworldstats.com/stats3.htm [accessed 2019-11-27]
20. Taiwan Institute for Information Industry FIND, Ministry of Economic Affairs. Service system driven and new enterprise research and development plan (2/4). 2016. URL: https://www.iii.org.tw/Press/NewsDtl.aspx?nsp_sqno=1934&fm_sqno=14 [accessed 2019-04-24] [WebCite Cache ID 77s78WPB7]

21.   Ministry of Health and Welfare, ROC. 2018 Taiwan Health and Welfare Report. 2019. URL: https://www.mohw.gov.tw/cp-3196-46105-1.html [accessed 2019-08-09]

22.   Yang PC, Lee WC, Liu HY, Shih MJ, Chen TJ, Chou LF, et al. Use of Facebook by hospitals in Taiwan: a nationwide survey. Int J Environ Res Public Health 2018 Jun 06;15(6) [FREE Full text] [doi: 10.3390/ijerph15061188] [Medline: 29882819]

23.   Elrod JK, Fortenberry JL. Driving brand equity in health services organizations: the need for an expanded view of branding. BMC Health Serv Res 2018 Dec 14;18(Suppl 3):924 [FREE Full text] [doi: 10.1186/s12913-018-3679-4] [Medline: 30545362]

24.   Kordzadeh N, Young DK. Exploring hospitals' use of Facebook: thematic analysis. J Med Internet Res 2018 May 16;20(5):e190 [FREE Full text] [doi: 10.2196/jmir.9549] [Medline: 29769173]

25.   Bhattacharya S, Srinivasan P, Polgreen P. Social media engagement analysis of U.S. Federal health agencies on Facebook. BMC Med Inform Decis Mak 2017 Apr 21;17(1):49 [FREE Full text] [doi: 10.1186/s12911-017-0447-z] [Medline: 28431582]

26.   Maganty N, Ilyas M, Ginsberg Z, Sharma A. Social media as a platform for information and support for melanoma patients: analysis of melanoma Facebook groups and pages. JMIR Dermatol 2018 May 16;1(1):e2. [doi: 10.2196/derma.8482]

27.   Reynolds BJ. Building trust through social media. CDC's experience during the H1N1 influenza response. Mark Health Serv 2010;30(2):18-21. [Medline: 20550001]

28.   Struik LL, Baskerville NB. The role of Facebook in Crush the Crave, a mobile- and social media-based smoking cessation intervention: qualitative framework analysis of posts. J Med Internet Res 2014 Jul 11;16(7):e170 [FREE Full text] [doi: 10.2196/jmir.3189] [Medline: 25016998]

29.   Gittelman S, Lange V, Gotway Crawford CA, Okoro CA, Lieb E, Dhingra SS, et al. A new source of data for public health surveillance: Facebook likes. J Med Internet Res 2015 Apr 20;17(4):e98 [FREE Full text] [doi: 10.2196/jmir.3970] [Medline: 25895907]

30.   Kosinski M, Bachrach Y, Kohli P, Stillwell D, Graepel T. Manifestations of user personality in website choice and behaviour on online social networks. Mach Learn 2013 Oct 19;95(3):357-380. [doi: 10.1007/s10994-013-5415-y]

31.   Van de Belt TH, Berben SAA, Samsom M, Engelen LJLPG, Schoonhoven L. Use of social media by Western European hospitals: longitudinal study. J Med Internet Res 2012 May 01;14(3):e61 [FREE Full text] [doi: 10.2196/jmir.1992] [Medline: 22549016]

32.   Fulco M. Taiwan Business Topics. Taipei, Taiwan: American Chamber of Commerce; 2017 Apr 12. Facebook still dominates Taiwan's social media. URL: https://topics.amcham.com.tw/2017/04/Facebook-still-dominates-taiwans-social-media/ [accessed 2019-04-24]

Original Paper

Use of Artificial Intelligence for Medical Literature Search: Randomized Controlled Trial Using the Hackathon Format

Dominik Schoeb[1], MD; Rodrigo Suarez-Ibarrola[1], MD; Simon Hein[1], MD; Franz Friedrich Dressler[1], MD; Fabian Adams[1], MD; Daniel Schlager[1], MD; Arkadiusz Miernik[1], MD, PhD

Medical Center – Department of Urology, Faculty of Medicine, University of Freiburg, Freiburg, Germany

Corresponding Author:
Dominik Schoeb, MD
Medical Center – Department of Urology
Faculty of Medicine
University of Freiburg
Freiburg,
Germany
Phone: 49 076127025823
Email: dominik.stefan.schoeb@uniklinik-freiburg.de

Abstract

Background: Mapping out the research landscape around a project is often time consuming and difficult.

Objective: This study evaluates a commercial artificial intelligence (AI) search engine (IRIS.AI) for its applicability in an automated literature search on a specific medical topic.

Methods: To evaluate the AI search engine in a standardized manner, the concept of a science hackathon was applied. Three groups of researchers were tasked with performing a literature search on a clearly defined scientific project. All participants had a high level of expertise for this specific field of research. Two groups were given access to the AI search engine IRIS.AI. All groups were given the same amount of time for their search and were instructed to document their results. Search results were summarized and ranked according to a predetermined scoring system.

Results: The final scoring awarded 49 and 39 points out of 60 to AI groups 1 and 2, respectively, and the control group received 46 points. A total of 20 scientific studies with high relevance were identified, and 5 highly relevant studies ("spot on") were reported by each group.

Conclusions: AI technology is a promising approach to facilitate literature searches and the management of medical libraries. In this study, however, the application of AI technology lead to a more focused literature search without a significant improvement in the number of results.

Introduction

Mapping out the research landscape around a project is time consuming and often frustrating, as the sheer size of published data is often impossible to read and understand, much less be put in context. To overcome these challenges, the current scientific standard is a systematic literature review using standardized methodology in accordance with the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analysis) guidelines [1] and the Cochrane regulations [2]. This process is thorough and reliable; however, it also requires a significant amount of time and resources. To overcome these

limitations and make literature searches more effective, a promising approach is the application of artificial intelligence (AI) technology, which can screen vast amounts of data faster and more efficiently than a human researcher. When founded in 2016, the IRIS.AI search engine was the first application addressing the problem of screening vast amounts of scientific literature using AI. Other promising formats have emerged since then, like Google's new AI service Talk to Books [3]; however, IRIS.AI is still the only application that exclusively caters to the scientific field.

The current topics of interest for our research group are virtual reality (VR) and augmented reality (AR) in surgery. VR and

XSL•FO
RenderX

AR are new technologies with a high potential for application in the medical field. Besides use in diagnostic imaging applications and the facilitation of minimally invasive procedures, the implementation of these technologies can substantially improve medical education, particularly when it comes to teaching practical tasks. Many studies and projects have been conducted on this topic; however, more research and development work is necessary to routinely implement those technologies into medical education. To investigate the most promising approaches and questions pertaining to the implementation of VR and AR, intensive literature research is necessary. The intention of this study was to evaluate the commercially available AI search engine IRIS.AI in the application of medical literature research on these topics.

Methods

Study Design

To evaluate the AI search engine, we gathered surgery experts from different specialties (ie, general surgery; ear, nose and throat surgery; neurosurgery; orthopedics and trauma surgery; and urology), information technology (IT) experts, and medical engineers to perform a systematic literature search on a specific scientific question in a predefined time frame. Participants were randomized to three different teams. All literature research was done in a 1-day event, and the three teams competed for the best overall result. All participants were provided similar computers for the study, and all search activities were performed at the same location. Communication between the different teams was not permitted. This format, called a "science hackathon" by IRIS.AI, was originally applied in the IT industry with the goal of finding a solution to a programming challenge. This format has been adapted for multiple fields, including scientific purposes and general topics such as marketing [4].
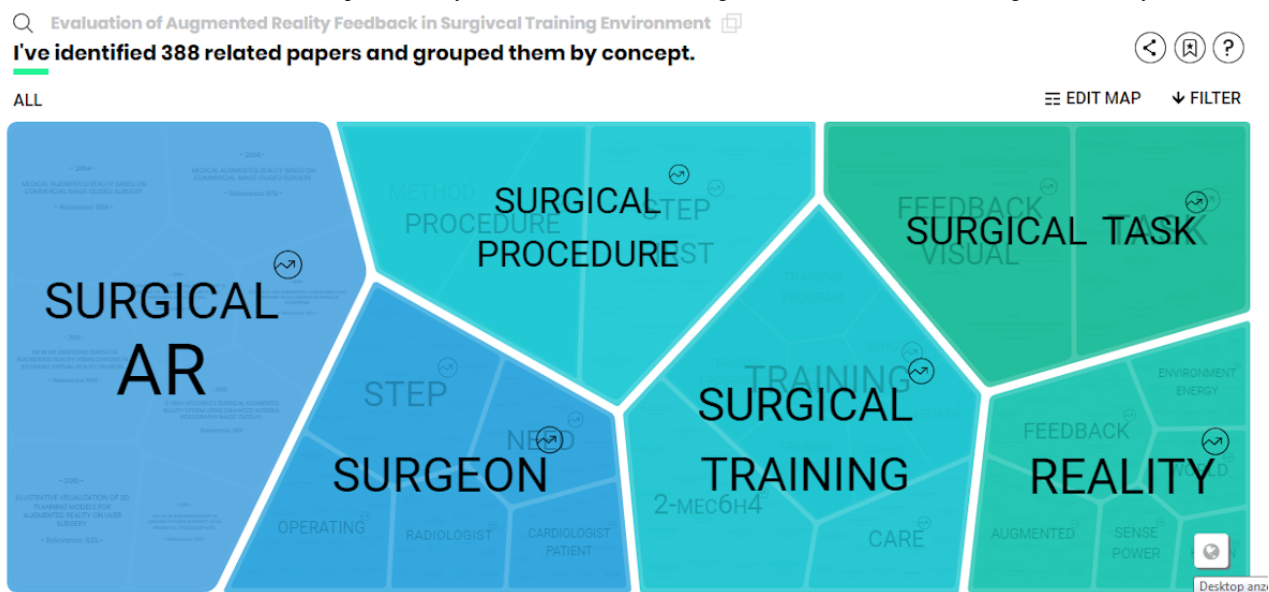
Group Participants

A total of 17 scientists with expertise in medical education, surgery, IT, or engineering took part in the literature research.

One participant cancelled on the day of the event and prevented an even distribution of group members. All scientists had an extensive background in scientific research and, therefore, a high proficiency in literature research using scientific databases. We recruited participants through public announcements and information given to collaborating research groups and institutes (Multimedia Appendix 1). All participants of this study consented to participate. The need for ethical approval was waived since no personal data was used in this study. The participants were divided according to their field of expertise to ensure an even distribution of medical and engineering experts. They were then randomly assigned to one of three groups. A detailed description of the group members' scientific backgrounds can be found in Multimedia Appendix 2. This resulted in two groups with 6 members and one group with 5 members. After the groups were determined, one was randomly designated as the control group without access to AI technology.

Literature Search

The IRIS.AI search engine can understand key concepts of a given scientific article and can search for relevant data that is similar or connected to the article by using an AI algorithm. The search results are displayed in a visual map divided in subcategories related to the main topic (Figure 1). A detailed description on how the AI software operates can be found at the website of the providing company (https://the.iris.ai/). Research time was limited to 5 hours and divided into two research sessions. All groups were instructed to document their research results in a standardized format using a report form (Multimedia Appendix 3), which was issued prior to the commencement of the research time. In addition, searches were monitored using a key-logging program on the search computers, and remote surveillance was used on the PC screen. The control group used Google Scholar, Web of Science, and PubMed in their literature research.

Figure 1. Screenshot of a search result "map" created by the IRIS.AI artificial intelligence research assistant. AR: augmented reality.

Definition of Research Task

All participants were issued a written problem statement at the beginning of the event with a clearly defined research task (Multimedia Appendix 4). The general question was: "What research and development work is necessary to build a 'ready to use' adaptive augmented reality (AR) system for surgeons to teach surgical residents how to perform a surgical procedure?" The three groups were asked to focus their literature search on the following three main problems defined by the posed question: (1) the recognition of reality such as anatomy or surgical instruments, which is necessary to fulfill the task of supporting a surgical resident; (2) the hardware necessary to provide the proposed system such as voice guidance or AR glasses; and (3) the pedagogical concept behind an AR system that "teaches" a surgeon how to perform a surgical procedure "step-by-step". The researchers were asked to find literature on the potential solutions to these problems, the application of similar approaches, and the capability of available technologies. Additionally, they were asked to find conceptual studies on how to combine available technologies.

Documentation of Results

Before the research time ended, all groups were asked to document their research results using a standardized report form (Multimedia Appendix 3). In addition, the groups were supposed to differentiate between the three main problems and sort their results on the report form accordingly.

Final Evaluation

A committee of three experts with backgrounds related to the posed scientific question evaluated and ranked the search results according to scientific quality and quantity criteria using a standardized scoring sheet (Multimedia Appendix 5). The experts included a urological surgeon with experience in medical and surgical education, a software engineer with specific background in the field of AR technology, and an engineer in charge of research and development at a large medical technology company. The results of the quantitative and qualitative evaluation were added to a final score, and teams were ranked accordingly (Multimedia Appendix 6).

Results

All participating teams were able to detect literature relevant to the topic. The first team, using the IRIS.AI search tool, listed 13 relevant papers on their score sheet, which were all judged as relevant for the proposed topic. The quality of the found studies was ranked very high, with a total score of 27. Multiple studies applying AR in surgical fields were detected, which shows that AR can provide an accurate visual representation of anatomical structures and can be a helpful tool during surgical procedures [5-10]. The second team supported by AI listed 15 relevant papers, but only 8 were found as related to the field by the judges. The quality of the ranked articles was awarded with 19 points. The literature review performed by this team followed a problem-oriented strategy that focused directly on researching the challenges proposed by the scientific question and the potential solutions available in the literature. Regarding the perception of reality and the tracking of deformable objects,

available technology was detected, which technically addresses the challenges of an adaptive AR system for intraoperative guidance, but the necessity of high computational power and the limited reliability are still challenges that need to be addressed before a clinical use is possible [11-13]. The control group listed the highest number of results with 46 identified scientific studies, but only 10 studies were found to be related to the field. The quality of the studies was graded with a total of 25 points. This group provided a broad selection of literature with an overview of the capabilities of existing technologies for intraoperative support of the operating surgeon, as well as the limitations and problems that need to be addressed when implementing the proposed system into clinical practice. In total, 20 relevant scientific studies were identified throughout the event covering all aspects of the question. Of those 20 studies, the AI groups contributed 7 findings each, while the control group contributed 6 studies. Final evaluation of the study content showed a result of 5 highly relevant studies ("spot on") in each of the groups, which suggests a more focused search strategy through the application of AI. The final scores were 49 and 39 points out of 60 for AI groups 1 and 2, respectively, and 46 points for the control group. Data from all relevant studies was extracted, structured, and divided by search teams. All data are presented and summarized in a descriptive manner in Multimedia Appendix 6.

Discussion

This study was, to our knowledge, the first experiment in the medical field using a science hackathon format and an AI tool to perform a literature search. Both AI groups showed similar quantitative results, but the quality differed between these groups which lead to a difference in their total scores, with a score of 27 points for AI group 1 and a score of 19 points for AI group 2. The control group showed a similar degree of quality as AI group 1 with 25 points awarded, but only 10 of the 46 studies found were categorized as related to the field. AI has been an expanding research area in the last decade, and AI programs have been implemented in various fields, including the medical practice [14]. However, little research has been done on the topic of applying this technology to facilitate literature research. Conventional literature search engines like Google Scholar and PubMed are able to perform key word searches as well as full text searches to some extent, but their ability to recognize similarities is limited to search phrases, available publication titles, or abstract contents [15]. AI, as used by IRIS.AI, has the potential to "understand" the posed research question as well as the screened scientific articles, thus providing the opportunity to filter out the relevant data more effectively. According to the 2015 report [16] of the International Association of Scientific, Technical, and Medical Publishers, approximately 2.5 million scientific papers are published per year, and there are approximately 28,000 active scholarly peer-reviewed research journals. Therefore, a more effective way to perform literature research is highly relevant. In this study, experienced scientific professionals were asked to perform a comprehensive literature review on the question of AR in surgery by applying the artificial search engine of IRIS.AI. When evaluating search results, we focused on the quantity of the relevant articles as

well the quality of the results, following the recommendations found in existing literature for the comparison of literature search engines [15,17,18]. Although all participating teams were able to retrieve relevant data on the topic, the results indicate that the application of AI has the potential to focus a scientific search more directly on relevant data and decrease the time necessary to screen the retrieved articles for the actual information needed. Both AI groups showed a significant difference in quality of the detected results, but the AI technology did not provide an increase in search quality. Our study, however, does have some limitations to be considered. First, although all teams were comprised of experts with similar backgrounds, individual differences in know-how and skill can never be entirely excluded, and might pose a bias since only a small sample size of three teams was provided. Second, the qualitative evaluation of search results and the categorization of "relevant" results performed by three judges of high expertise in the relevant fields has the potential to be subjectively biased, and individual preference of the judges in the evaluation can never be entirely excluded. However, to our knowledge, no completely unbiased tool is available. Third, all participants had no previous contact with the IRIS.AI search engine, whereas extensive experience in literature review was present in all participants. Therefore, pre-existing individual preference for certain search engines and a high degree of familiarity in their use were potential biases. Last, although it provides access to most major research databases, the full text access to some databases is not available for the IRIS.AI tool, which poses a potential barrier for the retrieval of relevant information.

This study found a lot of relevant research data on the proposed topic of the development of an adaptive AR system for surgeons to teach surgical residents how to perform a surgical procedure, which makes the format of a science hackathon a useful research tool to gain a fast and effective overview of the literature available for a certain topic.

AI technology is a promising approach to facilitate literature research and comparable to current conventional search engines. In this study, the application of AI technology lead to a more focused literature search without, however, a significant improvement of search results. The format of a scientific hackathon is an efficient tool for literature research, and provides scientists a faster method to gather relevant literature compared to conventional review methods.

Authors' Contributions

D Schoeb contributed towards protocol and project development, data analysis, and manuscript and figure writing and editing. RS-I provided manuscript editing and data analysis. SH contributed towards protocol development and manuscript editing. D Schlager provided data analysis and manuscript editing. AM contributed towards protocol and project development, manuscript and figure writing and editing, and supervision.

Conflicts of Interest

None declared.

This randomized study was not prospectively registered, justified by the authors as recruiting for the study was conducted within the authors' institution and partner institutions only and therefore was not registered in public registries. The editor granted an exception of ICMJE rules for prospective registration of randomized trials because the risk of bias appears low. However, readers are advised to carefully assess the validity of any potential explicit or implicit claims related to their primary outcomes or effectiveness, as the lack of registration means that authors could change their outcome measures retrospectively.

Multimedia Appendix 1
Event invitation used in the recruitment process.
[PDF File (Adobe PDF File), 39 KB - ijmr_v9i1e16606_app1.pdf]

Multimedia Appendix 2
Team composition and evaluation of literature search results of all teams.
[DOCX File , 19 KB - ijmr_v9i1e16606_app2.docx]

Multimedia Appendix 3
Report form for literature search results.
[DOCX File , 15 KB - ijmr_v9i1e16606_app3.docx]

Multimedia Appendix 4
Problem statement and definition of research task issued to all participants.

XSL•FO
RenderX

[DOCX File , 17 KB - ijmr_v9i1e16606_app4.docx]

Multimedia Appendix 5
Scoring sheet used in the evaluation of team results.
[PDF File (Adobe PDF File), 83 KB - ijmr_v9i1e16606_app5.pdf]

Multimedia Appendix 6
Summary of search results divided by search teams.
[DOCX File , 52 KB - ijmr_v9i1e16606_app6.docx]

Multimedia Appendix 7
CONSORT-eHEALTH checklist (v 1.6.1).
[PDF File (Adobe PDF File), 1427 KB - ijmr_v9i1e16606_app7.pdf]

References

1. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. Open Med 2009;3(3):e123-e130 [FREE Full text] [Medline: 21603045]
2. Higgins J, Green S. Cochrane Handbook for Systematic Reviews of Interventions: Cochrane Book Series. 2011 Mar. URL: https://training.cochrane.org/handbook/archive/v5.1/
3. Google Inc. Talk to Books URL: https://books.google.com/talktobooks/ [accessed 2019-08-08]
4. Calco M, Veeck A. The markathon: adapting the hackathon model for an introductory marketing class project. Marketing Education Review 2015 Mar 26;25(1):33-38. [doi: 10.1080/10528008.2015.999600]
5. Davis L, Hamza-Lup F, Daly J, Ha Y, Frolich S, Meyer C. Application of augmented reality to visualizing anatomical airways. International Society for Optics and Photonics 2002. [doi: 10.1117/12.478890]
6. Khor WS, Baker B, Amin K, Chan A, Patel K, Wong J. Augmented and virtual reality in surgery-the digital surgical environment: applications, limitations and legal pitfalls. Ann Transl Med 2016 Dec;4(23):454 [FREE Full text] [doi: 10.21037/atm.2016.12.23] [Medline: 28090510]
7. Mekni M, Lemieux A. Augmented reality: applications, challenges and future trends. Applied Computational Science 2014.
8. Ota D, Loftin B, Saito T, Lea R, Keller J. Virtual reality in surgical education. Comput Biol Med 1995 Mar;25(2):127-137. [doi: 10.1016/0010-4825(94)00009-f] [Medline: 7554831]
9. Quellec G, Charrière K, Lamard M, Droueche Z, Roux C, Cochener B, et al. Real-time recognition of surgical tasks in eye surgery videos. Med Image Anal 2014 Apr;18(3):579-590. [doi: 10.1016/j.media.2014.02.007] [Medline: 24637155]
10. Sielhorst T, Feuerstein M, Navab N. Advanced Medical Displays: A Literature Review of Augmented Reality. J. Display Technol 2008 Dec;4(4):451-467. [doi: 10.1109/jdt.2008.2001575]
11. Maier-Hein L, Mountney P, Bartoli A, Elhawary H, Elson D, Groch A, et al. Optical techniques for 3D surface reconstruction in computer-assisted laparoscopic surgery. Med Image Anal 2013 Dec;17(8):974-996. [doi: 10.1016/j.media.2013.04.003] [Medline: 23837969]
12. Van Krevelen D, Poelman R. A Survey of Augmented Reality Technologies, Applications and Limitations. IJVR 2010 Jan 01;9(2):1-20. [doi: 10.20870/ijvr.2010.9.2.2767]
13. Zhou S. Medical Image Recognition, Segmentation and Parsing. Massachusetts, US: Academic Press; 2015.
14. Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, et al. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol 2017 Dec;2(4):230-243 [FREE Full text] [doi: 10.1136/svn-2017-000101] [Medline: 29507784]
15. Samadzadeh GR, Rigi T, Ganjali AR. Comparison of Four Search Engines and their efficacy with emphasis on literature research in addiction (prevention and treatment). Int J High Risk Behav Addict 2013 Feb 03;1(4):166-171 [FREE Full text] [doi: 10.5812/ijhrba.6551] [Medline: 24971257]
16. Ware M, Mabe M. The STM Report: An Overview of Scientific and Scholarly Journal Publishing. 2015 Mar. URL: https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1008&context=scholcom
17. Anders ME, Evans DP. Comparison of PubMed and Google Scholar literature searches. Respir Care 2010 May;55(5):578-583 [FREE Full text] [Medline: 20420728]
18. Shariff SZ, Bejaimal SA, Sontrop JM, Iansavichus AV, Haynes RB, Weir MA, et al. Retrieving clinical evidence: a comparison of PubMed and Google Scholar for quick clinical searches. J Med Internet Res 2013 Aug 15;15(8):e164 [FREE Full text] [doi: 10.2196/jmir.2624] [Medline: 23948488]

Abbreviations

AI: artificial intelligence
AR: augmented reality
IT: information technology

XSL•FO
RenderX

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analysis
VR: virtual reality

XSL•FO
RenderX

Original Paper

Defining Symptom Concepts in Chronic Subjective Tinnitus: Web-Based Discussion Forum Study

Alice Hibbert[1,2], BSc; Markku Vesala[3]; Micky Kerr[1,2], BA; Kathryn Fackrell[1,2], BSc, PhD; Stephen Harrison[3]; Harriet Smith[1,2], BSc; Deborah Ann Hall[1,2,4], BSc, PhD

[1]Hearing Sciences, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, United Kingdom

[2]National Institute for Health Research Nottingham Biomedical Research Centre, Nottingham, United Kingdom

[3]Tinnitus Hub Ltd, Hemsworth, United Kingdom

[4]University of Nottingham Malaysia, Semenyih, Malaysia

Corresponding Author:
Deborah Ann Hall, BSc, PhD
National Institute for Health Research Nottingham Biomedical Research Centre
The Ropewalk
Nottingham,
United Kingdom
Phone: 44 8232600
Email: deborah.hall@nottingham.ac.uk

Abstract

Background: A minimum standard based upon consensus decision making recommends a core set of tinnitus-specific health complaints (outcome domains) that should be assessed and reported in all clinical trials as this enables comparisons to be made across studies as well as data pooling for meta-analysis.

Objective: This study aimed to further clarify how the outcome domain concepts should be defined for 5 of the core set: tinnitus intrusiveness, sense of control, acceptance of tinnitus, concentration, and ability to ignore. This step requires a clear and fully elaborated definition for each outcome domain, moving from an abstract or a vague concept to an operationalized and measurable health-related construct, so that a suitable measurement instrument can then be identified.

Methods: A series of 5 focus group–style semistructured discussions were conducted via a Web-based discussion forum, each open for 2 weeks and ending with a vote. The participants included 148 tinnitus experts who completed a preceding e-Delphi survey that had generated the original set of minimum standards. The participants were health care users living with tinnitus, health care professionals, clinical researchers, commercial representatives, and funders.

Results: The Web discussions led to a revision of all 5 original plain language definitions that had been used in the preceding e-Delphi survey. Each revised definition was voted by 8 to 53 participants and reached the prespecified threshold of 70% consensus for all except tinnitus intrusiveness. Although a single definition was not agreed upon for tinnitus intrusiveness, the majority of participants shared the view that the concept should be sufficiently broad to encapsulate a range of subdomains. The examples included tinnitus awareness, unpleasantness, and impact on different aspects of everyday life. Thematic analysis of the 5 Web-based discussion threads gave important insights into expert interpretations of each core outcome domain, generating an operationalized and measurable health construct in each case.

Conclusions: The qualitative data gathered during the Web-based discussion forum provided an important in-depth understanding of the health concepts that had raised a debate during earlier face-to-face meetings. The descriptive summaries and definitions provide sufficient operationalization of those concepts to proceed to the second stage of core outcome set development that is to identify and evaluate suitable measurement instruments. This study supports the use of Web-based peer discussion forums in defining health concepts.

(Interact J Med Res 2020;9(1):e14446)   doi:10.2196/14446

KEYWORDS

patient outcome assessment; treatment outcome; concept formation; qualitative research; patient participation; community participation; stakeholder participation; Web social networking

XSL•FO
RenderX

Introduction

Background

Chronic subjective tinnitus is a condition characterized by a persistent auditory sensation (eg, ringing, whistling, hissing, and buzzing) experienced only by the individual, with no corresponding external sound or source. The characteristics and impacts of tinnitus are highly variable from person to person [1], and the outcomes reported in clinical trials of tinnitus interventions are similarly diverse [2]. This prevents the comparison of findings across trials and pooling data in meta-analyses, leading to a waste of research resources and an unreliable evidence base for making decisions about which interventions are most effective [3].

The development of core outcome sets (COSs) can tackle this issue by establishing a common standard and minimum set of recommended core outcomes for use in clinical trials of a specific condition or intervention type as well as for use in other types of research and clinical audit [4]. The Core Outcome Measures in Effectiveness Trials (COMET) handbook [5] outlines guidelines for best practice in COS development and advocates a 2-step approach to COS development. The first step considers *what* condition-related complaints should always be collected and reported. In this paper, the *what* is henceforth referred to as an *outcome domain*. Once an agreement has been reached regarding what should be measured, *how* those outcomes should be measured is then determined [6]. This 2-step process has the advantage of being able to define each outcome domain so that it is understood by patients and clinicians in a consistent way and also to identify gaps where further research would be needed, for example, if an outcome domain is seen to be of core importance but no adequate outcome measurement instrument yet exists.

Research Leading up to This Study

The research reported in this paper is part of a longer-term program by the Core Outcomes Measures in Tinnitus (COMiT) initiative that aims to establish a COS for clinical trials assessing interventions for chronic subjective tinnitus in adults [7]. The first step of COS development has recently been completed by the Core Outcome Measures in Tinnitus: International Delphi (COMiT'ID) [7-12] study. This first step developed minimum recommendations for what all tinnitus clinical trials should measure. The methods included a series of international electronic Delphi surveys and face-to-face consensus meetings in which core outcome domains were defined for the 3 most common tinnitus intervention types: sound, psychology, and pharmacology-based approaches [11,12]. A total of 73 candidate outcome domains were considered during this process and, for each outcome domain, a plain language concept definition was cocreated with patient and public representatives via qualitative methods [8]. These domains and definitions were evaluated by 719 stakeholders with expertise in tinnitus, including both health care users and professionals [10]. The result was an agreement on 9 distinct core outcome domains across the 3 intervention types [11], which is summarized in Figure 1 (adapted from [11]). These core outcome domains were then ratified through Web (email) votes opened to all of the original electronic Delphi survey participants [11].

Figure 1. Graphic illustrating the Core Outcomes Measures in Tinnitus: International Delphi recommendations for core outcome domain sets for each family of interventions widely available for chronic subjective tinnitus in adults. Core outcome domains highlighted in bold are those considered in detail within this paper and represent 5 out of the 9 distinct domains.
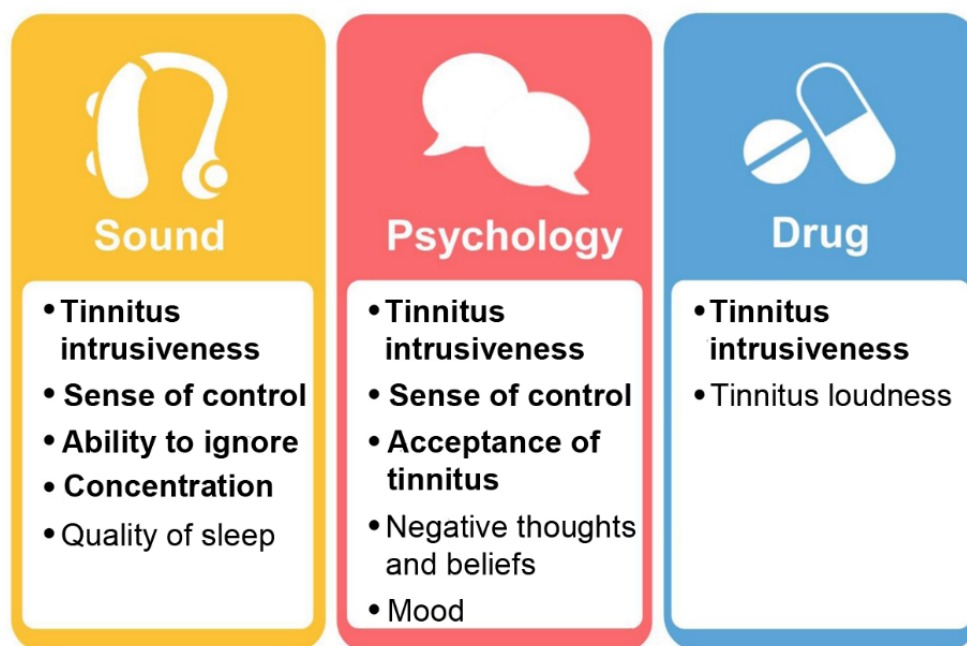
Defining Symptom Concepts

In accordance with the COMET handbook [5], the second step of COS development is to identify how each core outcome domain should be measured. This step requires a clear and fully elaborated definition for each, moving from an abstract or a vague *concept* to an operationalized and a measurable *construct*. This is emphasized by the consensus-based standards for the

selection of health measurement instruments initiative [13] that explains: "When selecting an outcome measurement instrument for research or clinical practice, first the outcome to be measured should be clearly defined... For example, when measuring a broad construct such as health-related quality of life, it should be clarified which subdomains are relevant for the target population in the specific context of interest. Sometimes several definitions exist for an outcome…Without explicitly defining or describing the intended outcome, people may have different ideas about it and interpret it differently." A detailed definition of the construct based on the specific area of health to which the core outcomes are to apply is a prerequisite for selecting an appropriate outcome measurement instrument [6]. The definitions and interpretations of 4 of the 9 core outcome domains (*quality of sleep, mood, negative thoughts and beliefs,* and *tinnitus loudness)* reached agreement by health care users and professionals during the COMiT'ID consensus meetings and so did not require further exploration and elaboration [11]. Reviewing those meeting discussions was considered sufficient to consolidate and finalize the definitions and conceptualizations of those 4 core outcome domains (see Multimedia Appendix 1). In contrast, those same discussions had highlighted the need for further work to specify and define the remaining 5 core outcome domains. First, it became evident that there were individual differences in the personal meaning attributed to certain domains and a lack of consistency in how they were understood by all stakeholders. Second, respondents made decisions to support certain outcome domains on the condition that their definition would be expanded to capture other outcome domain concepts (as subdomains). The purpose of this study was to specify and define these remaining 5 core outcome domains.

This study used some important new knowledge relating to the 5 core outcome domains of interest that had been gathered during the previous stakeholder discussions [11,12]. This can be summarized as follows: For *tinnitus intrusiveness*, there were a diverse range of interpretations about precisely what the concept entails. *Tinnitus intrusiveness* is commonly measured using a single-item numerical rating scale in which the concept is not defined [2]. Attempts to measure *tinnitus intrusiveness* as a construct using multiple questions assess a range of subdomains such as tinnitus awareness, loudness, unpleasantness, annoyance, and ability to ignore [14,15]. There was some discussion about potential negative connotations and misinterpretations of both *acceptance of tinnitus* and *sense of control*. These domains were often discussed together and compared with one another, but the exact relation and association between the 2 was unclear. Similarly, *ability to ignore* and *concentration* were often thought of as interacting with each other and were considered broad concepts that captured the essence of some of the other outcome domains that had been set aside from the core set. Furthermore, differing viewpoints emerged about whether *ability to ignore* should refer to change in the tinnitus itself or refer to an individual's personal capabilities. Resolving these debates is required to reach a common understanding of each construct so that it can be operationalized and mapped onto appropriate measurement instruments to ensure that the instrument has good content validity [6].

Currently, there are no formal guidelines to assist COS developers on how to further conceptualize and define outcome domains [5,6], and so COS developers have proceeded using different methods. For example, the World Health Organization's International Classification of Function, Disability, and Health has been used as a general reference framework for conceptualizing health-related quality of life in chronic pain [16] and in rheumatic conditions [17]. To help define atopic eczema flares, the Harmonising Outcome Measures for Eczema initiative first conducted systematic reviews of the literature [18,19] followed by a statistical evaluation of the performance of instruments measuring 2 alternative definitions of atopic eczema flares [20]. The first method relies upon symptoms being linked to concepts in the reference framework, whereas the second relies on a body of literature assessing the construct of interest. Neither method is suitable for chronic subjective tinnitus, where the core outcome domains of interest do not map well onto the World Health Organization's framework [1] and are not represented by an adequate body of literature [2].

Web-Based Discussion Forums in Core Outcome Set Development Work

Although the uptake of Web discussion forums as a platform for COS development is somewhat in its infancy, they have been successfully applied to evaluate the face validity of a new patient-reported outcome measure of treatment response in vitiligo [21], to explore patient perceptions of proposed core outcome domains for eczema [22,23], and to investigate patient priorities for a COS for pediatric acute respiratory illness [24]. Web-based platforms are now growing in popularity within the tinnitus research community, and self-help discussion forums are starting to be used for recruitment [25] and research data collection [26].

Aim

In summary, the aim of this study was to specify and define the 5 least well-defined core outcome domains recommended for clinical trials evaluating the effect of sound-, psychology-, and drug-based tinnitus interventions. The 5 core outcome domains in need of discussion were (1) tinnitus intrusiveness, (2) sense of control, (3) acceptance of tinnitus, (4) concentration, and (5) ability to ignore, and these were to be explored using a moderated Web discussion forum with representative stakeholders from the COMiT'ID study. The goals were to establish agreement on a single plain language definition describing each of the core outcome domains and to gain a more in-depth understanding of each concept that would then indicate what sort of questions would need to be asked when assessing each outcome.

Methods

Design

This qualitative study used a series of 5 focus group–style discussions conducted via a Web discussion forum. The Web discussion forum was chosen as a practical and cost-efficient research method for engaging with a large and geographically

distributed sample of participants, which could not be achieved by face-to-face methods.

Each Web discussion focused on a single core outcome domain that included *tinnitus intrusiveness*, *sense of control*, *acceptance of tinnitus*, *concentration*, and *ability to ignore*. The discussions were semistructured and ended with a voting phase that was focused on the executive summary of the discussion and the resulting concept definition, following recommendations by Im and Chee [27].

This study was conducted under a substantial amendment to the ethical approval originally granted for the COMiT'ID electronic Delphi and consensus studies by the West Midlands—Solihull Research Ethics Committee and Health Research Authority (reference 17/WM/0095, March 2017). This amendment was approved on September 18, 2017.

Recruitment and Participants

All registered COMiT'ID participants were invited by email to join the Web discussion forum. We had taken a number of steps to safeguard the relevant expertise (and hence representativeness) of these participants, and the details are published elsewhere [8,12]. We did not contact those who had explicitly withdrawn, and so 627 individuals were invited from a total of 641 unique individuals who had registered [10]. Regular updates via twitter [28], direct email, and at the international Tinnitus Research Initiative conference in March 2018 continued to encourage registration throughout the study.

The invitation and reminder emails contained a link to the discussion forum website [29] and a verification code that was required to register an account on the website. The code maintained privacy and security, ensuring that only those individuals who registered for the electronic Delphi survey were able to access the forum.

Eligible participants included members of the public with lived experience of tinnitus, health care practitioners, clinical researchers, and commercial representatives or funders. All the participants were targeted using a purposive sampling approach and had signed a self-declaration statement confirming that they met our eligibility criteria for having expertise on tinnitus. For full details see previous studies by Hall et al [11,12].

Procedure

Design of the Web-Based Discussion Forum

The website for the discussion forum was developed in partnership with the Tinnitus Hub [30]. Tinnitus Hub is a nonprofit organization that provides peer-to-peer support for those living with tinnitus and connects health care users with professionals conducting research. Tinnitus Hub hosts a peer support forum called Tinnitus Talk [31] that is one of the largest international tinnitus discussion forums and was selected for our research forum for its widespread reputation, secure platform, anonymity, and proven track record of engagement by the tinnitus community. This latter reason is particularly important given that people with tinnitus are older adults [32] who may be less familiar or comfortable with using Web discussion forums [33].

The Tinnitus Talk platform offered a number of positive design features well suited to the research aims and encouraging active engagement in the discussion. The participants could be individually distinguished, but their anonymity was preserved through the use of a pseudonym rather than their true name. The posts were automatically ordered chronologically, which allowed the discussion to be read as a conversation. A direct reply feature quoted the original post, avoiding the need to scroll back and forth through the discussion, and sent a notification to the author of the original post, nudging them to return to the forum and encouraging a natural flow back and forth, similar to face-to-face conversation. The participants could not alter their responses after posting but could add further comments to clarify or change their opinions.

Instructional videos were created to improve usability regardless of experience in Web discussion forums and technical ability. An introductory video on the homepage [29] guided participants through how to register and create an account (see Multimedia Appendix 2). Once logged in, a second video guided participants on how to write posts, reply to others, and set email notifications (see Multimedia Appendix 3).

Overall, 3 informational threads were open throughout the study: (1) an *Introduction,* which set out how the forum would run, with recommendations on how to take part, (2) *Guidelines and Ground Rules,* which stipulated rules such as respect for one another's opinions and expertise, and (3) *Tech Support FAQ,* which provided advice and solutions for common technical problems that might be encountered while using the Web forum. All participants were encouraged to read these threads and post a reply to practice using the forum software interface and to confirm that they had understood and agreed to follow the ground rules.

In addition, 5 further threads were used for each of the 5 focus group–style discussions, one for each core outcome domain (Figure 2). Individual threads overlapped in time so that the total discussion period was 6 weeks (Figure 2). All discussion threads were always visible, but they remained locked until the advertised opening date. The first discussion thread opened 2 weeks after the invitation email. Thereafter, 1 discussion thread opened each week. Discussion threads were open for 2 weeks, with semistructured discussion over 10 days and then a moderator-led summary and voting. Individual threads were purposefully ordered according to our expectation that engagement would be greatest during the middle of the study period. Therefore, we chose to place those outcome domains that had generated the widest debate across weeks 3 and 4.

Figure 2. Timeline illustrating the design of the Web discussion forum, with dates indicating the duration of each discussion topic. The Core Outcome Measures in Tinnitus International electronic Delphi survey was in its final (voting) stage at the launch of the forum. FAQ: frequently asked questions; Feb: February; Mar: March; Tech: technology.
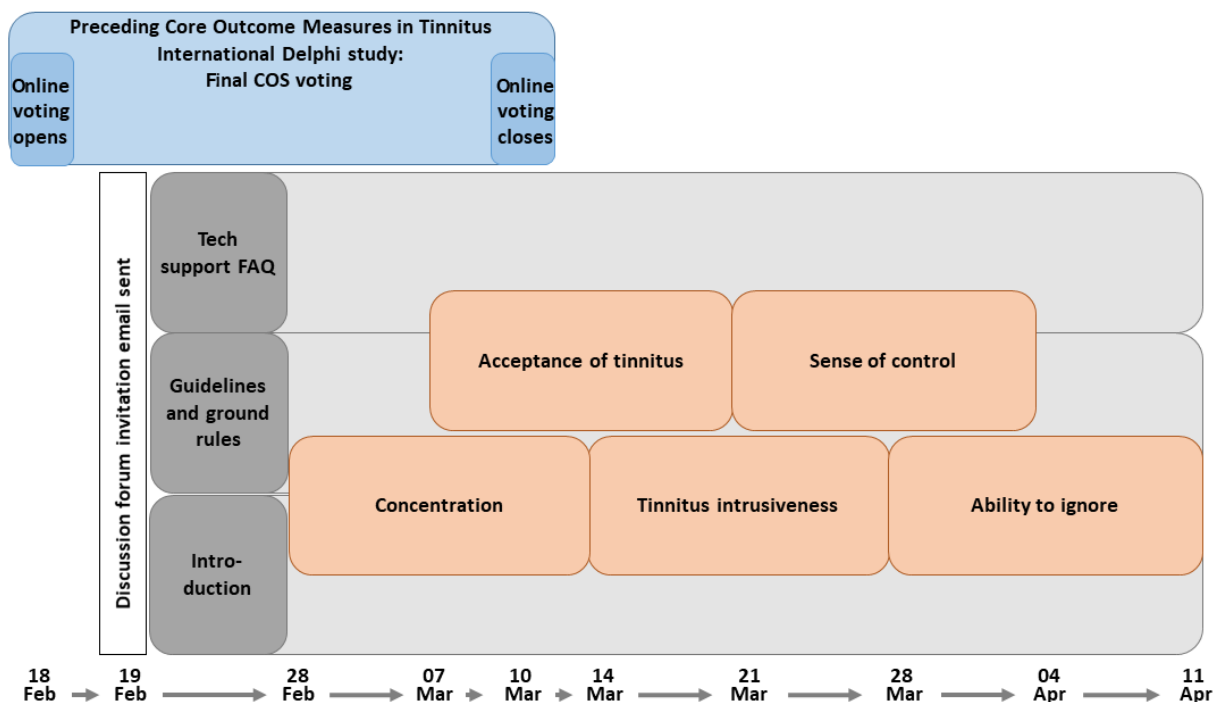
Methods for Reaching Agreement

All 5 discussion threads started with a reminder of the plain language definition for the outcome domain given during the electronic Delphi survey [8,9]. Most of the 2-week period comprised the semistructured discussion (see Multimedia Appendix 4) with a series of questions and discussion prompts that followed the natural flow of the conversation as much as possible. Throughout the semistructured discussion, participants were encouraged to not only answer the questions and discuss the concepts but also to suggest revisions for the plain language definitions.

The moderator then gave a brief executive summary of the discussion and proposed a final revised plain language definition. Wherever possible, the revised plain language definition was proposed, refined, and supported by participants during the discussion. Where this was not possible, the moderator developed a revised plain language definition based on the key themes raised during the discussion. These key themes were identified by a preliminary qualitative analysis of the discussion content conducted by the independent moderator as the discussion was unfolding.

Participants were asked to cast their vote according to 4 options: (1) agree with both the summary and definition, (2) agree with the summary but not the definition, (3) agree with definition but not the summary, and (4) disagree with both the summary and definition. Those who disagreed were asked to explain their reason and to recommend any changes. The Web platform used for the forum restricted voting to holding 1 per discussion thread. So, in cases where there was more than one plain language definition, participants were asked to choose their preferred definition. Any disagreements with the summary were collected

in the form of written viewpoints instead of an actual vote. Consistent with the preceding electronic Delphi survey, at least 70% agreement across respondents was considered the threshold for accepting the summaries and definitions [5,8].

Moderation Style

Each discussion was led by an independent moderator (AH) who had experience in leading focus groups and had not been involved in earlier stages of COMiT initiative work or experienced tinnitus herself. However, she did undergo a period of familiarization with the work conducted to date, including listening to recordings of the face-to-face consensus meetings. The moderating style was flexible, becoming more or less active depending on the degree of participant engagement. The aim was to foster a natural conversation style between participants as would happen during a face-to-face focus group. The most desirable style of conversation was one in which participants clearly explained and responded to one another's personal perspectives to reach a shared understanding. Posting of monologues or isolated messages to the moderator was discouraged. To promote this desired style of discussion, the moderator posed carefully worded (ie, nonleading) questions to encourage participants to elaborate and to give specific examples where more detail would be useful. The moderator also regularly reemphasized the key questions and topics to cover, reinforcing and thanking participants for their contributions and bringing them into conversation with one another on occasions when they had shared either similar or contrasting perspectives. When disagreements or tension arose, the moderator reminded participants of the purpose and context of the discussion forum and the ground rules.

The moderator also maintained a high degree of contact with participants to prompt and remind them of the structure and next stages of the procedure. For example, reminder emails were sent to highlight the next discussion thread opening and to encourage voting before a discussion thread closed. In response to participant feedback, the moderator updated the first post in each discussion thread so that it summarized each key question asked in the forum, with hyperlinks to the corresponding post. The intention was to ease the burden on participants, allowing them to make informed contributions without necessarily needing to read every single post in the thread.

Analysis

Executive Summary and Plain Language Definition

The first stage of the qualitative analysis relating to *tinnitus intrusiveness*, *sense of control*, *acceptance of tinnitus*, *concentration*, and *ability to ignore* was to generate a brief executive summary and, if needed, to revise the plain language definition for each of these outcome domains. These outputs were generated while the discussion thread was open so that they could be used in the Web-based voting phase.

The moderator (AH) prepared each executive summary to address 3 aspects of the discussion. One aspect of the discussion concerned key concepts that were discussed and new themes that emerged. The moderator determined what was key based upon themes that were most often mentioned and talked about by participants, and what appeared to be most relevant to the study aim. The second aspect concerned any recurring themes or strongly dissenting concerns that might necessitate a revision to the plain language definition. The third aspect concerned views where a concept for one core outcome domain seemed to converge with another concept that had previously been set aside from further discussion (either during the electronic Delphi survey or consensus meeting) [11]. The executive summary was written in such a way that a vote of agreement indicated support for all 3 aspects. These 3 discussion points might reasonably necessitate a revision to the plain language definition. Where possible, the revised plain language definition used phrases given by the participants. In cases where participants had offered multiple definitions, the moderator selected the one that appeared to reflect the majority's viewpoint. If this was not possible, then all of the candidates were asked to vote.

In-Depth Understanding of Each Concept

In addition to the moderator-led executive summary, thematic analysis was applied to a download of the entire 5 discussion threads conducted by 2 analysts (AH and MK). MK provided an independent perspective as he was naïve to the project. The 2 analysts independently examined each discussion thread separately and, in the order, that they took place. Methods for thematic analysis followed Braun and Clarke's 6-stage framework [34], and qualitative analysis was conducted using NVivo Pro 11 software (QSR International Pty Ltd, version 11). The text was read several times for familiarization before it was coded. Emerging themes were identified by grouping codes, and they were then refined and defined through an iterative process. Coded text segments provided the evidence corresponding to each theme. Once they had completed this process independently, the 2 analysts met together with the principal investigator (DAH) to compare their independent analyses and codebooks, with the intention of identifying and validating key themes to be reported for each outcome domain. To achieve this, similar themes were merged, any discrepancies were resolved, and those themes most relevant to the study objectives were identified.

Results

Participants and Engagement

Of the 627 individuals invited, 251 registered for the Web discussion forum, leading to a recruitment rate of 40.0% (251/627). Of these 251 participants, 119 submitted one or more posts to the discussion forum. Henceforth, these are referred to as *discussants*. An additional 29 participants did not submit any posts but took part by voting in at least one of the discussion threads, leading to an engagement rate of 59.0% (148/251). To preserve anonymity, we did not request information about stakeholder group membership. However, many discussants freely disclosed their stakeholder affiliation in the content of their posts. Although these data are indicative not definitive, 53.7% (64/119) participants identified themselves as health care users living with tinnitus and 25.2% (30/119) as professionals, and the remaining 21.0% (25/119) were unknown.

Similar degrees of activity were sustained across the 5 discussion threads (Table 1). Inevitably, some discussants were more active than others, but no individuals sought to dominate the conversation. Some held strong personal viewpoints, for example, "The reason I feel strongly about the negative impact side of intrusiveness is from my personal experience of tinnitus." However, these tended to be internally regulated by the forum discussants, for example, "I don't think you can compare one set of experiences to another fully because of the wild variety of coping strategies that people will use." Although the number of posts and discussants was broadly equivalent across the 5 discussion threads, the number of voters was markedly low for *concentration*, perhaps because it was the first thread and participants were still familiarizing themselves with the structure of the forum discussions and with the limited time window available for voting. The number of voters were also relatively low for *tinnitus intrusiveness*. The possible reasons for this are less clear, but several speculations can be offered. For example, the voting options for tinnitus intrusiveness required participants to choose between several potential definitions rather than simply agree or disagree with a single definition, and this different voting process may have been less appealing or confusing to participants. For example, 3 discussants very active in the forum when the vote was open preferred to share their view by submitting multiple written posts that commented on the voting options available rather than casting a vote.

Table 1. Participant activity across the 5 discussion threads evidenced by the total number of posts, number of unique individuals who submitted posts (discussants), and the number of unique individuals who voted (voters). Note that the number of posts includes those submitted by the moderator.

Participants (type)	Number of participants in each discussion thread				
	Tinnitus intrusiveness	Sense of control	Acceptance of tinnitus	Concentration	Ability to ignore
Posts	160	133	147	142	145
Discussants	49	44	54	49	47
Voters	20	33	53	8	44

The level of activity on the Web discussion forum is plotted in Figure 3 across the relevant 8-week period from the first email invitation being distributed to the closure of the last discussion thread. This displays that engagement was sustained fairly consistently throughout the procedure and demonstrates the benefit of sending reminder emails and regular updates via other social media.

Figure 3. Graph showing the number of user registrations and forum posts over the 8 weeks, from the initial invitation email to closure of the final discussion thread. The hashtag symbol (#) indicates the first mass invitation email, and asterisk symbols (*) indicate subsequent reminder emails. Reminders generally coincided near the start and midway through each discussion thread but were required less toward the end of the study. Apr: April; Feb: February; Mar: March.
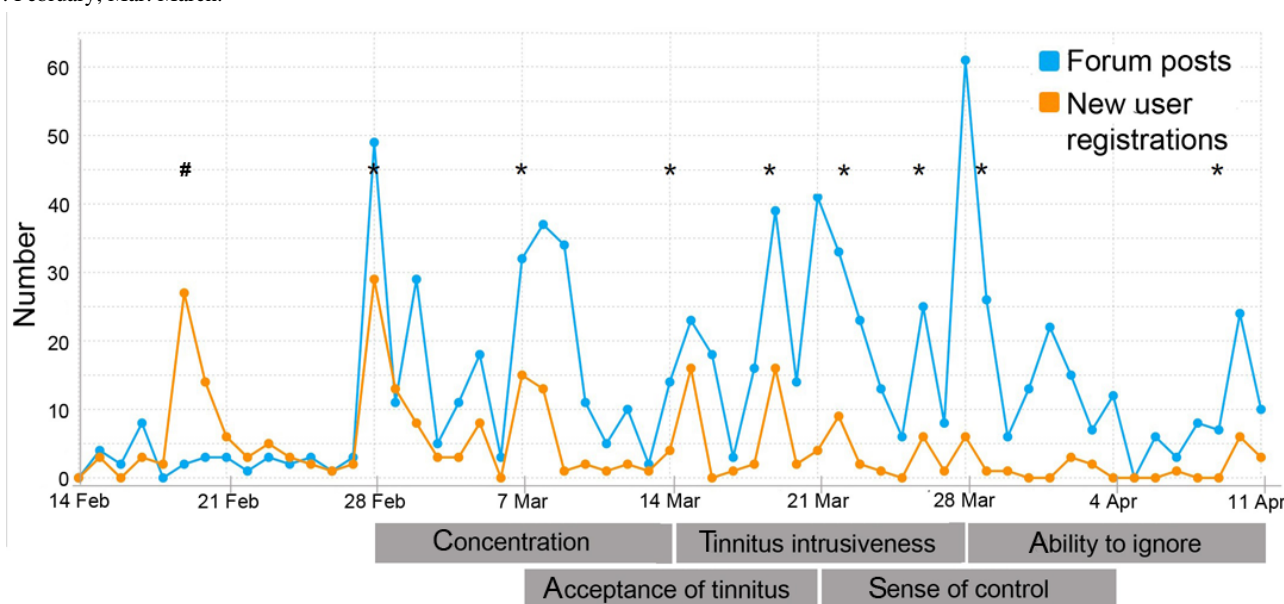
Executive Summaries and Plain Language Definitions

The executive summaries created from each Web discussion forum are presented in Multimedia Appendix 5, and the plain language definitions and percentage of supporting votes are presented in Table 2. Table 2 also includes expansions to each concept with regard to which subdomains should be included in the operational definition, as recommended by participants in the discussion forum. Subdomains denoted as *(maybe)* are those suggested by a few of the discussants only and so are not critically important to include in the operational definition.

In the preceding consensus meetings, what constitutes *tinnitus intrusiveness* had generated the widest debate, and this was also true for the Web discussion forum. Several different versions of a definition were proposed but none seemed to gain majority support. A total of 7 alternative versions were put to the Web vote, but none reached the required 70% agreement to be able to conclude that a consensus had been reached. These different views on what tinnitus intrusiveness means for a person with tinnitus were consistent with the numerous suggestions to broaden its description to encapsulate other domain concepts such as the impact on different aspects of life, unpleasantness, and annoyance (Table 2).

It was a recurring theme during both the consensus meetings discussing sound-based and psychology-based interventions that *sense of control* and *acceptance of tinnitus* might converge onto a similar concept, along with the set aside domain *coping*. The Web discussion forum gave greater clarity on the distinction between these 2 core outcome domains. There was consensus that *sense of control* refers to the feelings achieved once a treatment or coping strategy has been found that provides relief, whereas *acceptance of tinnitus* refers to the general feeling of being at peace with the tinnitus and no longer fighting against it. Some of the discussants indicated a temporal order whereby *acceptance* might follow the (re)gaining of a *sense of control*. Although these 2 concepts are undeniably related, the definition for a *sense of control* perhaps encapsulates a situation-specific feeling that is related to actively managing the tinnitus, whereas the definition for *acceptance of tinnitus* encapsulates a more general feeling about letting go of the resistance to and distress caused by the tinnitus.

Table 2. The plain language definitions for each of the final core outcome domains, and percentage of voters who supported the revised plain language definition where this was changed during the course of the Web discussion forum.

Outcome domain	Original plain language definition	Revised plain language definition	Number of agreements among voters, n (%)	Subdomains
Tinnitus intrusiveness (n=20)[a]	Noticing the sound of tinnitus is there and it is invading your life or your personal space	• The extent to which tinnitus invades your life, stresses you in daily situations and prevents you from doing things you want to do • The unacceptable and unwelcome interference of internal head and body noise heard only by the individual • Being acutely aware of the sounds of tinnitus, feeling that it is invading your life or your personal space, changing your thoughts or actions and negatively impacting on your life	• 7 (35) • 3 (15) • 4 (20)[b]	• Tinnitus awareness • Tinnitus unpleasantness • Impact on individual activities • Impact on social life • Impact on relationships • Impact on work (maybe) • Annoyance (maybe)
Sense of control (n=33)[a]	Whether or not you feel you have a choice in how to manage the impact of tinnitus and feelings caused by tinnitus	• Feeling that you have effective options for managing the impacts of and feelings caused by tinnitus, through an understanding of your condition, learned strategies, and/or available resources	• 27 (82)	• Coping
Acceptance of tinnitus (n=53)[a]	Recognizing that tinnitus is a part of your life without having a negative reaction to it	• Recognizing that tinnitus is a part of your life and staying neutral toward it in both thoughts and actions	• 44 (83)	• N/A[c]
Concentration (n=8)	Ability to keep your attention focused	• The ability to keep your attention focused on whatever you wish	• 8 (100)	• Impact on work (maybe)
Ability to ignore (n=44)	Ability to continue as normal as if tinnitus were not there	• The ability to continue as if tinnitus were not there	• 35 (80)	• Conversations • Listening

[a]Number of voters.

[b]Synthesized from comments made by 4 (20) voters who could not choose between the given definitions.

[c]Not available.

Similarly, the Web discussion forum gave greater clarity on the distinction between *concentration* and *ability to ignore*. There was consensus that *concentration* refers to the ability to focus on a specific task or activity that requires full attention, whereas *ability to ignore* describes the capacity to focus away from tinnitus in most if not all situations. Some of the discussants suggested that *ability to ignore* may be more relevant to low-demand tasks, whereas *concentration* may be more relevant to high-demand, complex tasks.

In-Depth Understanding of Each Health Concept

The themes emerging from the discussion thread under each core outcome domain are reported in this section (but in no particular order).

Tinnitus Intrusiveness

Overall, 2 major themes emerged from the discussion thread: one highlighting the importance of the negative impact of tinnitus on everyday functioning and another highlighting the strength of its negative emotional impact.

Negative Functional Impacts of Tinnitus, Not Just Awareness of Its Presence

A prominent debate concerned whether the concept of *tinnitus intrusiveness* should reflect the ongoing presence of tinnitus and the individual's awareness of it, or whether it should go beyond this to capture the negative impacts of tinnitus on everyday life. Although the discussants' initial preferences seemed divided, by the end of the discussion thread, the overwhelming majority of posts supported the latter interpretation. For example, to argue against awareness of the presence of tinnitus, one discussant said: "it's not the tinnitus itself that's the problem it's the fact that it's invasive and having invaded it causes problems—specifically that it makes doing certain things (sleeping, socialising, whatever) either less do-able or at least less enjoyable." (This participant later stated that they had meant to use the word "intrusive" rather than "invasive").

The course of the discussion was guided by the ultimate objective of measuring *tinnitus intrusiveness* to assess whether an intervention is effective. For example, to argue in favor of the negative impacts, 1 discussant said:

I know it's possible to have noticeable but not disabling tinnitus... A treatment should aim to combat the negative effects. Begs the academic question of whether a treatment might work for people for whom tinnitus was never a problem, just something they noticed.

Consistent with this, discussants voiced support for *tinnitus intrusiveness* including facets of at least 3 other functional outcome domains (*impact on individual activities, impact on social life,* and *impact on relationships*). These had previously been agreed as critically important during the electronic Delphi survey, but had been set aside from the final core outcome domain sets. For example, 1 discussant wrote: "By its very nature, tinnitus can 'intrude' anywhere so the core outcome of 'intrusiveness' could quite easily apply to all of the impact outcome domains."

Negative Emotional Impact and Its Potency

Another prominent theme concerned the emotional reaction to the intrusiveness of tinnitus. Discussants emphasized the potential for extreme suffering by describing how insufferable, intolerable, and invasive tinnitus could be. Different definitions of intrusiveness were drawn upon to help understand the depth of the concept, particularly exploring its personification. For example, 1 discussant wrote:

Intrusiveness itself is the inability to keep something unwelcomed be it physical or non-physical such as a thought, from infiltrating your mind/personal space without your permission. Tinnitus is like a burglar that enters your home, holds you hostage, but instead of stealing your belongings, it robs you of your sense of peace. It is therefore, intrusive and dominating.

Consistent with this, discussants voiced support for the inclusion of *tinnitus unpleasantness,* which had previously been agreed as a critically important domain during the electronic Delphi survey but had been set aside from the final core outcome domain sets. For example, 1 discussant wrote:

I am content with the definition of tinnitus intrusiveness covering all the impacts, unpleasantness and awareness. As I am writing I can hear my tinnitus very clearly. It is not unpleasant as such, but a nuisance I could do without. As I have already said at first my tinnitus was always unpleasant and at its worst I had to really concentrate on not giving way to panic. Those days have gone and I hope they never return.

To conclude, there was no majority agreement on a revised definition, but *tinnitus intrusiveness* was generally agreed by discussants to be a broad concept with subdomains, referring not only to the extent to which tinnitus has an unwanted presence and is deeply unpleasant but also the extent to which it negatively impacts upon daily life and activity. Hence, the construct *tinnitus intrusiveness* embodies both interference with functioning and psychological distress associated with tinnitus.

Sense of Control

Overall, 3 major themes emerged from the discussion thread: 1 highlighting the importance of autonomy in how an individual manages his/her own tinnitus, 1 highlighting the associated sense of empowerment, and 1 highlighting the importance of actively adopting management strategies.

Autonomy

The concept of free agency was related to the *sense of control*; the importance of an individual's capacity to act independently and to make free choices and about how and when to apply management strategies. It was generally felt that this facet of *sense of control* required different options to be available and to be offered (by the health care professional). Autonomy was a phrase used by several discussants. For some, it was also important that the options were effective. It is important to note that making the choice *not* to do anything was considered equally relevant and acceptable. For example, 1 discussant wrote: "In some instances my tinnitus definitely defines some actions, entry into very loud shopping spaces and restaurants for instance are a step too far... Ear plugs help, but then leaving and being in control of that is a better choice. No [sic] convenient but better."

Empowerment

A sense of free agency promotes self-efficacy, and so, one emerging theme emphasized the emotional consequence of feeling in control and being able to make personal choices. One example is the post: "When you are able to manage the impact/response to some extent and continue to enjoy life (by focusing on what you can do and reframing what you can't) you overcome the helplessness and achieve some sense of control." This positive sense of self was described by discussants as self-confidence, empowerment, and self-efficacy, whereas the converse was described as being at the mercy of tinnitus, leading to despair.

Active Management Strategies, Not Just State of Mind

Another emerging theme emphasized that *sense of control* should reflect a context in which the individual practices an active tinnitus management strategy. Some discussants expressed strong opinions that the concept of *sense of control* should not be restricted to the view that tinnitus suffering can be alleviated simply by changing negative thoughts or by adjusting one's state of mind. One discussant said: "This control has to have been achieved as the result of an intervention—as others have said—it is not a 'state of mind' it doesn't come about by 'positive' thinking." Instead, it was agreed during the voting phase that *sense of control* should refer to an active, practical, autonomous approach to tinnitus, which may involve an array of interventions, strategies, treatments, tools, techniques, resources, and aids.

To conclude, discussants considered that *sense of control* referred to being in possession of active management options to cope with the impacts of and feelings caused by tinnitus. The revised definition was as follows: "Feeling that you have effective options for managing the impacts of and feelings caused by tinnitus, through an understanding of your condition, learned strategies and/or available resources."

Acceptance of Tinnitus

Overall, 3 major themes emerged from the discussion thread: 1 highlighting the importance of handling negative reactions, 1 highlighting reconciliation of one's own identity, and 1 emphasizing the ongoing struggle to maintain acceptance.

Becoming Less Reactive to Negative Reactions

The original plain language definition stipulated that *acceptance of tinnitus* was never having a negative reaction toward it. Some discussants felt this definition was "unrealistic...[and] inadequate in its understanding of the ups and downs of living with constant ringing in your ears," particularly, those that interpreted not having negative reactions as implying the need for positive reactions. It was instead argued that a healthy level of acceptance is just getting better at managing and not responding to negative thoughts, bringing yourself back to a place of neutral balance, and choosing not to fight against the tinnitus. Useful comparisons were drawn to practicing meditation, acknowledging when your mind wanders, and calmly bringing it back to where you want it to be until it gradually becomes easier to do so and wanders less often. As 1 discussant stated, sometimes "it's alright to have a negative reaction to tinnitus. Just let it be. It will pass."

Self-Identity When Living With Tinnitus

During the discussion of *acceptance of tinnitus*, the theme of identity emerged as a complex and individualized issue. For some, recognizing tinnitus as a part of themselves was central to acceptance, whereas for others, it was important to distinguish between themselves and their tinnitus, maintaining their own identity as more than that. Compromise was found in the stance: "tinnitus is just small part of who you are but that it doesn't define or control you." This battle to rediscover and reconcile one's own identity, classifying tinnitus as either internal or external to the self, seems integral to reaching acceptance, which was spoken about as "coming to terms with what is and not fighting against it," and becoming able to "coexist with [their] tinnitus rather than see it as the enemy." This seems to mark a turning point where those living with tinnitus begin to allow themselves to make empowered choices taking their tinnitus into consideration rather than struggling to always act in spite of it. One professional defined *acceptance of tinnitus* from their experience with patients as: "people choosing to explore living well even if tinnitus may be with them permanently."

The Ongoing Struggle of Acceptance

Some discussants had very strong negative reactions to this concept, demonstrating hostility and frustration toward the idea of being told to *accept* their tinnitus which to them seemed completely intolerable or even harmful. To these people, notions of endurance, resilience, and tolerance seemed to be more conceivable milestones along the way. Even among those individuals who spoke more favorably of the concept of *acceptance of tinnitus*, it was clear that this is not something that is achieved and then easily maintained forever: "Acceptance is a process that comes and goes. It is a continuous process. The definition is too static." It became apparent that any measurement of *acceptance of tinnitus* would have to understand and integrate an element of time to be accurate and recognize that those living with tinnitus "will have good days and bad days."

To conclude, the revised definition of *acceptance of tinnitus* was "Recognising that tinnitus is part of your life, and staying neutral towards it in both thoughts and actions." The concept was described by discussants as a highly individual and challenging experience, marked by one or many turning points by which achieving a sense of self-identity as someone living with tinnitus gives a greater sense of peace, and the struggle of living with tinnitus becomes somehow easier.

Concentration

Overall, 3 major themes emerged from the discussion thread: 1 emphasizing the importance of intentional control, 1 highlighting the unavoidable prominence of tinnitus, and 1 drawing attention to the resulting cognitive effort and mental fatigue required when concentrating.

Intentional Control

One emerging theme was the importance of being able to control your *concentration,* including the ability to focus on whatever you wish or need to focus on at any given time and the ability to control the switching of attention. One discussant said: "So concentration is the ability to exercise attentional control, to stay focused on something of our own choice. This ability is impeded when you have tinnitus."

Centrality of Tinnitus

Discussion of *concentration* demonstrated the unavoidable prominence of tinnitus. Discussants struggled to suggest plain language definitions for the outcome domain that did not include the word tinnitus, despite specifying that this was undesirable as it gave more power or contradicted the notion of *concentration* being the ability to focus elsewhere and on other things. For example, 1 discussant wrote "I feel with the wording 'away from the tinnitus', just by virtue of that wording and sentiment implies you are trying so hard to distract yourself that the tinnitus remains the focus."

Effortfulness

The concept of *concentration* was agreed to encompass not only the ability to focus and function cognitively, but also the additional mental effort required because of the presence of tinnitus and the subsequent fatigue caused by that effort. One discussant said:

> I would consider a sound-based treatment successful if it restored, even partially, my ability to immerse myself in a task and for this to feel less of an effort than it is now. Ideally, the treatment should reduce the occurrence of cognitive tiredness that makes sustained concentration difficult.

To conclude, the revised definition of *concentration* was as follows: "The ability to keep your attention focused on whatever you wish." This was described by discussants to be referring to the ability to control your attention and sustain focus on whatever it is you intend to focus upon, with successful tinnitus interventions enhancing the ability to concentrate by making it easier and less effortful.

XSL•FO

RenderX

Ability to Ignore

Overall, 2 major themes emerged from the discussion thread: one debating whether the *ability to ignore* should be attributed to the tinnitus sound itself or to individual capabilities and another raising concerns about the negative connotations of the choice of wording, including whether or not ignoring tinnitus represents a realistic goal.

Changes in the Tinnitus Percept, Not Just in the Individual's Capabilities to Ignore

A key debate focused on whether the concept referred to any change in the tinnitus percept (ie, making the noise easier to ignore) or in the individual's capability (ie, making the person better at ignoring tinnitus). Some discussants expressed no strong preference: "Either outcome/effect (change in tinnitus or in person) would be good." However other discussants agreed that a reasonable expectation for sound-based treatments, the only intervention type that this outcome domain is recommended for, is to improve *ability to ignore* by a third mechanism that falls somewhere in between, "a sound-based treatment can't really affect your personal abilities (in the same way as a psychological treatment might) but nor can it change the tinnitus itself (from my experience). To give an analogy, glasses don't change my eyesight or improve my ability to be observant—they just allow me to see better whilst I am wearing them." As it was not unanimously resolved, this theme did not lead to any change to the plain language definition of *ability to ignore.* However, it was still a substantial theme that emerged from the discussion and may be informative in next steps deciding how the *ability to ignore* should be measured, as what sort of questions should be asked is best governed by what sort of change a successful intervention is likely to create and what sort of change would be meaningful to patients.

Negative Connotations and Unrealistic Goals

Concerns were voiced about negative connotations of the word *ignore* and discussants were worried that it could be seen as dismissive toward tinnitus or as blaming people living with tinnitus for their own suffering. As 1 discussant reasoned: "Replace tinnitus with any other disease: as if cancer were not there, depression were not there... and suddenly it makes no sense at all." There were also criticisms of the original definition of this outcome domain as an unrealistic or unachievable goal, and for its use of the term *normal*: "What is normal? We all change and adapt to what life throws at us."

To conclude, discussants seemed most supportive of an understanding of the *ability to ignore* as adapting to tinnitus, adjusting daily life and routine activities so as to minimize the negative impact of tinnitus, and maintaining healthy and realistic goals for how a treatment may be able to help make tinnitus easier to ignore in certain situations or to certain extents. The revised definition was as follows: "The ability to continue as if tinnitus were not there."

Discussion

Principal Findings

The 5 Web discussion forums brought together a self-selected subset of survey participants, including health care users and professionals with experience of tinnitus. Participants took part in semistructured discussions of 5 complex concepts relating to patient-reported tinnitus-specific complaints that had been voted during a preceding e-survey. The qualitative data collected during these discussion threads provide an important in-depth understanding of each health-related concept, which had not been possible hitherto. The descriptive summaries and revised definitions also provide clarification on aspects of similarity and distinctiveness between core outcome domains. These findings are informative for identifying outcome instruments that putatively assess these concepts and for evaluating their content validity [6,13,35].

Despite the paucity of qualitative data from people with tinnitus [1], 1 study using interviews to explore preferences for outcomes and treatments confirms 3 of the present construct descriptions [36]. In response to a question about what they were hoping for in a treatment, patients described their preference for a "reduction of conscious awareness of the tinnitus (to reduce time listening to tinnitus)." This notion reflects a theme from the Web forum discussion on the *ability to ignore*. Preferences for how they would like to receive treatments highlighted the importance of "choice in personalising their care and determining the best course of action for them," which is synonymous with our interpretation of the construct *sense of control.* Psychological adjustment (described by 1 participant as "you've got to learn to accept it") was understood to be an active part of coping with tinnitus in a similar way that we have described the construct *acceptance of tinnitus.* Although the study by Pryce et al [36] was exploratory and only interviewed 41 patients based in the United Kingdom, it is nevertheless important because it presents an independent look at similar issues and was unpublished at the time this study was ongoing.

Strengths of Using a Web-Based Peer Discussion Forum for Core Outcome Set Development

An increasing number of social and health science researchers are recognizing the internet as a rich source of information. A Web discussion forum facilitates participation by any number of individuals in a way that is not constrained by geographical location or time zone. It offers a rapid and easy way to engage with a large number of participants whilst being more flexible and cost-effective than conventional face-to-face methods. Furthermore, the data are already transcribed, and so they are less likely to contain errors and are immediately ready to analyze (see the study by Ferrante et al [37] for a review). Unsurprisingly, an increasing number of investigators are designing Web discussion forums to collect qualitative research data from patients and using thematic analysis to evaluate forum posts [38-40]. This study contributes to the increasing use of electronic communication to support group decision making and consensus making [41].

A major strength of using such a virtual environment is that it allows investigators to conduct real-time qualitative analyses as part of an iterative process in which the participants are actively involved in determining the meaning and significance of findings and where the moderators are able to consolidate, clarify, and resolve any misunderstandings for the purposes of concept definition. The Web format seemed to provide a suitable space that enabled participants to reflect and share ideas about word choice and semantics. With the exception of *concentration*, all discussion threads had a rich debate on choice between alternative wordings. For example, 1 discussant said: "Clearly the same words can mean or imply very different things to different people. That's inevitably going to be a big issue with tinnitus which is so individual to each person affected by it." Written forum posts, perhaps more so than spoken conversations, lend themselves to greater deliberation over the selection of a particular word according to how it might be interpreted. Discussion about the words *acceptance* and *ignore* are good examples. As 1 discussant said:

> *Imagine not being part of this forum, but being told by your doctor that some new treatment leads to your being able to ignore the tinnitus, only to find out that it doesn't for you, and then only worked in the research because the definition was engineered in a particular way... I suppose what I'm saying is that ability to ignore (and the other definitions) have to reflect what most tinnitus sufferers would think if they heard that phrase from their doctor or therapist.*

In this way, discussants were not just research participants, they also played an important role in shaping the research *product*. The Web discussion forum confirms the usefulness of each 2-week discussion period, to consolidate, clarify, and resolve any misunderstandings.

One of the positive themes emerging from the Web forum discussion was the therapeutic benefit of participation. Several discussants living with tinnitus thanked the moderator and research team, expressing a sense of reward in having taken part and describing how it had been personally enlightening and therapeutic. One said: "This is better than I felt 2 years ago, and I must admit that this academic research group with fellow sufferers has been a part of that improvement." Another said: "Lastly I just wanted to say thanks to 'Manager' and to COMIT'ID and everyone else for all of this. It's been really good/therapeutic for me to have been a part of it, and to be able to hear all the experiences of everyone else." Although Web discussion forums are commonplace for peer support groups [42], similar benefits through sharing ideas and experiences during a research study should not be overlooked. It was not necessary for the moderator to offer support, as the participants took on that role themselves (see also the study by Ferrante et al [37]). We feel it is important to acknowledge that this happened despite the identity of health care users or professionals not being known. This may have helped to create a space without hierarchies where interpersonal relationships were on a level playing field. For example, 1 discussant responded to another:

> *I don't think anyone is nit-picking. The purpose of these discussions is to better understand what the core outcomes mean to everyone and reach a better agreement about their definitions. Everyone's views are valid, equally important to hear and worthy of respect. The whole point of the research is to achieve as broad a consensus as possible and this can only be achieved by exploring and discussing where and how our views align and where they differ.*

A mix of health care users and professionals should also help to avoid some of the potential for bias in the design and interpretation of the study if carried out only by a particular stakeholder group (eg, health care users) [5].

Limitations of This Web-Based Peer Discussion Forum for Core Outcome Set Development

A potential limitation of using Web forums for data collection is that participants need to be computer literate and able to communicate adequately in written English. This may limit the population somewhat and bias self-selection toward those who are more health literate. For this particular study design, it is possible that opening and closing discussion threads in sequence could have resulted in participants who joined later on in the 6-week process missing the opportunity to share their viewpoints on earlier outcome domains. It is also possible that within a different context or procedure, more themes could emerge as it was not possible to ascertain whether data saturation was reached by the forum discussions [43,44]. However, the main practical study objective was to enable robust decision making for the 5 core outcome domains in a time-limited way. We believe that this was successfully achieved at least for the sense of control, acceptance of tinnitus, concentration, and ability to ignore. In the case of tinnitus intrusiveness, the Web discussion forum was perhaps more limited in its ability to converge opinions onto a plain language definition as the concept appears to be particularly complex and viewpoints are more variable.

Implications for Future Research

Expanding the Subdomains Encompassed by the Concept Tinnitus Intrusiveness

During preparation for the electronic Delphi survey, the COMiT team, with input from health care users, had made a decision to narrowly define the outcome domains, removing broad concepts that were reflected in a number of more narrowly focused outcome domains [9]. However, during the face-to-face consensus meetings and these Web discussion forums, participants argued for a different approach, noticing where concepts were interrelated and favoring to nest those interrelated concepts under a broader construct definition. For example, health-related quality of life was originally deemed to be a broad concept encompassing subdomains such as *impact on relationships, impact on individual activities*, *impact on social life*, and *impact on work* [9]. We had considered these subdomains as distinct outcome domains in their own right, but the overwhelming opinion of stakeholders was that they should be incorporated into the construct *tinnitus intrusiveness* [11]. The next challenge will be to examine the spectrum of symptoms and aspects of functioning and health that health care users,

health care practitioners, clinical researchers, and other stakeholders, such as regulatory agencies, expect to be covered in a measure of *tinnitus intrusiveness*. Precedents for this next step exist in other disciplines, such as deciding how to measure quality of life in chronic pain [16] and quality of life in adults with eczema [45].

Evaluating the Content Validity of Existing Instruments

Content validity refers to a number of key attributes of a measurement instrument, namely, how relevant the items are for the construct and target population of interest, how comprehensively those items reflect the construct, how comprehensible the instrument is, and whether it is understood by patients as intended [35]. Content validity is often considered to be the most important measurement property of a patient-reported outcome measure because its lack can undermine all other measurement properties. The rich personal insights reported in this study provide a firm foundation for defining exactly what symptoms and aspects of functioning that health care users might expect to be covered in the measurement tools that assess the 5 concepts of interest. One start would be to use the themes emerging from this study to create a thematic *checklist* that can then be compared against the item content of available instruments [45], while evaluating the adequacy of the published evidence for their content validity [13,35]. To assist this process, the in-depth executive summaries and proposed subdomains can be taken alongside the plain language definitions. We further suggest that this approach could be applied to outcome domains where the Web discussion forum did not reach an agreement on the concept definition (such as *tinnitus intrusiveness*) and outcome domains that are more typically measured using performance-based tests rather than questionnaires (such as *concentration*).

Conclusions

Our experience leads us to strongly advocate the use of qualitative methods to ensure concepts are defined to support clear and consistent interpretation by all end users and agreed upon before looking to map outcome domains to measurement instruments. The vast range of different interpretations held for the same domains became apparent during the study, and some major decisions were made as to how the core outcome domains should be conceptualized, defined, and distinguished going forward. Any COS development study following the recommendations of COMET [5] should place substantial emphasis on patient and public involvement. This necessitates involving stakeholders in detailed concept definition as it cannot be assumed that those who contribute to the consensus decision are all speaking the same language based upon research literature and professional terminology. Our findings support the acceptability and feasibility of using Web discussion forums as a research method to achieve this.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Supplementary analysis of consensus meeting discussion on the 4 core outcome domains that were not part of the Web discussion forum.
[DOCX File , 24 KB - ijmr_v9i1e14446_app1.docx]

Multimedia Appendix 2
Instructional Video 1. Hosted on the homepage, walking participants through how to register and create an account.
[MP4 File (MP4 Video), 18165 KB - ijmr_v9i1e14446_app2.mp4]

Multimedia Appendix 3
Instructional Video 2. Hosted on the homepage once logged in, walking participants through how to post, comment on others posts, and adjust their notification settings.
[MP4 File (MP4 Video), 28931 KB - ijmr_v9i1e14446_app3.mp4]

Multimedia Appendix 4
Moderator's semistructured plans for each discussion thread, including general template of prompts, questions, and posts and tailored discussion packs for each core outcome domain.
[DOCX File , 27 KB - ijmr_v9i1e14446_app4.docx]

Multimedia Appendix 5
Executive summaries and revised plain language definitions for each core outcome domain, including voting results for each discussion thread.
[DOCX File , 21 KB - ijmr_v9i1e14446_app5.docx]

References

1. Hall DA, Fackrell K, Li AB, Thavayogan R, Smith S, Kennedy V, et al. A narrative synthesis of research evidence for tinnitus-related complaints as reported by patients and their significant others. Health Qual Life Outcomes 2018 Apr 11;16(1):61 [FREE Full text] [doi: 10.1186/s12955-018-0888-9] [Medline: 29642913]

2. Hall DA, Haider H, Szczepek AJ, Lau P, Rabau S, Jones-Diette J, et al. Systematic review of outcome domains and instruments used in clinical trials of tinnitus treatments in adults. Trials 2016 Jun 1;17(1):270 [FREE Full text] [doi: 10.1186/s13063-016-1399-9] [Medline: 27250987]

3. Clarke M, Williamson PR. Core outcome sets and systematic reviews. Syst Rev 2016 Jan 20;5:11 [FREE Full text] [doi: 10.1186/s13643-016-0188-6] [Medline: 26792080]

4. Gargon E, Williamson PR, Altman DG, Blazeby JM, Clarke M. The COMET Initiative database: progress and activities from 2011 to 2013. Trials 2014 Jul 10;15:279 [FREE Full text] [doi: 10.1186/1745-6215-15-279] [Medline: 25012001]

5. Williamson PR, Altman DG, Bagley H, Barnes KL, Blazeby JM, Brookes ST, et al. The COMET Handbook: version 1.0. Trials 2017 Jun 20;18(Suppl 3):280 [FREE Full text] [doi: 10.1186/s13063-017-1978-4] [Medline: 28681707]

6. Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a 'Core Outcome Set' - a practical guideline. Trials 2016 Sep 13;17(1):449 [FREE Full text] [doi: 10.1186/s13063-016-1555-2] [Medline: 27618914]

7. Hall DA, Haider H, Kikidis D, Mielczarek M, Mazurek B, Szczepek AJ, et al. Toward a global consensus on outcome measures for clinical trials in tinnitus: report from the first international meeting of the COMiT initiative, November 14, 2014, Amsterdam, the Netherlands. Trends Hear 2015 Apr 24;19 [FREE Full text] [doi: 10.1177/2331216515580272] [Medline: 25910505]

8. Fackrell K, Smith H, Colley V, Thacker B, Horobin A, Haider HF, et al. Core Outcome Domains for early phase clinical trials of sound-, psychology-, and pharmacology-based interventions to manage chronic subjective tinnitus in adults: the COMIT'ID study protocol for using a Delphi process and face-to-face meetings to establish consensus. Trials 2017 Aug 23;18(1):388 [FREE Full text] [doi: 10.1186/s13063-017-2123-0] [Medline: 28835261]

9. Smith H, Horobin A, Fackrell K, Colley V, Thacker B, Hall DA, Core Outcome Measures in Tinnitus (COMiT) initiative. Defining and evaluating novel procedures for involving patients in Core Outcome Set research: creating a meaningful long list of candidate outcome domains. Res Involv Engagem 2018;4:8 [FREE Full text] [doi: 10.1186/s40900-018-0091-5] [Medline: 29507772]

10. Hall DA, Smith H, Heffernan E, Fackrell K, Core Outcome Measures in Tinnitus International Delphi (COMiT'ID) Research Steering Group. Recruiting and retaining participants in e-Delphi surveys for core outcome set development: Evaluating the COMiT'ID study. PLoS One 2018;13(7):e0201378 [FREE Full text] [doi: 10.1371/journal.pone.0201378] [Medline: 30059560]

11. Hall DA, Smith H, Hibbert A, Colley V, Haider HF, Horobin A, Core Outcome Measures in Tinnitus (COMiT) initiative. The COMiT'ID Study: developing core outcome domains sets for clinical trials of sound-, psychology-, and pharmacology-based interventions for chronic subjective tinnitus in adults. Trends Hear 2018;22:2331216518814384 [FREE Full text] [doi: 10.1177/2331216518814384] [Medline: 30488765]

12. Hall DA, Hibbert A, Smith H, Haider HF, Londero A, Mazurek B, Core Outcome Measures in Tinnitus (COMiT) initiative. One size does not fit all: developing common standards for outcomes in early-phase clinical trials of sound-, psychology-, and pharmacology-based interventions for chronic subjective tinnitus in adults. Trends Hear 2019;23:2331216518824827 [FREE Full text] [doi: 10.1177/2331216518824827] [Medline: 30803389]

13. Mokkink LB, Prinsen CA, Bouter LM, Vet HC, Terwee CB. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. Braz J Phys Ther 2016 Jan 19;20(2):105-113 [FREE Full text] [doi: 10.1590/bjpt-rbf.2014.0143] [Medline: 26786084]

14. Hallam RS. Manual of the Tinnitus Questionnaire. Revised and updated. London: Polprera Press; 2008.

15. Meikle MB, Henry JA, Griest SE, Stewart BJ, Abrams HB, McArdle R, et al. The tinnitus functional index: development of a new clinical measure for chronic, intrusive tinnitus. Ear Hear 2012;33(2):153-176. [doi: 10.1097/AUD.0b013e31822f67c0] [Medline: 22156949]

16. Stucki G, Cieza A. IMMPACT. HRQL Measures in Chronic Pain Trials. URL: http://www.immpact.org/static/meetings/Immpact2/background/Stucki_review.pdf [accessed 2019-11-17] [WebCite Cache ID 778ePaI9Y]

17. Boers M, Kirwan JR, Wells G, Beaton D, Gossec L, d'Agostino M, et al. Developing core outcome measurement sets for clinical trials: OMERACT filter 2.0. J Clin Epidemiol 2014 Jul;67(7):745-753 [FREE Full text] [doi: 10.1016/j.jclinepi.2013.11.013] [Medline: 24582946]

18. Langan SM, Thomas KS, Williams HC. What is meant by a 'flare' in atopic dermatitis? A systematic review and proposal. Arch Dermatol 2006 Sep;142(9):1190-1196. [doi: 10.1001/archderm.142.9.1190] [Medline: 16983006]

19. Langan SM, Schmitt J, Williams HC, Smith S, Thomas KS. How are eczema 'flares' defined? A systematic review and recommendation for future studies. Br J Dermatol 2014 Mar;170(3):548-556. [doi: 10.1111/bjd.12747] [Medline: 24266741]

20. Thomas KS, Stuart B, O'Leary CJ, Schmitt J, Paul C, Williams HC, et al. Validation of treatment escalation as a definition of atopic eczema flares. PLoS One 2015;10(4):e0124770 [FREE Full text] [doi: 10.1371/journal.pone.0124770] [Medline: 25897763]

21. Tour SK, Thomas KS, Walker D, Leighton P, Yong AS, Batchelor JM. Survey and online discussion groups to develop a patient-rated outcome measure on acceptability of treatment response in vitiligo. BMC Dermatol 2014 Jun 14;14:10 [FREE Full text] [doi: 10.1186/1471-5945-14-10] [Medline: 24929563]

22. Howells LM, Chalmers JR, Cowdell F, Ratib S, Santer M, Thomas KS. 'When it goes back to my normal I suppose': a qualitative study using online focus groups to explore perceptions of 'control' among people with eczema and parents of children with eczema in the UK. BMJ Open 2017 Nov 15;7(11):e017731 [FREE Full text] [doi: 10.1136/bmjopen-2017-017731] [Medline: 29146642]

23. Howells L, Thomas KS, Sears AV, Nasr I, Wollenberg A, Schuttelaar ML, long-term control of eczema working group for the HOME initiative. Defining and measuring 'eczema control': an international qualitative study to explore the views of those living with and treating atopic eczema. J Eur Acad Dermatol Venereol 2019 Jun;33(6):1124-1132 [FREE Full text] [doi: 10.1111/jdv.15475] [Medline: 30720903]

24. Dyson MP, Shave K, Gates A, Fernandes RM, Scott SD, Hartling L. Which outcomes are important to patients and families who have experienced paediatric acute respiratory illness? Findings from a mixed methods sequential exploratory study. BMJ Open 2017 Dec 14;7(12):e018199 [FREE Full text] [doi: 10.1136/bmjopen-2017-018199] [Medline: 29247097]

25. Probst T, Pryss RC, Langguth B, Spiliopoulou M, Landgrebe M, Vesala M, et al. Outpatient tinnitus clinic, self-help web platform, or mobile application to recruit tinnitus study samples? Front Aging Neurosci 2017;9:113 [FREE Full text] [doi: 10.3389/fnagi.2017.00113] [Medline: 28484389]

26. Simoes J, Neff P, Schoisswohl S, Bulla J, Schecklmann M, Harrison S, et al. Toward personalized tinnitus treatment: an exploratory study based on internet crowdsensing. Front Public Health 2019;7:157 [FREE Full text] [doi: 10.3389/fpubh.2019.00157] [Medline: 31294010]

27. Im E, Chee W. Practical guidelines for qualitative research using online forums. Comput Inform Nurs 2012 Nov;30(11):604-611 [FREE Full text] [doi: 10.1097/NXN.0b013e318266cade] [Medline: 22918135]

28. Twitter. COMIT'ID Study. URL: https://twitter.com/COMITIDStudy [accessed 2019-11-17]

29. COMIT'ID. URL: http://www.comitid.com/ [accessed 2019-11-17]

30. Tinnitus Hub. URL: https://www.tinnitushub.com/ [accessed 2019-11-17]

31. Tinnitus Talk Support Forum. URL: http://www.tinnitustalk.com/ [accessed 2019-11-17]

32. McCormack A, Edmondson-Jones M, Somerset S, Hall D. A systematic review of the reporting of tinnitus prevalence and severity. Hear Res 2016 Jul;337:70-79 [FREE Full text] [doi: 10.1016/j.heares.2016.05.009] [Medline: 27246985]

33. Githens RP. Older adults and e-learning: opportunities and barriers. Q Rev Dist Educ 2007;8(4):329-338 [FREE Full text]

34. Braun V, Clarke V. Chapter 4: Thematic analysis. In: Cooper H, editor. APA Handbook of Research Methods in Psychology: Vol. 2. Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological. Washington, DC: American Psychological Association; 2012:57-71.

35. Terwee CB, Prinsen CA, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res 2018 May;27(5):1159-1170 [FREE Full text] [doi: 10.1007/s11136-018-1829-0] [Medline: 29550964]

36. Pryce H, Hall A, Shaw R, Culhane B, Swift S, Straus J, et al. Patient preferences in tinnitus outcomes and treatments: a qualitative study. Int J Audiol 2018 Oct;57(10):784-790. [doi: 10.1080/14992027.2018.1484184] [Medline: 30388941]

37. Ferrante JM, Friedman A, Shaw EK, Howard J, Cohen DJ, Shahidi L. Lessons learned designing and using an online discussion forum for care coordinators in primary care. Qual Health Res 2016 Nov;26(13):1851-1861 [FREE Full text] [doi: 10.1177/1049732315609567] [Medline: 26481942]

38. Horgan A, McCarthy G, Sweeney J. An evaluation of an online peer support forum for university students with depressive symptoms. Arch Psychiatr Nurs 2013 Apr;27(2):84-89. [doi: 10.1016/j.apnu.2012.12.005] [Medline: 23540518]

39. Im E, Lee B, Hwang H, Yoo KH, Chee W, Stuifbergen A, et al. 'A waste of time': Hispanic women's attitudes toward physical activity. Women Health 2010 Sep;50(6):563-579 [FREE Full text] [doi: 10.1080/03630242.2010.510387] [Medline: 20981637]

40. Im E, Lee SH, Chee W. 'Being conditioned, yet becoming strong': Asian American women in menopausal transition. J Transcult Nurs 2011 Jul;22(3):290-299. [doi: 10.1177/1043659611404429] [Medline: 21519062]

41. Benford P, Walker DM. The internet as a research medium. In: Walker DM, editor. An Introduction to Health Services Research: A Practical Guide. London: Sage; 2014.

42. Welch V, Petkovic J, Pardo Pardo J, Rader T, Tugwell P. Interactive social media interventions to promote health equity: an overview of reviews. Health Promot Chronic Dis Prev Can 2016 Apr;36(4):63-75 [FREE Full text] [doi: 10.24095/hpcdp.36.4.01] [Medline: 27077792]

43. Ando H, Cousins R, Young C. Achieving saturation in thematic analysis: development and refinement of a codebook. Compr Psychol 2014;3(4). [doi: 10.2466/03.cp.3.4]

44. Saunders B, Sim J, Kingstone T, Baker S, Waterfield J, Bartlam B, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. Qual Quant 2018;52(4):1893-1907 [FREE Full text] [doi: 10.1007/s11135-017-0574-8] [Medline: 29937585]

45. Heinl D, Prinsen CA, Deckert S, Chalmers JR, Drucker AM, Ofenloch R, et al. Measurement properties of adult quality-of-life measurement instruments for eczema: a systematic review. Allergy 2016 Mar;71(3):358-370. [doi: 10.1111/all.12806] [Medline: 26564008]

Abbreviations

COMET: Core Outcome Measures in Effectiveness Trials
COMiT: Core Outcomes Measures in Tinnitus
COMiT'ID: Core Outcomes Measures in Tinnitus: International Delphi
COS: core outcome set
NIHR: National Institute for Health Research

XSL·FO
RenderX

Original Paper

Perceived Treatment Satisfaction and Effectiveness Facilitators Among Patients With Chronic Health Conditions: A Self-Reported Survey

Kathryn Volpicelli Leonard[1*], MPH; Courtney Robertson[1*], PhD; Amrita Bhowmick[1,2*], MPH, MBA; Leslie Beth Herbert[1*], PhD

[1]Health Union LLC, Philadelphia, PA, United States

[2]Department of Health Behavior, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, NC, United States

[*]all authors contributed equally

Corresponding Author:
Kathryn Volpicelli Leonard, MPH
Health Union LLC
1 International Plaza Dr
#550
Philadelphia, PA,
United States
Phone: 1 4849859715
Email: Kate.Leonard@health-union.com

Abstract

Background: Approximately 50% of patients are nonadherent to prescribed medications. Patient perception regarding medication effectiveness has been linked to improved adherence. However, how patients perceive effectiveness is poorly understood.

Objective: The aim of this study was to elucidate factors associated with perceived treatment satisfaction and effectiveness among patients with chronic health conditions.

Methods: We conducted a descriptive study using a cross-sectional survey design. We administered a Web-based survey to participants with migraine, multiple sclerosis (MS), or rheumatoid arthritis (RA). Patients were recruited from established online communities of Health Union. Descriptive statistics, correlations, and comparison tests were used to examine outcomes.

Results: Data were collected from 1820 patients: 567 with migraine, 717 with MS, and 536 with RA. The majority of participants were female (1644/1820, 90.33%), >40 years old (1462/1820, 80.33%), and diagnosed >5 years ago (1189/1820, 65.33%). Treatment satisfaction and perceived medication effectiveness were highly correlated ($r=0.90$, $P<.01$). Overall, three temporal factors were positively correlated with satisfaction or perceived effectiveness: time on current medication (satisfaction $r_s=0.22$, $P<.01$; effectiveness $r_s=0.25$, $P<.01$), time since diagnosis (satisfaction $r_s=0.07$, $P<.01$; effectiveness $r_s=0.09$, $P<.01$), and time on treatment (effectiveness $r_s=0.08$, $P<.01$).

Conclusions: Findings validated the strong relationship between treatment satisfaction and perceived effectiveness. Understanding the (1) positive relationship between time and treatment satisfaction and effectiveness and (2) factors associated with determining medication effectiveness can help clinicians better understand the mindset of patients regarding treatment. Clinicians may be better prepared to elicit patient beliefs, which influence medication adherence, for people diagnosed with chronic health conditions.

(Interact J Med Res 2020;9(1):e13029)   doi:10.2196/13029

KEYWORDS

treatment effectiveness; patient satisfaction; migraine; multiple sclerosis; rheumatoid arthritis

XSL•FO
RenderX

Introduction

Background

The World Health Organization reported that in developed countries, approximately 50% of patients with chronic health conditions do not adhere to the medication they have been prescribed [1]. Adherence and compliance to treatment are important in any disease context but can be particularly challenging in the context of chronic health conditions that require sustained adherence, even in the absence of acute symptoms and with regimens that can be logistically and economically challenging [1,2]. In turn, low levels of adherence and compliance can have a dramatic detrimental impact on symptomology, overall disease course, and health care costs [1-3]. For patients with chronic conditions, noncompliance can mean reduced quality of life and swifter disease progression; for the health care system as a whole, nonadherence increases the societal cost burden (eg, for avoidable health care professional [HCP] visits and hospitalizations) [3-7]. Adherence and compliance to treatment are, therefore, widely researched topics, with a complex set of predictors summarized by Jin et al [5] into categories related to patient-determined factors (demographics, beliefs, motivations, etc), treatment logistics, social and economic factors, health care availability and accessibility, and disease experiences.

Several studies have looked closely at one particular component of that matrix: patient beliefs about therapy. For example, Rajpura and Nayak [8] reported that positive beliefs regarding medication predicted adherence to medication among elders with hypertension. Patient satisfaction with treatment (in addition to the more general construct of patient satisfaction with care) has also been associated with better adherence and compliance, including research done among patients seeking chronic pain treatment, patients with type 2 diabetes, patients with chronic obstructive pulmonary disease, patients with cystic fibrosis, patients with depression, and patients with hypertension [9-14].

In addition, research has supported a role for treatment efficacy perceptions in predicting better adherence and compliance. Bender and Bender [15] found that, among patients with asthma, concerns about diminishing treatment effectiveness over time played a secondary role in adherence behaviors, behind more frequently mentioned factors such as safety, cost, and perceived disease severity. Although not treatment efficacy *per se*, Horne and Weinman [16] also found that patients who believed that their prescribed medication was necessary for maintaining health also reported higher compliance.

Research has also shown a strong connection between treatment satisfaction and perceived treatment effectiveness. For example, the Treatment Satisfaction Questionnaire for Medication (TSQM) incorporates an element of perceived treatment effectiveness as 1 of the 4 domains in determining treatment satisfaction. Specifically, the TSQM includes questions related to the impact of the medication on disease and symptoms that, along with all other dimensions, have been shown to be highly reliable and valid constructs. Relevant to the current research questions, the TSQM was also originally validated among patients with chronic conditions (arthritis, asthma, depression, type 1 diabetes, hypercholesterolemia, hypertension, migraine, and psoriasis) [17], and was found to be a useful tool for measuring treatment satisfaction among multiple sclerosis (MS) patients [18].

Objectives

In summary, several studies have demonstrated a relationship between treatment satisfaction and adherence and compliance, as well as between perceived treatment effectiveness and adherence and compliance. Moreover, a strong link has been established between treatment satisfaction and perceived treatment effectiveness.

However, although some research has demonstrated the importance patients place on efficacy over tolerability and ease of administration [19], less research has focused on the basis of patient perceptions of treatment effectiveness.

The aim of this study was, therefore, to better understand the factors that patients rely on when making personal evaluations of treatment effectiveness, including the role of quality of life improvements, symptoms, and HCP assessment. Secondarily, the research further explored the relationship between perceived treatment effectiveness and treatment satisfaction among patients with chronic conditions (Figure 1). The analyses focused on patients with 1 of 3 chronic conditions (migraine, MS, or rheumatoid arthritis [RA]).

Figure 1. Relationships explored in research on how patients determine if a treatment is working for them. HCP: health care practitioner; MS: multiple sclerosis; RA: rheumatoid arthritis.
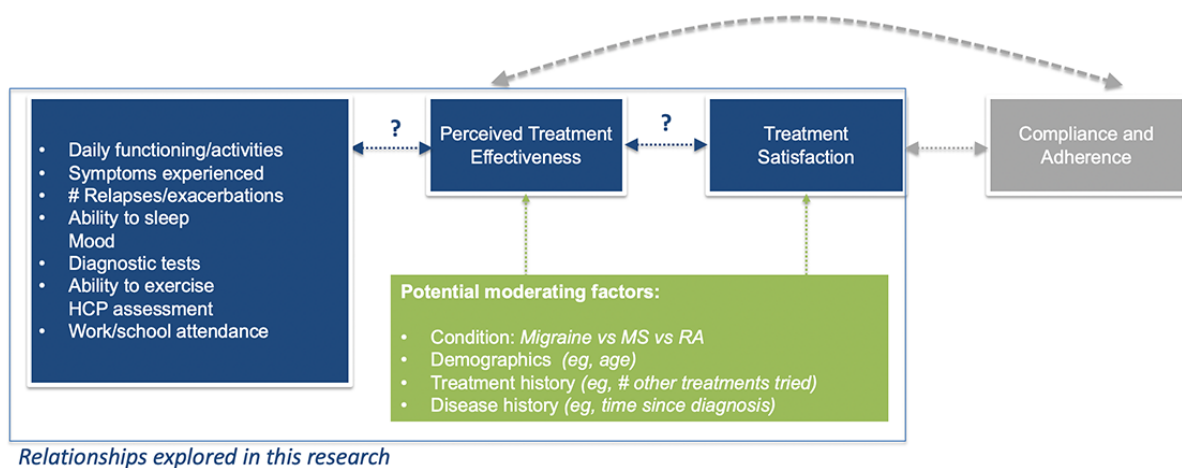
Methods

Recruitment

We conducted a descriptive study using a cross-sectional survey design. A Web-based survey was administered from June 17, 2017, to July 30, 2017, among 3 online communities of Health Union (Philadelphia, PA, The United States). Potential participants were recruited as a convenience sample. We employed this common nonprobability sampling technique given its efficiency, ease of implementation, low cost, and the exploratory nature of the questions posed herein.

Links to the survey were posted on the 3 Health Union community sites and associated Facebook pages: Migraine.com, MultipleSclerosis.net, and RheumatoidArthritis.net. We chose to survey these specific patient communities as these patients are apt to have chronic conditions that often require >1 medication to treat, as well as the potential for patients to need to try several medications before finding the one that works for them. Participants were eligible to participate in the survey if they were at least 18 years old, lived in the United States, and previously received a diagnosis from an HCP of the respective health condition of the community (ie, migraine, MS, and RA).

The research was conducted in accordance with all applicable regulations as outlined in the Declaration of Helsinki and was exempt from institutional review. Informed consent was obtained from all participants before completing the Web-based survey. Participants were informed about the voluntary nature of the survey, information being collected, anonymous nature of data collection, and the expected time for survey completion. In addition, no identifiers were collected from participants; however, Internet Protocol (IP) addresses of participants were used to ensure that the questionnaire could only be completed once by any individual. Duplicate entries were avoided by preventing users with the same IP address from accessing the survey more than once during the study period.

Data Collection

Electronic data collection was conducted through SurveyMonkey (SMVK Inc, California), with data protection provided through its security measures. After reviewing information about the nature of the survey, participants completed a survey consisting of up to 30 multiple-choice questions and an optional free-response item. Responses were completely anonymous, and no compensation was provided for participation.

Measures

The survey included questions on demographics, treatment journey, and treatment satisfaction and effectiveness. Targeted questions were asked based on where participants were in their journey, that is, having never taken medication, taken medication in the past but stopped, and currently taking medication. In addition, questions were asked to gain a deeper understanding of the number of medications tried and what prompted switching medications. Satisfaction was measured on a 7-point scale, ranging from 1=Not at all satisfied to 7=Extremely satisfied. Perceived treatment effectiveness was also measured on a 7-point scale, ranging from 1=Not at all effective to 7=Extremely effective.

After completing the multiple-choice questions, participants were given the option of completing an open-ended response to share additional observations or concerns about treating their condition. Participants were instructed that *medication* referred to both over-the-counter and prescription medications throughout the survey.

Statistical Analysis

Several of the collected responses to demographic and other types of questions were utilized to stratify participants for further comparison: age (<40, 40-49, 50-59, and >60), time since diagnosis calculated based on current age and age at diagnosis (<2 years, 2-5 years, and >5 years), number of medications ever taken for health condition (< median for condition and > median for condition), and frequency of HCP visits (once a month or more, every 2-3 months, twice a year, and once a year or as needed). Demographic categories were reported using descriptive statistics, and comparisons between categories were evaluated using chi-squared analyses. Differences in mean Likert scores among categories were evaluated using analysis of

variance. Correlation between Likert scores and ordinal variables was conducted using the Spearman rank-order correlation, whereas correlation between Likert scores was conducted using the Pearson correlation. Data analysis used $P<.05$ to determine statistical significance.

Results

The survey was completed by 1820 participants—567, 717, and 536 diagnosed with migraine, MS, and RA, respectively (Table 1). For purposes of this research, data of participants currently using a medication for their condition (N=1641) were analyzed (migraine, n=524; MS, n=617; RA, n=500).

Table 1. Demographics of survey respondents currently using a medication for their condition.

Demographic factors	Migraine (n=524), n (%)	MS[a] (n=617), n (%)	RA[b] (n=500), n (%)	All (N=1641), n (%)
Age				
<40	124 (23.7)	125 (20.3)	81 (16.2)	330 (20.11)
40-49	159 (30.3)	141 (22.9)	110 (22.0)	410 (24.98)
50-59	165 (31.5)	233 (37.8)	157 (31.4)	555 (33.82)
60+	76 (14.5)	118 (19.1)	152 (30.4)	346 (21.08)
Sex				
Female	501 (95.6)	513 (83.1)	478 (95.6)	1492 (90.92)
Male	23 (4.4)	104 (16.9)	22 (4.4)	149 (9.08)
Number of medications				
< Median[c]	290 (55.3)	356 (57.7)	261 (52.2)	907 (55.27)
> Median[c]	234 (44.7)	261 (42.3)	239 (47.8)	734 (44.73)
Frequency of health care professional visits				
Once a month or more	117 (22.3)	34 (5.5)	86 (17.2)	237 (14.44)
Every 2-3 months	229 (43.7)	225 (36.5)	355 (71.0)	809 (49.30)
Twice a year	74 (14.1)	271 (43.9)	52 (10.4)	397 (24.19)
Once a year, when I relapse/need to	104 (19.8)	87 (14.1)	7 (1.4)	198 (12.07)

[a]MS: multiple sclerosis.

[b]RA: rheumatoid arthritis.

[c]The median number of medications for patients with migraine was 7, for patients with multiple sclerosis was 2, and for patients with rheumatoid arthritis was 4; across all respondents the median number was 4.

Medication Satisfaction and Perceived Treatment Effectiveness

In examining satisfaction with a participant's current medication, the average rating across all participants was 4.6. When examining satisfaction by condition, participants significantly differed among each of the conditions (migraine=4.2, MS=5.1, and RA=4.4; $F_{2,1638}=62.97$, $P<.01$). Similarly, when exploring perceived effectiveness with current medication, the average rating was 4.6 across all participants. A significant difference was also seen between each of the conditions in participants' perceived effectiveness ratings (migraine=4.4, MS=5.0, and RA=4.3; $F_{2,1638}=41.23$, $P<.01$). Given the parallels in the satisfaction and perceived efficacy ratings, correlation was conducted between these ratings to verify the relationship. As expected, both within and across all conditions, ratings of current medication satisfaction and current medication perceived efficacy were strongly correlated (all participants, $r_{1639}=0.90$, $P<.01$; migraine, $r_{522}=0.91$, $P<.01$; MS, $r_{615}=0.85$, $P<.01$; RA, $r_{498}=0.93$, $P<.01$).

Relationship Between Time and Treatment Satisfaction and Perceived Effectiveness

Time on treatment for a participant's condition and time on current medication were examined to identify if these factors are related to treatment satisfaction and perceived efficacy. It was found that as the duration that the participant was taking medication for the condition increased, so did the participant's satisfaction with and perceived efficacy of the current medication. Satisfaction was significantly correlated with time on treatment for the participant's condition within each of the three surveyed conditions, but not across the entire sample (all participants, $r_{1639}=0.04$, $P=.09$; migraine, $r_{522}=0.10$, $P=.03$; MS, $r_{615}=0.08$, $P=.047$; RA, $r_{498}=0.17$, $P<.01$; Table 2). Similarly, perceived treatment efficacy was significantly correlated with time on treatment within each of the three surveyed conditions, as well as across the entire sample (all participants, $r_{s1639}=0.08$, $P<.01$; migraine, $r_{s522}=0.09$, $P=.04$; MS, $r_{s615}=0.11$, $P=.01$; RA, $r_{s498}=0.17$, $P<.01$.

XSL•FO
RenderX

Table 2. Medication satisfaction and perceived effectiveness by time on treatment.

Medication ratings	Migraine	MS[a]	RA[b]	All
Mean satisfaction ratings by time on treatment				
Less than 1 year	4.1	4.7	4.0	4.4
1 to 4 years	3.9	5.2	4.2	4.6
5 to 10 years	4.2	5.2	4.5	4.7
More than 10 years	4.3	5.2	4.7	4.6
Mean perceived effectiveness ratings by time on treatment				
Less than 1 year	4.2	4.6	4.0	4.3
1 to 4 years	4.0	5.1	4.2	4.5
5 to 10 years	4.4	5.2	4.5	4.7
More than 10 years	4.5	5.2	4.8	4.7

[a]MS: multiple sclerosis.

[b]RA: rheumatoid arthritis.

Within each of the three conditions, as well as across the entire sample, time on current medication was significantly correlated with both satisfaction (all participants, $r_{s1639}=0.22$, $P<.01$; migraine, $r_{s522}=0.24$, $P<.01$; MS, $r_{s615}=0.14$, $P<.01$; RA, $r_{s498}=0.29$, $P<.01$) and perceived efficacy of the participant's current medication (all participants, $r_{s1639}=0.25$, $P<.01$; migraine, $r_{s522}=0.26$, $P<.01$; MS, $r_{s615}=0.21$, $P<.01$; RA, $r_{s498}=0.28$, $P<.01$) (Table 3).

Table 3. Medication satisfaction and perceived effectiveness by time on current medication.

Medication ratings	Migraine	MS[a]	RA[b]	All
Mean rating of satisfaction by time on current medication				
Less than 3 months	3.5	4.8	3.6	3.8
3 to 6 months	4.0	4.8	4.1	4.3
6 months to 1 year	4.0	5.0	4.7	4.6
1 to 2 years	4.2	5.2	4.3	4.8
More than 2 years	4.5	5.3	4.9	4.9
Mean rating of perceived effectiveness by time on current medication				
Less than 3 months	3.6	4.6	3.6	3.8
3 to 6 months	4.1	4.6	4.1	4.3
6 months to 1 year	4.1	4.8	4.7	4.6
1 to 2 years	4.2	5.0	4.3	4.7
More than 2 years	4.7	5.3	4.8	5.0

[a]MS: multiple sclerosis.

[b]RA: rheumatoid arthritis.

Relationship Between Medication Experience and Satisfaction and Perceived Effectiveness

The median number of medications ever taken for each participant's health condition was calculated per condition: all conditions = 4, migraine=7, MS=2, and RA=4. From these medians, participants were split into groups by number taken: Up through the median for each condition and over the median for each condition. Satisfaction and perceived effectiveness ratings were then correlated with these categories to find that there is a negative relationship with those who have tried more medications being less satisfied (Table 4). Significant negative correlations were found within migraine and MS participants, as well as across the entire sample for both satisfaction (all participants, $r_{s1639}=-0.09$, $P<.01$; migraine, $r_{s522}=-0.12$, $P<.01$; MS, $r_{s615}=-0.10$, $P=.01$; RA, $r_{s498}=-0.02$, $P=.59$) and perceived effectiveness (all participants, $r_{s1639}=-0.12$, $P<.01$; migraine, $r_{s522}=-0.15$, $P<.01$; MS, $r_{s615}=-0.15$, $P<.01$; $r_{s498}=-0.03$, $P=.50$).

Table 4. Medication satisfaction and perceived effectiveness by medication experience.

Mean rating	Migraine	MS[a]	RA[b]	All
Mean rating of satisfaction by medication experience				
Up through median	4.4	5.3	4.4	4.7
Over median	4.0	5.0	4.3	4.5
Mean rating of perceived effectiveness by medication experience				
Up through median	4.6	5.2	4.4	4.8
Over median	4.1	4.8	4.3	4.4

[a]MS: multiple sclerosis.

[b]RA: rheumatoid arthritis.

Demographics in Relation to Medication Satisfaction and Perceived Effectiveness

In addition to exploring treatment experience and time in relation to satisfaction and perceived effectiveness, the participant's age and frequency of seeing an HCP were also examined. As age increased, so did the participant's satisfaction (all participants, $r_{s1639}=0.09$, $P<.01$; migraine, $r_{s522}=0.13$, $P<.01$; MS, $r_{s615}=0.04$, $P=.30$; RA, $r_{s498}=-0.12$, $P<.01$) and perceived effectiveness (all participants, $r_{s1639}=0.11$, $P<.01$; migraine, $r_{s522}=0.17$, $P<.01$; MS, $r_{s615}=0.07$, $P=.11$; RA, $r_{s498}=0.13$, $P<.01$) with the current medication. These correlations are significant for the entire sample, as well as for migraine and RA participants, but not for MS participants (Table 5).

Table 5. Medication satisfaction and perceived effectiveness by participant age.

Mean rating	Migraine	MS[a]	RA[b]	All
Mean rating of satisfaction by participation age (years)				
<40	3.9	5.1	4.0	4.4
40-49	4.2	5.1	4.5	4.6
50-59	4.2	5.2	4.2	4.6
60+	4.7	5.3	4.7	4.0
Mean rating of perceived effectiveness by participation age (years)				
<40	4.1	4.8	4.0	4.3
40-49	4.3	5.1	4.5	4.6
50-59	4.4	5.1	4.1	4.6
60+	4.9	5.2	4.7	4.9

[a]MS: multiple sclerosis.

[b]RA: rheumatoid arthritis.

A negative relationship was found between the frequency of HCP visits with both satisfaction and perceived effectiveness, indicating that with fewer HCP visits come higher contentment and perceived value (Table 6). A significant negative correlation was found for satisfaction within migraine and RA participants, as well as the entire sample (all participants, $r_{s1639}=-0.22$, $P<.01$; migraine, $r_{s522}=-0.23$, $P<.01$; MS, $r_{s615}=-0.08$, $P=.05$; RA, $r_{s498}=-0.14$, $P<.01$); perceived efficacy was significant within all conditions and across the entire sample (all participants, $r_{s1639}=-0.24$, $P<.01$; migraine, $r_{s522}=-0.28$, $P<.01$; MS, $r_{s615}=-0.10$, $P=.01$; RA, $r_{s498}=-0.18$, $P<.01$).

XSL•FO
RenderX

Table 6. Medication satisfaction and perceived effectiveness by frequency of health care professional visits.

Mean ratings	Migraine	MS[a]	RA[b]	All
Mean rating of satisfaction by frequency of HCP[c] visits				
Once a month or more	3.7	4.9	4.0	4.0
Every 2-3 months	4.2	5.0	4.4	4.5
Twice a year	4.6	5.2	4.8	5.1
Once a year, when I relapse/need to	4.7	5.2	4.9	4.9
Mean rating of perceived effectiveness by frequency of HCP visits				
Once a month or more	3.7	4.8	3.9	4.0
Every 2-3 months	4.3	4.9	4.4	4.5
Twice a year	4.9	5.2	4.9	5.1
Once a year, when I relapse/need to	4.9	5.2	4.9	5.0

[a]MS: multiple sclerosis.

[b]RA: rheumatoid arthritis.

[c]HCP: health care professional.

Determining Medication Efficacy

When asked to select from a list of factors that helped participants to determine how well their current medication is working, participants across all three conditions surveyed were most likely to select their ability to perform day-to-day activities and the number of symptoms experienced (Table 7).

Table 7. Top factors selected by participants for determining medication effectiveness (N=1641).

Factors for determining medication effectiveness	X^2 (2)	P value	Migraine, n (%)	MS[a], n (%)	RA[b], n (%)	All, n (%)
My ability to perform my day-to-day activities	187.4	<.01	462 (88.2)	384 (62.2)	461 (92.2)	1307 (79.65)
The number of symptoms I experience	74.1	<.01	358 (68.3)	339 (54.9)	396 (79.2)	1093 (66.61)
The number of relapses/exacerbations I experience	46.9	<.01	319 (60.9)	393 (63.7)	222 (44.4)	934 (56.92)
My ability to sleep	118.1	<.01	221 (42.2)	112 (18.2)	232 (46.4)	565 (34.43)
My mood or level of happiness/depression	62.9	<.01	205 (39.1)	147 (23.8)	168 (33.6)	520 (31.69)
My ability to exercise	100.5	<.01	118 (22.5)	143 (23.2)	238 (47.6)	499 (30.41)

[a]MS: multiple sclerosis.

[b]RA: rheumatoid arthritis.

Discussion

Principal Findings

Further validating the relationship between medication satisfaction and perceived treatment effectiveness, large effect sizes (ranging from 0.85 to 0.93) were observed across conditions between ratings of medication satisfaction and perceived treatment effectiveness. Interestingly, however, mean treatment satisfaction and treatment effectiveness ratings were higher among patients with MS than patients with migraine or RA, which may be related to differences in the demographic composition of the subgroups (eg, more male patients were represented in the MS sample compared with the other groups) and/or which may reflect differences in overall condition management (eg, patients with MS report more frequent HCP visits compared with the other groups).

Across all three conditions, a time dimension—regardless of how it was measured (time since diagnosis, time since starting medication for condition, and time on current medication)—also showed a positive relationship with medication satisfaction and perceived treatment effectiveness. More experience with the condition (and with medication for the condition) may lead to differing/more realistic expectations for treatment effectiveness and/or may give patients time to find a treatment that *works*. With that said, the research also demonstrates that patients who cycle through multiple medications and/or who are in more regular contact with their HCP are more likely to report lower perceived treatment effectiveness (the former was only significant for patients with migraine and MS).

Overall, these findings suggest that clinicians' understanding of disease history (time since diagnosis) and treatment history (time since starting on medication for condition, number of over-the-counter and prescription medications tried, and time on current medication) will help inform a perspective on patient mindset regarding treatment effectiveness. Are patients at a point where they have accepted the diagnosis, they understand expectations for treatment effectiveness, and they have found a treatment that works for them? Or, have patients cycled

through multiple medications without success and made frequent visits to discuss next steps?

The research also found that the ability to perform daily activities was, by far, the most likely way in which patients with migraine or RA determined whether a treatment was working for them, whereas patients with MS were almost equally as likely to factor in both ability to perform daily activities and number of relapses/exacerbations. However, it is important that clinicians not forget to inquire about potential other factors that may have an impact on perceived treatment effectiveness. These secondary factors included number of symptoms experienced (across all three conditions), ability to sleep (for patients with migraine and RA, in particular), and ability to exercise (for patients with RA, in particular). It is also interesting to note that patients are less likely to indicate that the assessment of their HCP influenced their determination of treatment effectiveness—instead relying more on their personal experience.

Given the link between perceived treatment effectiveness and treatment adherence, an understanding of the factors that patients use to determine whether a treatment *works* can help clinicians tailor their patient interactions to more specifically discuss these factors. Questions around level of daily functioning and symptoms, as well as sleep and exercise patterns, should be used to help clinicians better understand the treatment experience. As needed, clinicians should be prepared to initiate conversations early if patients point to issues with these aspects to ensure continued adherence if appropriate. There should also be recognition that the clinician's assessment that a treatment is working may not be mirrored by the patient, and that treatment goals may differ between the clinician and patient [20-22]. Recent views expressed by Crum and Zuckerman [23] reinforce the importance of clinician-patient conversations around perceived treatment effectiveness, including identifying the origin of a patient's mindset around treatment and how fixed or malleable that mindset may be. Brown [24] makes a similar point regarding the importance of close and careful listening to patients, in her discussion of *What Patients Say, What Doctors Hear* by Danielle Ofri [25].

Limitations

The study's sample size had an adequate level of statistical power; however, interpretation of these data was limited by design issues inherent with using convenience sampling and self-report data, which is subject to recall and participation bias. Respondents represent those who are engaged with online health communities and may not represent or be generalizable to the broader patient populations of each condition. For example, more women completed the survey than men, which should be noted as a limitation. In addition, patients recruited through these methods may have had increased knowledge about the progression, treatment, and coping strategies of the disease, which may have influenced study results. More research needs to be done to determine if study findings are consistent across patient populations recruited in other ways.

Additional research into the factors underlying perceived treatment effectiveness is needed, incorporating aspects not directly assessed in this research (eg, side effect/tolerability experiences) and/or focusing more explicitly on prescription (rather than over-the-counter) treatments. This research also did not directly address the relationship between perceived treatment effectiveness and compliance, although other research has suggested a strong relationship between the two (eg, for use of mental health services) [26].

Conclusions

Time since diagnosis, time since starting medication for condition, and time on current medication showed a positive relationship with medication satisfaction and perceived treatment effectiveness. In addition, more patient experience with the condition (and with medication for their condition) may lead to more realistic treatment expectations and/or may give patients time to find a treatment that they believe works for them. Conversely, lower perceived treatment effectiveness and multiple medication attempts appear to prompt more frequent HCP visits, likely increasing health care costs and placing potential strain on the HCP-patient relationship.

Given the link between perceived treatment effectiveness and adherence, an understanding of the factors patients used to determine whether a treatment is effective can help clinicians tailor their patient interactions to more specifically discuss these factors. Clinicians should ask more questions around symptoms and daily functioning, as well as perceived treatment effectiveness, to better understand the treatment experience. Additional research into the factors underlying perceived treatment effectiveness is needed, incorporating aspects not directly assessed in this research, such as tolerability.

Conflicts of Interest

The authors of this paper are employees of Health Union, LLC, a company that creates condition-specific online ecosystems to support, educate and connect millions of people with challenging, chronic health concerns.

References

1. World Health Organization. Adherence to Long-term Therapies: Evidence for Action. Geneva: World Health Organization; 2003.

2. Lavsa SM, Holzworth A, Ansani NT. Selection of a validated scale for measuring medication adherence. J Am Pharm Assoc (2003) 2011;51(1):90-94. [doi: 10.1331/JAPhA.2011.09154] [Medline: 21247831]

3. Hughes DA, Bagust A, Haycox A, Walley T. The impact of non-compliance on the cost-effectiveness of pharmaceuticals: a review of the literature. Health Econ 2001 Oct;10(7):601-615. [doi: 10.1002/hec.609] [Medline: 11747044]

4. Scarlett W, Young S. Medical noncompliance: the most ignored national epidemic. J Am Osteopath Assoc 2016 Aug 1;116(8):554-555. [doi: 10.7556/jaoa.2016.111] [Medline: 27455106]

5. Jin J, Sklar GE, Oh VM, Li SC. Factors affecting therapeutic compliance: a review from the patient's perspective. Ther Clin Risk Manag 2008 Feb;4(1):269-286 [FREE Full text] [doi: 10.2147/tcrm.s1458] [Medline: 18728716]

6. Lam WY, Fresco P. Medication adherence measures: an overview. Biomed Res Int 2015;2015:217047 [FREE Full text] [doi: 10.1155/2015/217047] [Medline: 26539470]

7. Svarstad BL, Shireman TI, Sweeney JK. Using drug claims data to assess the relationship of medication adherence with hospitalization and costs. Psychiatr Serv 2001 Jun;52(6):805-811. [doi: 10.1176/appi.ps.52.6.805] [Medline: 11376229]

8. Rajpura J, Nayak R. Medication adherence in a sample of elderly suffering from hypertension: evaluating the influence of illness perceptions, treatment beliefs, and illness burden. J Manag Care Pharm 2014 Jan;20(1):58-65 [FREE Full text] [doi: 10.18553/jmcp.2014.20.1.58] [Medline: 24511766]

9. Hirsh AT, Atchison JW, Berger JJ, Waxenberg LB, Lafayette-Lucey A, Bulcourf BB, et al. Patient satisfaction with treatment for chronic pain: predictors and relationship to compliance. Clin J Pain 2005;21(4):302-310. [doi: 10.1097/01.ajp.0000113057.92184.90] [Medline: 15951647]

10. Guisasola FA, Povedano ST, Krishnarajah G, Lyu R, Mavros P, Yin D. Hypoglycaemic symptoms, treatment satisfaction, adherence and their associations with glycaemic goal in patients with type 2 diabetes mellitus: findings from the Real-Life Effectiveness and Care Patterns of Diabetes Management (RECAP-DM) Study. Diabetes Obes Metab 2008 Jun;10(Suppl 1):25-32. [doi: 10.1111/j.1463-1326.2008.00882.x] [Medline: 18435671]

11. Chrystyn H, Small M, Milligan G, Higgins V, Gil EG, Estruch J. Impact of patients' satisfaction with their inhalers on treatment compliance and health status in COPD. Respir Med 2014 Feb;108(2):358-365 [FREE Full text] [doi: 10.1016/j.rmed.2013.09.021] [Medline: 24209768]

12. Regnault A, Balp M, Kulich K, Viala-Danten M. Validation of the Treatment Satisfaction Questionnaire for Medication in patients with cystic fibrosis. J Cyst Fibros 2012 Dec;11(6):494-501 [FREE Full text] [doi: 10.1016/j.jcf.2012.04.007] [Medline: 22583743]

13. Jneid S, Jabbour H, Hajj A, Sarkis A, Licha H, Hallit S, et al. Quality of life and its association with treatment satisfaction, adherence to medication, and trust in physician among patients with hypertension: a cross-sectional designed study. J Cardiovasc Pharmacol Ther 2018 Nov;23(6):532-542. [doi: 10.1177/1074248418784292] [Medline: 29916266]

14. Aljumah K, Hassali AA, AlQhatani S. Examining the relationship between adherence and satisfaction with antidepressant treatment. Neuropsychiatr Dis Treat 2014;10:1433-1438 [FREE Full text] [doi: 10.2147/NDT.S67008] [Medline: 25120364]

15. Bender BG, Bender SE. Patient-identified barriers to asthma treatment adherence: responses to interviews, focus groups, and questionnaires. Immunol Allergy Clin North Am 2005 Feb;25(1):107-130. [doi: 10.1016/j.iac.2004.09.005] [Medline: 15579367]

16. Horne R, Weinman J. Patients' beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness. J Psychosom Res 1999 Dec;47(6):555-567. [doi: 10.1016/s0022-3999(99)00057-4] [Medline: 10661603]

17. Atkinson MJ, Sinha A, Hass SL, Colman SS, Kumar RN, Brod M, et al. Validation of a general measure of treatment satisfaction, the Treatment Satisfaction Questionnaire for Medication (TSQM), using a national panel study of chronic disease. Health Qual Life Outcomes 2004 Feb 26;2:12 [FREE Full text] [doi: 10.1186/1477-7525-2-12] [Medline: 14987333]

18. Vermersch P, Hobart J, Dive-Pouletty C, Bozzi S, Hass S, Coyle PK. Measuring treatment satisfaction in MS: Is the Treatment Satisfaction Questionnaire for Medication fit for purpose? Mult Scler 2017 Apr;23(4):604-613 [FREE Full text] [doi: 10.1177/1352458516657441] [Medline: 27364322]

19. González CM, Carmona L, de Toro J, Batlle-Gualda E, Torralba AI, Arteaga MJ, et al. Perceptions of patients with rheumatic diseases on the impact on daily life and satisfaction with their medications: RHEU-LIFE, a survey to patients treated with subcutaneous biological products. Patient Prefer Adherence 2017;11:1243-1252 [FREE Full text] [doi: 10.2147/PPA.S137052] [Medline: 28790806]

20. Rubin DT, Dubinsky MC, Martino S, Hewett KA, Panés J. Communication between physicians and patients with ulcerative colitis: reflections and insights from a qualitative study of in-office patient-physician visits. Inflamm Bowel Dis 2017 Apr;23(4):494-501 [FREE Full text] [doi: 10.1097/MIB.0000000000001048] [Medline: 28296817]

21. Pascoe K, Lobosco S, Bell D, Hoskin B, Chang DJ, Pobiner B, et al. Patient- and Physician-reported Satisfaction With Systemic Lupus Erythematosus Treatment in US Clinical Practice. Clin Ther 2017 Sep;39(9):1811-1826 [FREE Full text] [doi: 10.1016/j.clinthera.2017.07.039] [Medline: 28803702]

22. Tintoré M, Alexander M, Costello K, Duddy M, Jones DE, Law N, et al. The state of multiple sclerosis: current insight into the patient/health care provider relationship, treatment challenges, and satisfaction. Patient Prefer Adherence 2017;11:33-45 [FREE Full text] [doi: 10.2147/PPA.S115090] [Medline: 28053511]

23.   Crum A, Zuckerman B. Changing mindsets to enhance treatment effectiveness. J Am Med Assoc 2017 May 23;317(20):2063-2064. [doi: 10.1001/jama.2017.4545] [Medline: 28418538]
24.   Brown T. The power of conversation. Am J Nurs 2017;117(8):67. [doi: 10.1097/01.naj.0000521980.58909.0d]
25.   Ofri D. What Patients Say, What Doctors Hear. Boston, MA: Beacon Press; 2017.
26.   Lippens T, Mackenzie CS. Treatment satisfaction, perceived treatment effectiveness, and dropout among older users of mental health services. J Clin Psychol 2011 Dec;67(12):1197-1209. [doi: 10.1002/jclp.20842] [Medline: 22105535]

Abbreviations

HCP: health care professional
IP: Internet Protocol
MS: multiple sclerosis
RA: rheumatoid arthritis
TSQM: Treatment Satisfaction Questionnaire for Medication

XSL•FO
RenderX

XSL•FO

RenderX