Attempts to Limit Censoring in Measures of Patient Satisfaction

Cindy Nguyen, BSc[1], Joost T P Kortlever, MD[1] ⓘ,
Amanda I Gonzalez, MD[1], David Ring, MD, PhD[1] ⓘ,
Laura E Brown, PhD[2], and Jason R Somogyi, MD[3]

Abstract

Background: Measures of patient satisfaction are increasingly used to measure patient experience. Most satisfaction measures have notable ceiling effects, which limits our ability to learn from variation among relatively satisfied patients. This study tested a variety of single-question satisfaction measures for their mean overall score, ceiling and floor effect, and data distribution. In addition, we assessed the correlation between satisfaction and psychological factors and assessed how the various methods for measuring satisfaction affected net promoter scores (NPSs). **Methodology:** A total of 212 patients visiting orthopedic offices were enrolled in this randomized controlled trial. Patients were randomized to 1 of 5 newly designed, single-question satisfaction scales: (a) a helpfulness 11-point ordinal scale from 0 to 10, (b) a helpfulness ordinal 11-point scale from 0 to 5 (ie, with 1.5, 2.5, etc), (c) a helpfulness 100-point slider, (d) a satisfaction 11-point ordinal scale from 0 to 10, and (e) a willingness to recommend 11-point ordinal scale from 0 to 10. Additionally, patients completed the 2-item Pain Self-Efficacy Questionnaire (PSEQ-2), 5-item Short Health Anxiety Inventory (SHAI-5) Scale, and Patient-Reported Outcomes Measurement Information System (PROMIS) Depression. We assessed mean and median score, ceiling and floor effect, and skewness and kurtosis for each scale. Spearman's correlation tests were used to test correlations between satisfaction and psychological status. Finally, we assessed the NPS for the various scales. **Results:** Ceiling effects ranged from 29% to 68%. The 11-point ordinal helpfulness scale from 0 to 10 had the least ceiling effect (29%). All of the scales were asymmetrically distributed, with the 11-point ordinal scale from 0 to 5 having the most Gaussian distribution (skew = 0.64 and kurtosis = 2.3). Satisfaction scores did not correlate with psychological factors: PSEQ-2 ($r = 0.04$; $P = .57$), SHAI-5 ($r = 0.01$; $P = .93$), and PROMIS Depression ($r = -0.04$; $P = .61$). Net promoter scores varied substantially by scale design, with higher scores corresponding with greater ceiling effects. **Conclusions:** Variations in scale types, text anchors, and lead-in statements do not eliminate the ceiling effect of single-question measures of satisfaction with a visit to an orthopedic specialist. Further studies might test other scale designs and labels.
Level of Evidence: Diagnostic; Level II

Keywords
patient satisfaction scales, ceiling effect, skewness and kurtosis, quality improvement

Introduction

Patient satisfaction is an increasingly used measure of patient experience (1–7). Patients have high expectations regarding when they present for medical care. They expect a high level of customer experience with their medical care and rate their satisfaction on many aspects (ie, medical staff, clinic location, parking). A high level of patient satisfaction largely depends on the interaction and communication with the physician (8). Access to physician and medical center patient satisfaction data might influence a person's decision on where to seek care. Waters et al identified 7 themes influencing satisfaction in orthopedic outpatient clinics,

[1] Department of Surgery and Perioperative Care, Dell Medical School—The University of Texas at Austin, Austin, TX, USA
[2] Center for Health Communication, Dell Medical School—The University of Texas at Austin, Austin, TX, USA
[3] Orthopaedic Surgery, Texas Orthopedics, Austin, TX, USA

Corresponding Author:
David Ring, Department of Surgery and Perioperative Care, Dell Medical School—The University of Texas at Austin, 1701 Trinity Street, Austin, TX 78712, USA.
Email: david.ring@austin.utexas.edu

including trust, relatedness—extent to which a patient feels connected to, respected, or understood by the clinician—, expectations, wait time, visit duration, communication, and empathy (7). Satisfaction is associated with adherence to clinician recommendations (8).

There are several measures of patient satisfaction. An 11-point ordinal measure of satisfaction is the most commonly used. Measures of patient experience tend to have strong ceiling effects (more than half give one of the top 2 scores) that hinder attempts to learn from patient experience and evolve and improve care (9–14). There have been many attempts to limit ceiling effects (9,11,12,14,15). Using a 5-point very positively worded scale significantly lowered mean scores in comparison to a moderately positively worded scale, but there was no difference in skewness and ceiling effects between the scales (9). A 4-point labeled scale and an 11-point numeric scale showed both strong floor and ceiling effects, but ceiling effects were more pronounced in the 4-point labeled scale (11). Comparing a 5-point scale with descriptors with a 10-point scale with descriptors, favored the 5-point scale when assessing means, floor, and ceiling effects (12). A visual analog scale avoided the ceiling effect better than a Likert scale (14). A 10-item visual analog format showed more variability than a 5-item Likert format, 5-item satisfaction format, 5-item valuation format, or 4-item Chernoff faces (15). These previous studies were able to reduce ceiling effects somewhat, but were not able to eliminate them.

The purpose of this study is to compare various scales to measure patient satisfaction in musculoskeletal specialty care. The primary null hypothesis was that there is no difference in mean and median satisfaction, ceiling or floor effect, and data distribution (by looking at skewness and kurtosis) of various satisfaction scales. The secondary hypothesis we assessed was that there is no correlation between scaled satisfaction and psychological status. Finally, we assessed how the satisfaction scores compared to the net promoter scores (NPSs).

Methodology

Study Design

After obtaining approval by our institutional review board (The University of Texas at Austin; protocol number 2018-04-0039), a total of 212 patients from multiple orthopedic practices were asked to participate in this randomized controlled trial. The patients included both new and return visits. Enrollment took place over a 2-month period in 4 orthopedic offices in a large urban area. All English-speaking patients, aged 18 to 89 years, visiting an orthopedic surgeon were eligible for this study. Patients were randomized by research assistants, using an unblocked Excel random number generator, to 1 of 5 satisfaction scales. We were granted a waiver of written informed consent. Completing questionnaires indicated informed consent. All questionnaires were completed using an encrypted tablet via a secure, US Health Insurance Portability and Accountability Act-compliant electronic platform: REDCap (Research Electronic Data Capture: a secure web-based application for building and managing online surveys and databases) (16).

After the patients' clinic visit, the surgeon provided the clinical diagnosis. From there, patients completed a demographic survey, including sex, age, race/ethnicity, marital status, education, work status, insurance status, and comorbidities. Patients were then provided the randomly assigned satisfaction scale, the 2-item short form of the Pain Self-efficacy Questionnaire (PSEQ-2), the 5-item Short Health Anxiety Inventory (SHAI-5), and Patient-Reported Outcomes Measurement Information System (PROMIS) Depression.

Outcome Measures

The primary outcome was patient satisfaction directly after visiting the orthopedic surgeon. We used 5 different satisfaction scales: (a) a helpfulness 11-point ordinal Likert scale from 0 to 10 and 5 anchor points, (b) a helpfulness ordinal 11-point Likert scale from 0 to 5 (ie, with 1.5, 2.5, etc) and 5 anchor points, (c) a helpfulness 100-point slider with 5 anchor points, (d) a satisfaction 11-point ordinal Likert scale from 0 to 10 and 5 anchor points, and (e) a willingness to recommend 11-point ordinal Likert scale from 0 to 10 and 5 anchor points (Figure 1). For all scales, higher scores indicated more satisfaction.

The PSEQ-2 is a measure to assess the confidence in performing activities while in pain (17–19). Each item is scored from 0 to 6, with higher scores suggesting more self-efficacy (17).

The SHAI-5 is a measure to assess health anxiety (20). Each item can be scored 0 to 3, with higher total scores representing more health anxiety.

The PROMIS Depression was used to measure symptoms of depression (21).

Patient Characteristics

Two hundred twelve patients were enrolled in the study, including 90 (42%) men and patient had a median age of 50 (interquartile range [IQR]: 38-62). One hundred and forty-three patients (67%) were self-reported of white ethnicity (Table 1). Median PSEQ-2 score was 11 (9-12), 4 (3-6) for SHAI-5, and 48 (43-53) for PROMIS Depression. Patients had a variety of diagnoses (Supplemental Appendix 1), while the most common diagnoses were trigger finger, carpal tunnel syndrome, ganglion cyst, and knee arthritis.

Statistical Analysis

Continuous data (both psychological measures and satisfaction scales) showed both normal and non-normal distributions. We reported continuous variables using mean, SD, and
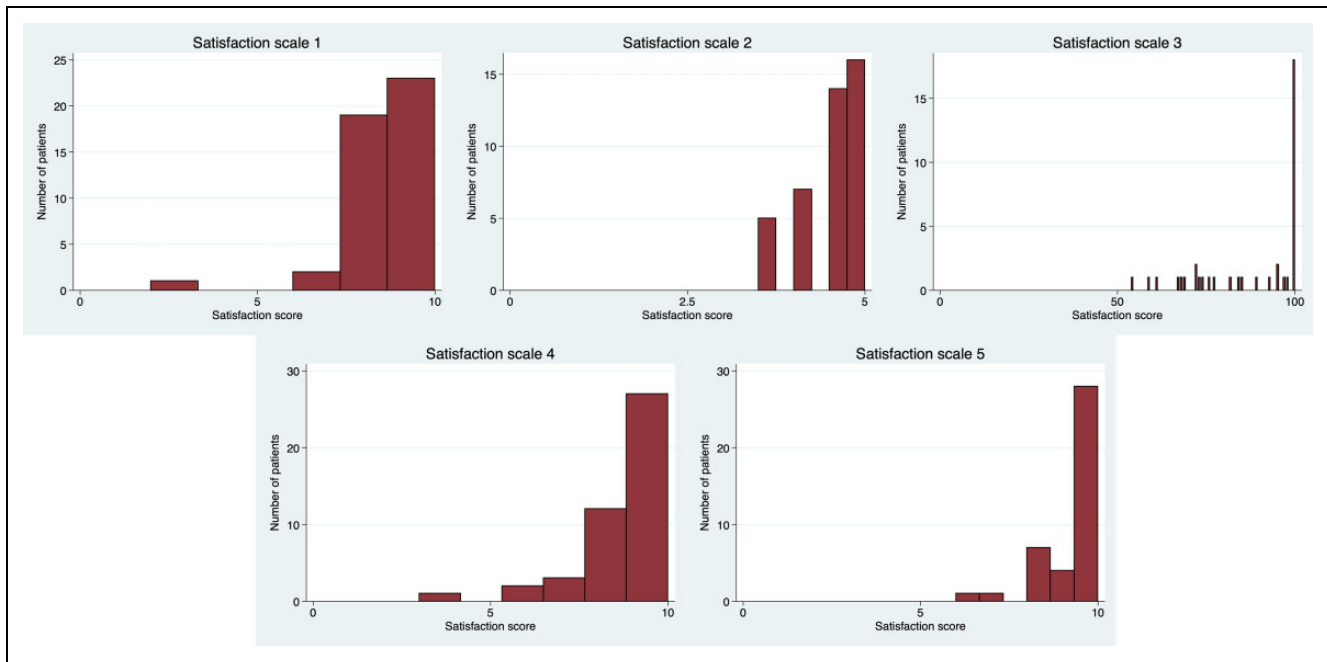
Figure 1. Score Distributions for All Scales.

Table 1. Patient and Clinical Characteristics.[a]

Variables	N = 212
Age in years	50 (38-62)
Male	90 (42)
Race	
White	143 (67)
Other	69 (33)
Marital status	
Married	123 (58)
Single	49 (23)
Divorced/separated/widowed	40 (19)
Level of education	
High school or less	45 (21)
2-year college	42 (20)
4-year college	71 (33)
Postcollege graduate degree	54 (25)
Work status	
Employed	132 (62)
Retired	45 (21)
Homemaker/student/other	35 (17)
Insurance status	
Private/military	155 (73)
Other	57 (27)
Diabetes	21 (9.9)
Other comorbidities	43 (20)
PSEQ-2	11 (9–12)
SHAI-5	4 (3–6)
PROMIS Depression	48 (43-53)

Abbreviations: PROMIS: Patient-Reported Outcomes Measurement Information System; PSEQ-2, 2-item Pain Self-Efficacy Questionnaire; SHAI-5, Short Health Anxiety Inventory.
[a] Continuous variables as median (interquartile range); discrete variables as number (percentage).

median (IQR). Categorical data are presented as frequencies and percentages. We calculated floor and ceiling effect and the skewness and kurtosis of every scale. We scaled every scale to 10 and also standardized every scale. Difference in satisfaction scores were analyzed using Kruskal-Wallis tests and differences between floor and ceiling effects were calculated using Fisher's exact test. When a ceiling effect occurs, there could be a normal distribution, but we are not able to detect this because of a threshold to the measurement. We want to limit loss of data above the threshold of our measurement, which is known as censoring (22). Skewness and kurtosis are rough indicators for a normal distribution of values. Skewness (γ_1) is an index of the symmetry of a distribution. Symmetric distributions have a skewness of 0. If skewness has a positive value, it suggests relatively many low values, having a long right tail. Negative skewness suggests relatively many high values, having a long left tail (23,24). Kurtosis (γ_2) is a measure to describe tailedness of a distribution. Kurtosis of a normal distribution is 3. Negative kurtosis represents little peaked distribution, and positive kurtosis represents more heavy peaked distribution (23,24). If skewness is 0 and kurtosis is 3, there is a normal or Gaussian distribution. Correlation between scaled satisfaction and psychological status is tested by Spearman's correlation tests for satisfaction scores with PSEQ-2, SHAI-5, and PROMIS Depression. We calculated the NPS using our scaled scores (25). The NPS is widely used in the service industry to assess customer satisfaction (26). Respondents are grouped as promotors if they score 9 or 10, as passives if they score 7 or 8, or as detractors if they score from 0 to 6. The NPS is calculated by subtracting the

Table 2. Score Distributions of the Scales.[a]

Scale	Number completed	Visible anchors (N)	Visible numbers	Mean score	Median score	Range	Possible range	Mean scaled score	Median scaled score	Mean scaled range	Possible scaled range	P value scaled score	P value standardized score	Net promoter score
1	45 (21)	Yes (5)	Yes	8.6 ± 1.4	9.0 (8.0-10)	2.0-10	0-10	8.6 ± 1.4	9.0 (8.0-10)	2.0-10	0-10	**.026**	**<.001**	44
2	42 (20)	Yes (5)	Yes	4.5 ± 0.51	4.5 (4.0-5.0)	3.5-5.0	0-5.0	9.0 ± 1.0	9.0 (8.0-10)	7.0-10	0-10			60
3	39 (18)	Yes (5)	Yes	88 ± 15	97 (74-100)	54-100	0-100	8.8 ± 1.5	9.7 (7.4-10)	5.4-10	0-10			44
4	45 (21)	Yes (5)	Yes	8.7 ± 1.4	9.0 (8.0-10)	3.0-10	0-10	8.7 ± 1.4	9.0 (8.0-10)	3.0-10	0-10			53
5	41 (19)	Yes (5)	Yes	9.4 ± 1.0	10 (9.0-10)	6.0-10	0-10	9.4 ± 1.0	10 (9.0-10)	6.0-10	0-10			76

[a]Bold indicates statistically significant difference; continuous variables as mean ± standard deviation (range); or median (interquartile range); discrete variables as number (percentage).

Table 3. Floor and Ceiling Effect and Skewness and Kurtosis of the Scales.[a]

Scale	Floor effect	P value	Ceiling effect	P value	Skewness	Kurtosis
1	0 (0)	–	13 (29)	**.003**	−2.2	11
2	0 (0)		16 (38)		−0.64	2.3
3	0 (0)		18 (46)		−0.83	2.3
4	0 (0)		16 (36)		−1.6	6.9
5	0 (0)		28 (68)		−1.6	4.6

[a]Bold indicates statistically significant difference; discrete variables as number (percentage).

percentage detractors from the percentage promoters. The NPS ranges between −100 and 100.

An a priori power analysis indicated that 200 patients would be needed to assess a difference in satisfaction of 0.5 on a 0 to 10 scale with an effect size of 90%, and α set at .05. In order to account for incomplete responses, we aimed for a sample size of 210 patients.

Results

Difference in Satisfaction Scores

Mean scaled satisfaction scores (range 0-10) were 8.6 ± 1.4 for the helpfulness 11-point ordinal scale from 0 to 10, 9.0 ± 1.0 for the helpfulness ordinal 11-point scale from 0 to 5, 8.8 ± 1.5 for the helpfulness 100-point slider, 8.7 ± 1.4 for the satisfaction 11-point ordinal scale from 0 to 10, and 9.4 ± 1.0 for the willingness to recommend 11-point ordinal scale from 0 to 10 (Table 2)

Because of non-normal distributions, we tested for a difference using median scores. We found a difference in median scaled scores (range 0-10) for the 5 different satisfaction scales: helpfulness 11-point ordinal scale from 0 to 10: 9.0 (8.0-10), helpfulness ordinal 11-point scale from 0 to 5: 9.0 (8.0-10), helpfulness 100-point slider: 9.7 (7.4-10), satisfaction 11-point ordinal scale from 0 to 10: 9.0 (8.0-10), and willingness to recommend 11-point ordinal scale from 0 to 10 (10 [9.0-10]; $P = .026$; Table 2).

Difference in Ceiling and Floor Effects

We found a difference in ceiling effect for the satisfaction scales ($P = .003$), with the willingness to recommend 11-point ordinal scale from 0 to 10 showing the highest ceiling effect (68%) and the helpfulness 11-point ordinal scale from 0 to 10 showing the lowest (29%; Table 3). None of the scales showed a floor effect. No patient scored the lowest rating on any scale.

Skewness and Kurtosis

We found negative skewness for every satisfaction scale (the helpfulness 11-point ordinal scale from 0 to 10: γ_1 −2.2; the helpfulness ordinal 11-point scale from 0 to 5: γ_1 −.64; the

helpfulness 100-point slider: γ_1 −.83; the satisfaction 11-point ordinal scale from 0 to 10: γ_1 −1.6; the willingness to recommend 11-point ordinal scale from 0 to 10: γ_1 −1.6; Table 3). We found positive kurtosis for all satisfaction scales (the helpfulness 11-point ordinal scale from 0 to 10: γ_2 11; the helpfulness ordinal 11-point scale from 0 to 5: γ_2 2.3; the helpfulness 100-point slider: γ_2 2.3; the satisfaction 11-point ordinal scale from 0 to 10: γ_2 6.9; the willingness to recommend 11-point ordinal scale from 0 to 10: γ_2 4.6). The helpfulness ordinal 11-point scale from 0 to 5 had the most Gaussian distribution (γ_1 −.64 and γ_2 2.3).

Correlation Satisfaction and Psychological Status

Scaled satisfaction scores were not significantly correlated with PSEQ-2 ($r = 0.04$; $P = .57$), SHAI-5 ($r = 0.01$; $P = .93$), or PROMIS Depression ($r = -0.04$; $P = .61$; not in table).

Net Promoter Scores

Net promoter scores were 44 for the helpfulness 11-point ordinal scale from 0 to 10, 60 for the helpfulness ordinal 11-point scale from 0 to 5, 44 for the helpfulness 100-point slider, 53 for the satisfaction 11-point ordinal scale from 0 to 10, and 76 for the willingness to recommend 11-point ordinal scale from 0 to 10.

Discussion

Patient satisfaction is increasingly emphasized, beginning to be tied to reimbursement, and even board certification. Patient satisfaction is a component of health care quality and reflects the ability of health care professionals to meet the needs and expectations of their patients (27). It is difficult to distinguish satisfied patients from less-satisfied patients because of ceiling effects of current satisfaction scales (10–12). The purpose of this study was to assess ceiling (and floor) effects of various satisfaction scales. Secondarily, to assess skewness and kurtosis and means of various scaled satisfaction scales.

As with all survey measures, an inherent bias is present that induces limitations to our study. Despite a sample size of 212 patients, we utilized 5 satisfaction scales, which resulted in 39 to 45 patients for each satisfaction scale. A larger sample size may give more variety in scores on every scale and result in a more normal distribution. Secondly, this study was conducted in one large urban area in 4 separate orthopedic offices. Results may not be generalizable to other subspecialty offices or countries, nor to nonspecialty care. Third, most patients of this sample size were self-reported as white, did not have self-reported comorbidities, and were fluent in English. Results may vary in a more mixed study population. For example, Menendez et al found that Spanish-speaking patients are less satisfied than English-speaking patients (28). Fourth, both new and follow-up patients were

eligible for this study. While new patients only have a single encounter from which to develop an opinion on satisfaction, follow-up patients may have a previous relationship with the physician and score accordingly for the entire relationship. Dissatisfied patients are less likely to return for follow-up, resulting in less response variability on the scales. On the contrary, this mix of new and follow-up patients could reflect the usual mix in a specialist's practice contributing to the heterogeneity of our cohort. Fifth, in our study, we did not ask specifically to rate satisfaction with either the process of care or treatment outcome. Asking to rate satisfaction of the process of care or treatment outcome may give different scores (29). Sixth, we did not perform separate analysis for different diagnoses because none of the diagnoses had sufficient numbers and—based on prior work—we did not feel diagnosis was likely to be important (30,31). Differentiating between preoperative and postoperative patients and between acute and chronic disorders may lead to different results. The results of this type of study might be different if restricted to people with a specific diagnosis. Patients with more pain, for example, have the tendency to rate their overall hospital experience lower, giving different satisfaction scores (32). Finally, we did not formally validate any of these measures. That step can be completed once we have successfully reduced or eliminated ceiling effects.

The satisfaction scales showed differences in mean and median satisfaction. Our mean scaled scores range from 8.6 to 9.4 and our median scaled scores range from 9.0 to 10, which highlights the issue we sought to address in evaluating the ceiling effect (5). We hypothesize that patients are likely being respectful, deferential, and appreciative upon completion of their office visit despite our earnest attempts to obtain nonbiased feedback to identify areas of improvement in the clinical setting.

The satisfaction scales showed differences in ceiling effects for our various satisfaction scales. We have not been able to eliminate the ceiling effect for measures of satisfaction with a visit to an orthopedic specialist. Consistent with a study of Dell-Kuster et al, a numeric 11-point Likert scale provides the least ceiling effect (11). However, this scale still has a substantial (29%) ceiling effect, implying loss of data beyond the threshold of this scale. In our study, we found that an 11-point Likert scale with 5 visible anchors seems to have less ceiling effect when inquiring about helpfulness compared to asking about satisfaction (36%) or willingness to recommend (68%). Floor effects are uncommon with satisfaction scores, except for the occasional very dissatisfied patient who chooses the lowest score, but otherwise scores below 5 are uncommon.

Differences in data distribution were identified by analyzing skewness and kurtosis for our various satisfaction scores. The scales were asymmetrically distributed. Every scale had a negative skewness reflecting censoring at the top values. The helpfulness ordinal 11-point scale from 0 to 5 had the most normal skewness (−0.64). None of the scales had a normal kurtosis of 3. The helpfulness ordinal 11-point scale

from 0 to 5 and helpfulness 100-point slider were closest to a normal kurtosis, with a kurtosis of 2.3. We were able to determine that the type of scale affects censoring, skewness, and kurtosis and therefore the distribution.

There was no significant correlation between scaled satisfaction and psychological status. Intrinsic factors that specifically drive patient satisfaction remain unclear (33). Smith et al studied extrinsic factors that influence patient satisfaction scores and found that patients are more satisfied if they feel that their physician provides them with compassionate, coordinated care (34). The characteristics of the individual provider were the largest factor influencing patient satisfaction.

Net promoter scores varied substantially by scale design, with higher scores corresponding with greater ceiling effects. It may be that organizations will need to choose between higher scores and learning more from people by having less censoring. Our 11-point ordinal "willingness to recommend the doctor" scale, which is similar to the Friends and Family Test (FFT; the FFT asks how likely a patient is to recommend the same service to a friend or family member with the same condition) (27) had the largest ceiling effect and the best NPS.

Conclusion

Future studies might consider using other scale methods to try to limit ceiling effects and censoring. One option would be developing a computer adaptive test to measure patient satisfaction. A computer adaptive test uses answers to previous questions, customizing the subsequent question, resulting in a higher level of precision using fewer questions, and better able to limit ceiling effects (35). Another option would be using open-ended questions where answers would be scored to assess satisfaction, rather than using numeric scales. We hope others will join us in attempts to reduce ceiling effects in patient experience measures. When effective techniques are identified, the resulting new scales can go through a more thorough validation.

Authors' Note

Declaration of Conflicting Interests

Funding

ORCID iD

Joost T P Kortlever https://orcid.org/0000-0003-3825-5546
David Ring, MD, PhD https://orcid.org/0000-0001-5947-5316

Supplemental Material

Supplemental material for this article is available online.

References

1. Anderson RT, Camacho FT, Balkrishnan R. Willing to wait?: The influence of patient wait time on satisfaction with primary care. BMC Health Serv Res. 2007;7:31.
2. Glaser KM, Markham FW, Adler HM, McManus RP, Hojat M. Relationships between scores on the Jefferson Scale of physician empathy, patient perceptions of physician empathy, and humanistic approaches to patient care: a validity study. Med Sci Monit. 2007;13:CR291-4.
3. Mathews M, Ryan D, Bulman D. What does satisfaction with wait times mean to cancer patients? BMC Cancer. 2015;15:1017.
4. Menendez ME, Chen NC, Mudgal CS, Jupiter JB, Ring D. Physician empathy as a driver of hand surgery patient satisfaction. J Hand Surg Am. 2015;40:1860-5.e2.
5. Parrish RC, Menendez ME, Mudgal CS, Jupiter JB, Chen NC, Ring D. Patient satisfaction and its relation to perceived visit duration with a hand surgeon. J Hand Surg Am. 2016;41:257-62.e1-4.
6. Teunis T, Thornton ER, Jayakumar P, Ring D. Time seeing a hand surgeon is not associated with patient satisfaction. Clin Orthop Relat Res. 2015;73:2362-8.
7. Waters S, Edmondston SJ, Yates PJ, Gucciardi DF. Identification of factors influencing patient satisfaction with orthopaedic outpatient clinic consultation: a qualitative study. Man Ther. 2016;25:48-55.
8. Esselman PC, White BF, Chimes GP, Kennedy DJ. Patient satisfaction surveys: tools to enhance patient care or flawed outcome measures? PM and R. 2013:9.
9. Bjertnaes O, Iversen HH, Garratt AM. The universal patient centeredness questionnaire: scaling approaches to reduce positive skew. Patient Prefer Adherence. 2016;10:2255-60.
10. Cleary PD, Edgman-Levitan S, McMullen W, Delbanco TL. The relationship between reported problems and patient summary evaluations of hospital care. QRB Qual Rev Bull. 1992;18:53-9.
11. Dell-Kuster S, Sanjuan E, Todorov A, Weber H, Heberer M, Rosenthal R. Designing questionnaires: healthcare survey to compare two different response scales. BMC Med Res Methodol. 2014;14:96.
12. Garratt AM, Helgeland J, Gulbrandsen P. Five-point scales outperform 10-point scales in a randomized comparison of item scaling for the patient experiences questionnaire. J Clin Epidemiol. 2011;64:200-7.

13. Hirsch O, Keller H, Albohn-Kuhne C, Krones T, Donner-Banzhoff N. Pitfalls in the statistical examination and interpretation of the correspondence between physician and patient satisfaction ratings and their relevance for shared decision making research. BMC Med Res Methodol. 2011;11:71.

14. Voutilainen A, Pitkaaho T, Kvist T, Vehvilainen-Julkunen K. How to ask about patient satisfaction? The visual analogue scale is less vulnerable to confounding factors and ceiling effect than a symmetric Likert scale. J Adv Nurs. 2016;72:946-57.

15. Castle NG, Engberg J. Response formats and satisfaction surveys for elders. Gerontologist. 2004;44:358-67.

16. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) – a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inf. 2009;42:377-81.

17. Bot AGJ, Nota SPFT, Ring D. The creation of an abbreviated version of the PSEQ: the PSEQ-2. Psychosomatics. 2014;55:381-5.

18. Briet JP, Bot AGJ, Hageman MGJS, Menendez ME, Mudgal CS, Ring DC. The Pain Self-Efficacy Questionnaire: validation of an abbreviated two-item questionnaire. Psychosomatics. 2014;55:578-85.

19. Nicholas MK, McGuire BE, Asghari A. A 2-item short form of the Pain Self-Efficacy Questionnaire: development and psychometric evaluation of PSEQ-2. J Pain. 2015;16:153-63.

20. Bot AGJ, Becker SJE, van Dijk CN, Ring D, Vranceanu AM. Abbreviated psychologic questionnaires are valid in patients with hand conditions. Clin Orthop Relat Res. 2013;471:4037-44.

21. Schalet BD, Pilkonis PA, Yu L, Dodds N, Johnston KL, Yount S, et al. Clinical validity of PROMIS depression, anxiety, and anger across diverse clinical samples. J Clin Epidemiol. 2016;73:119-27.

22. Spriensma AS, Eekhout I, De Boer MR, Luime JJ, De Jong PH, Bahçecitapar MK, et al. Analysing outcome variables with floor effects due to censoring: a simulation study with longitudinal trial data. Epidemiol Biostat Public Heal. 2018:9.

23. Ho AD, Yu CC. Descriptive statistics for modern test score distributions: skewness, kurtosis, discreteness, and ceiling effects. Educ Psychol Meas. 2015;75(3):365-88.

24. Kim H-Y. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restor Dent Endod. 2013;38:52-4.

25. Inc SS. What Is Net Promoter? [Internet]. 2019. Accessed March 11, 2019. https://www.netpromoter.com/know/

26. Hamilton DF, Lane J V, Gaston P, Patton JT, Macdonald DJ, Simpson AHRW, et al. Assessing treatment outcomes using a single question: the net promoter score. Bone Joint J. 2014;96-B:622-8.

27. Stirling P, Jenkins PJ, Clement ND, Duckworth AD, McEachan JE. The net promoter scores with friends and family test after four hand surgery procedures. J Hand Surg Eur Vol. 2018. doi:10.1177/1753193418819686

28. Menendez ME, Loeffler M, Ring D. Patient satisfaction in an outpatient hand surgery office. Qual Manag Health Care. 2015;24:183-9.

29. Hudak PL, McKeever P, Wright JG. The metaphor of patients as customers: implications for measuring satisfaction. J Clin Epidemiol. 2003;56:103-8.

30. Ottenhoff JSE, Derkzen L, Reichel LM, Vagner G, Loeb MD, Ring D. Satisfaction with specific and nonspecific diagnoses. J Hand Surg Am. 2019;44:460-466.e1.

31. Jayakumar P, Teunis T, Vranceanu A-M, Lamb S, Ring D, Gwilym S. Relationship between magnitude of limitations and patient experience during recovery from upper-extremity fracture. JB JS Open Access. 2019;4:3.

32. Otani K, Chumbler NR, Herrmann PA, Kurz RS. Impact of pain on patient satisfaction integration process: how patients with pain combine their health care attribute reactions. Heal Serv Res Manag Epidemiol. 2015;2. doi:10.1177/2333392815615103

33. Lafage R, Schwab F, Frangella NJ, Lafage V, Klineberg E, Horn SR, et al. Baseline mental status predicts happy patients after operative or non-operative treatment of adult spinal deformity. J Spine Surg. 2018;4:687-95.

34. Smith ML, Congiusta S, Kasabian AK, Thorne C, Nash IS, Coppa GF, et al. Factors influencing patient satisfaction in plastic surgery. Plast Reconstr Surg. 2018;142:820-5.

35. Fidai MS, Saltzman BM, Meta F, Lizzio VA, Stephens JP, Bozic KJ, et al. Patient-reported outcomes measurement information system and legacy patient-reported outcome measures in the field of orthopaedics: a systematic review. Arthroscopy. 2018;34:605-614.

Author Biographies

Cindy Nguyen is a student working in the field of Surgery in The Netherlands. She did research in 2019 with D. Ring at the University of Texas at Austin.

Joost T P Kortlever is PhD student working in the field of Surgery in The Netherlands. He worked 2 years with D. Ring as a PhD Student at the University of Texas at Austin.

Amanda I Gonzalez is an orthopaedic surgeon working in Switzerland. She did research in 2019 with D. Ring at the University of Texas at Austin.

David Ring is the Associate Dean for Comprehensive Care and Professor of Surgery and Psychiatry at Dell Medical School. His interests include trauma and post-traumatic reconstruction in the arm, quality and patient safety, common arm illnesses and psychosocial aspects of arm illness.

Laura E Brown is a Postdoctoral Fellow at the Center for Health Communication at Dell Medical School. She is involved with research with D. Ring at the University of Texas at Austin.

Jason R Somogyi is an orthopaedic surgeon working at Texas Orthopedics. He is involved with research with D. Ring at the University of Texas at Austin.