Patient Satisfaction/Patient Experience, Patient-Reported Outcomes, and Healthcare Quality: Are We Focusing on the Wrong Metrics?

by

Marsha Haley, M.D.

BS, Old Dominion University, 1998

MD, University of Virginia, 2004

Submitted to the Graduate Faculty of the

Department of Health Policy and Management

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Public Health

University of Pittsburgh

2022

UNIVERSITY OF PITTSBURGH

GRADUATE SCHOOL OF PUBLIC HEALTH

This essay is submitted

by

Marsha Haley, M.D.

on

April 15, 2022

and approved by

Essay Advisor: Mark Roberts, M.D., MPP, Professor, School of Public Health, Health Policy
and Management, University of Pittsburgh

Essay Reader: Janel Hanmer, M.D., PhD, Medical Director, UPMC Patient Reported Outcomes
Center, School of Medicine, University of Pittsburgh

Essay Reader: Amy Triola, Director of Patient Satisfaction, Office of Patient Experience, Wolff
Center at UPMC, University of Pittsburgh

Patient Satisfaction/Patient Experience, Patient-Reported Outcomes, and Healthcare Quality: Are We Focusing on the Wrong Metrics?

Marsha Haley, MPH

University of Pittsburgh, 2022

Abstract

Background: To meet the Institute of Medicine goal of patient-centered healthcare quality, Patient Satisfaction/Patient Experience (PS/PE) has evolved to become 25 percent of the U.S. government's Medicare and Medicaid Value-Based Purchasing (VBP) Program. PS/PE is most commonly measured by CAHPS® and Press Ganey® (PG) surveys. Patient-reported Outcomes (PROs) are also measured as part of the VBP, but are not used for reimbursement.

Materials and Methods: We performed a MEDLINE literature search to evaluate whether high-level evidence (randomized controlled trials, systematic reviews, or meta-analyses) exist to show that CAHPS®, PG®, and PROMIS® survey scores correspond with improved quality of healthcare.

Results: The number of publications on PS/PE and PRO has increased significantly since the initiation of patient surveys. One systematic review found that CAHPS® scores were inconsistently associated with patient-reported quality. Zero studies were found to show that PG® scores were associated with improved quality. Five studies included PROMIS® had convergent validity with legacy measures of PRO. Further review of the literature showed concerns inherent in the way CAHPS® and PG® are used which may adversely affect healthcare quality.

Conclusion: PS/PE is weighted highly for reimbursement. The most-utilized PS/PE surveys have psychometric and methodologic flaws and, using the above literature search

method, we found very little high-level data showing that CAHPS® and PG® scores correlate with improved healthcare quality as defined by the IOM. PROMIS® is a reliable, valid, and precise measure of patient-reported health status. CMS should consider a policy change to decrease the weight of PS/PE in the VBP program. At the same time, CMS should consider incorporating PRO data into reimbursement.

Table of Contents

List of Tables

List of Figures

Preface

I am most grateful for the comments and suggestions provided by my Advisor and Essay Readers.

1.0 Introduction

The Institute of Medicine (IOM) defines healthcare quality as "the degree to which healthcare services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge." In 2001, the IOM designated six quality domains for modern healthcare systems. One of these aims is patient centeredness, which "relates to meeting patients' needs and preferences and providing education and support" (1). The three core values of patient centeredness are the following: considering patients' needs, wants, perspectives, and individual experiences; offering patients opportunities to provide input into and participate in their own care; and enhancing partnerships and understanding in the patient-physician relationship (2).

Patient Satisfaction (PS) is a gauge of whether a patient's expectations about a health encounter were met. PS consists of three domains: medically necessary care that improves outcomes; interventions that patients or families desire that are medically unnecessary and may negatively affect health outcomes; and the humanistic aspects of healthcare including good communication and respect, as well as conveniences such as parking, designer hospital gowns, and hospital aesthetics and architecture (3). Patient Experience (PE) assesses the aspects of healthcare delivery that patients value highly, including timely appointments, easy access to information, and good communication with healthcare providers (4). Since PS and PE are not directly observable, they are measured by surveys in which patients report their perceptions of their healthcare experience (5). Although the Centers for Medicare and Medicaid Services (CMS) states that surveys are meant to measure PE and not PS, there is significant overlap in the definitions and in the literature; (6) therefore we will refer to this metric as one entity, PS/PE.

1

To help achieve the aim of patient centeredness, the Affordable Care Act (ACA) of 2010 began financially awarding or penalizing hospitals based on PS/PE scores as part of its Hospital Value-Based Purchasing (VBP) Program, which is essentially a pay-for-performance program. Initially, Medicare withheld one percent of base operating Medicare severity diagnosis-related group payments for an incentive bonus fund; as of fiscal year 2021, this amount is two percent. The incentive fund is then distributed based on the Total Performance Score (TPS). Currently, PS/PE is 25 percent of the TPS, expressed as the domain "Person and Community Engagement". (7). PS/PE is most commonly measured by Consumer Assessment of Healthcare Providers and Systems (CAHPS®) surveys, which are developed by CMS. CAHPS® is a required survey for inpatient acute care hospitals who participate in CMS healthcare funding/reimbursements (8). More recently, private survey firms such as Press Ganey® (PG) have been approved by CMS to administer outpatient surveys (9).

Patient-Reported Outcomes (PROs) are also a high priority for CMS and other organizations. PROs are defined as "any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else". PROs measure what patients are able to do and how they feel by having the patients answer questionnaires (10). Originally, there were many ways to measure PROs, but there was a lack of standardization and comparability between methods. In 2004, the Patient-Reported Outcomes Measurement Information System (PROMIS®) was developed as a web-based resource which addressed the shortcomings of the earlier methods. PROMIS® is a measure of PROs for physical, mental and social well-being that is increasingly being used in the United States (11). PROs, however, are not widely utilized in clinical practice (12), and are rarely used for reimbursement purposes (13).

In this paper, we explore how PE/PS and PROs are currently used to measure the quality of healthcare, and the benefits and limitations of each. We also examine the current use of PE/PS in determining reimbursement, and whether CMS should incorporate PROs into reimbursement.

1.1 Materials and Methods

Literature Search

We performed systematic searches using the National Library of Medicine's MEDLINE (PubMed). We searched for the phrases "patient satisfaction" and "patient experience." The earliest healthcare surveys started in 1985 (14), so we included articles published since 1985 in the English language. Patient satisfaction yielded 138,423 results, and patient experience yielded 394,691 results. We performed the same database search using the phrase "patient reported outcomes." This yielded 105,360 results.

To include those studies with the highest levels of evidence based on the new evidence pyramid (15), we narrowed the results to include randomized controlled trials, systematic reviews, and meta-analyses. Patient satisfaction yielded 20,616 results; patient experience yielded 18,375 results; and patient reported outcomes yielded 20,714 results.

Inclusion Criteria and Screening

Given the large number of studies on these topics, we focused on studies that utilized the two most widely utilized surveys for PS/PE (CAHPS® and PG®) (16), and the PROMIS® PRO measurement system, under the search terms above. We tailored the date range depending on when each survey was implemented: 1985 for PG®, 2002 for CAHPS® (14), and 2004 for PROMIS® (11). PG® yielded 8 results, CAHPS® yielded 23 results, and PROMIS® yielded 167 results.

Since our interest is in measuring the effect of PS/PE and PRO utilization on healthcare quality by increasing the likelihood of desired health outcomes, we excluded studies that did not include the surveys as an independent variable. Studies done outside of the U.S. were also excluded.

Using these criteria, CAHPS® yielded 1 result, Press Ganey® yielded 0 results, and PROMIS® yielded 5 results (Table 1).

1.1.1 Results

Publications on PS/PE began to appear in significant numbers in the early 1990s; at that time, there were approximately 4,000 studies on the topic. In 2010, when the ACA was implemented, there were approximately 20,000 studies. Since 2017, there have been 30,000 - 40,000 publications per year on the topic of PS/PE. (Figure 1). Similarly, PRO publications started appearing in the early 1990s; at that time, there were fewer than 100 studies on the topic. In 2010 there were over 2,000, and since 2018 there have been over 10,000 studies per year on PRO. (Figure 2). A similar trend is seen if we limit the results to high-level evidence, albeit with fewer numbers of studies (17).

Despite the large number of studies, there has been a paucity of data showing high-level evidence that the two most commonly used PS/PE surveys correlate with increased likelihood of desired health outcomes. There are significantly more studies showing that PRO helps achieve this goal. In fact, as we will explore in the Discussion section, PS/PE surveys have psychometric and methodologic flaws, and some data have shown an inverse relationship between scores and desired health outcomes.

In the article by Navarro et al, a systematic review of the literature was performed to summarize the association between CAHPS® ratings and healthcare clinical outcomes or quality measures of care. (Although there have been previous reviews on PE/PS, this was the first study to examine the CAHPS® survey specifically.) Higher ratings of patient-provider communication were associated with higher self-reported health scores, fewer emergency room visits, fewer and shorter hospital stays, and higher likelihood of being prescribed guideline management therapies for cardiovascular disease. Studies conflicted on whether better patient-provider communication was associated with decreased mortality and readmission. Higher ratings of customer service and getting care quickly were associated with earlier stage of breast cancer at diagnosis among Black patients only. Higher ratings of customer service were associated with improvement of antidepressant compliance, cholesterol testing after a cardiovascular event, and mammography. Higher ratings of getting needed care quickly were associated with higher percentages of diabetes eye exams. There were no associations between patient-provider communication ratings and compliance with diabetic, blood pressure, and antidepressant medication. There was no association between CAHPS® experience measures and kidney transplant failure, inpatient complications, hemoglobin A1c, cholesterol levels, and blood pressure. The authors concluded that the use of CAHPS® was substantiated for certain outcomes and independent measures of patient-reported quality; however, inconsistent findings point to the need for more research (18).

In the study by Hadlandsmyth et al, three PROMIS measures (anxiety, depression, and pain interference) demonstrated preliminary validity in 67 veterans three months after orthopedic surgery, with a high degree of overlap between PROMIS® depression and anxiety measures. The authors recommended replicating these results with a larger sample size (19).

Kroenke et al. assessed the effectiveness of providing PROMIS® symptoms scores to clinicians on symptom outcomes for SPADE symptoms (sleep, pain, anxiety, depression, low energy). After completing PROMIS® questionnaires, the 300 study participants were randomized to a feedback group where the clinician either did or did not receive a visual display of symptom scores. This intervention was no superior to usual care in improving symptom outcomes (20).

Wilford et al. evaluated the PROMIS® Emotional Distress-Depression and Anxiety Short forms in patients with cervical cancer. The short forms reliably and validly assessed cervical cancer-specific emotional distress, and performed as well or better than legacy measures including the Functional Assessment of Cancer Therapy-Cervical and the Impact of Event Scale (21).

In the study by Merriwether et al, the authors evaluated the reliability and construct validity of PROMIS® static short-form instruments in patients with fibromyalgia. PROMIS® demonstrated convergent validity with the Fibromyalgia Impact Questionnaire, a legacy measure of fibromyalgia disease severity. In addition, PROMIS® allowed for comparison of outcomes across diverse clinical populations of patients with and without pain (22).

Lazor et al. performed a systematic review to identify reliable and valid instruments to measure anxiety in children and adolescents undergoing hematopoietic stem cell transplantation. PROMIS® Pediatric Anxiety Short Form was one of the three multi-item instruments that was found to be reliable and valid (23).

1.1.2 Discussion

One of the difficulties in continuing to strive for the attainment of "quality healthcare" is defining precisely what this term means, since different definitions may result in different approaches. When the IOM initially published "To Err is Human," quality was framed in terms of

decreasing medical error (24). In the IOM's subsequent report, the meaning of healthcare quality was expanded to include six domains: care that is safe, effective, patient-centered, timely, efficient, and equitable (25). The IOM's definition has since evolved to the current one discussed earlier: "the degree to which healthcare services for individuals and populations increase the likelihood of desired health outcomes and are consistent with current professional knowledge" (1). Allen-Duck et al. employed a concept analysis framework to define healthcare quality; their definition is "the assessment and provision of effective and safe care, reflected in a culture of excellence, resulting in the attainment of optimal or desired health" (26). While these definitions have much in common, translating these ideas into effective practice to measure and improve quality has been challenging.

Quality is difficult to measure directly; therefore, surrogates are used. There is limited evidence that many surrogate quality measures - such as biomarkers, risk-factor control, or care processes – lead to improved health outcomes. CMS has therefore attempted to define how quality should be measured. CMS defines a quality measure as a "standard for measuring the performance of healthcare providers to care for patients and populations." The components of the measure should include a numerator, a denominator, and a denominator exclusion (members of the population who should not be included in the denominator) (27). Surveys are a common method to measure how patients perceive the impact of specific interventions on their health conditions and quality of life. The choice and use of measurement instruments should consider psychometric properties such as validity, reliability, and clinical utility. If psychometric properties are not considered, the results can lead to bias, increased treatment costs, inappropriate clinical decisions, and ethical issues (28).

To help achieve the IOM-designated quality goal of patient-centeredness, the ACA mandated the use of measures of the quality of care. The law repeatedly refers to patient-

centeredness, patient satisfaction, and patient experience in its provisions. These provisions have, in turn, become regulations for Medicare's VBP Program (29). Base on the CMS definition above, VBP measures should be reliable, valid, and associated with improved clinical outcomes. Starting in 2016, CMS began to link payment to 90 day complication rates, PS/PE, and PRO in joint replacements. In 2018, CMS implemented the Bundled Payments for Care Improvement (BPCI) Model, which reimburses participants for clinically appropriate services provided to Medicare patients using a bundled payment methodology. The goal of this model is to "promote seamless, patient-centered care throughout each Clinical Episode." At the end of the 90 day performance period, actual expenditures are reconciled against a target price based on historical claims data, after a 3% discount is applied. If the expenditures are below the target price, participants may receive a payment; conversely, if expenditures are above the target price, participants may owe money back to CMS. As of 2021, the model has expanded to include many common inpatient diagnoses (30).

CAHPS® was developed so that health plan members could provide feedback about access, provider communication, and other measures of quality. Launched in 1995, it was a standardized questionnaire that was initially used to compare results among different health plans (31). CAHPS® surveys are distributed to a random sample of patients, and they been found to have good reliability and validity in the literature (32). A psychometric report from the PG® vendor demonstrated validity and high reliability (33).

In parallel, in the 1990s, PROs, began to be collected and reported in standardized, validated formats that could provide feedback about ongoing treatment decisions (34). Health-related quality of life as measured by PROs is one way to measure attainment of optimal health.

PROMIS® is a National Institutes of Health (NIH) sponsored publicly available system of highly reliable, precise measures of patient-reported health status (11).

When considered as a portion of overall healthcare costs, the quest for patient-centered and patient-assessed quality is an expensive one. U.S. National Health Expenditure accounted for nearly 20 percent of the Gross Domestic Product in 2020, and is predicted to grow at an average annual rate of 5.4 percent annually. Medicare spending grew to 20% of total healthcare expenditures in 2020. Among major payors, Medicare is expected to experience the fastest spending growth, largely due to projected enrollment growth (35). With current withholding of 2 percent of hospitals' base operating payments, PS/PE, at 25 percent of this total, holds a significant amount of financial power over the day-to-day operations of healthcare facilities and physician practices.

Concerns have been raised about the unintended consequences of placing such significant weight on PS/PE. If a physician is faced with financial penalties for low PS/PE scores, he or she may be incentivized to offer treatments that are not recommended by the Choosing Wisely Initiative (36). There can be negative consequences for safety net hospitals, since they typically do worse on PS/PE metrics. Faced with penalties for low PS/PE scores, physicians could avoid working in these hospitals, exacerbating socioeconomic and racial disparities (37). Jerant et al. found that patients who received more opioid prescriptions gave higher PS/PE scores (38); this could potentially incentivize hospitals and physicians to prescribe more opioids to improve reimbursement. In fact, due to this concern, Medicare recently stopped using pain management questions as part of CAHPS® (39).

Other authors have raised the important question of what is actually being measured by PS/PE surveys. Chang et al. found that CAHPS® scores were not associated with technical quality

of care (40). Chen et al. found that poor PS/PE scores were associated with certain unmodifiable characteristics such as race and socioeconomic status, thus calling into question their validity in assessing provider performance (41). PS/PE scores are lower for women and underrepresented physicians (42).

Response bias can significantly impact the results of PS/PE surveys (43), and limited sample size can affect the validity of scores (36). Presson et al. performed a psychometric analysis of PG® and found that while it demonstrated suitable properties for most metrics, the high ceiling rate can have a notable impact on physician scores. For example, a 0.5 score decrease can result in a 30 percent drop in percentile rank (44). Hospital administrators seek a high score in PS/PE by counting only "top-box" scores. When using the PG® Likert scale, this essentially creates a system with one passing and four "failing" grades. If patients are unaware that a 4 is a low and not an intermediate-high grade, the usefulness of the data is brought into question (45). For HCAHPS®, the scale is 1-10, with 9 or 10 being tox box, 7 or 8 as being middle box, and 6 or below being bottom box. The higher a hospital's top box score, the higher it ranks among participating hospitals (46). Performing the measurements this way could be considered to be at odds with the CMS definition of a quality measure (27).

PS/PE scores have also been shown to be an inverse measure of quality in some scenarios. Fenton et al. found that in a nationally representative sample, higher PS/PE scores were associated with greater inpatient use, higher overall healthcare and prescription expenditures, and increased mortality (47). Manary et al. recognized the conflicts in the literature regarding PS/PE and opined that this could be due to the timing of the surveys; when there was a large gap between the experience and the surveys, as in the Fenton study, the association between PS/PE scores and outcomes skewed negative (48).

10

While PS/PE and PROs measure different aspects of healthcare quality, both seek to improve healthcare quality by incorporating patient feedback. PROMIS® provides assessments of patient-reported health status and well-being. It was developed and validated with scientific methods to have greater precision than most conventional measures, to enhance communication between clinicians and patients in diverse clinical settings, and to allow the physician to obtain a validated assessment of how the patient is feeling and functioning and track those values over time (49). Unlike PS/PE surveys, PROMIS® has minimal ceiling and floor effects. The clinician can repeat the survey after an intervention and objectively measure whether the patient's perception of the health outcome has improved. PROMIS® can also be modified for different clinical situations, including mental health, pediatrics, and surgery.

Previous authors have opined that the weight of PS/PE surveys in the VBP program is disproportionately high, and have recommending lowering the percentage of weight (50) and redesigning the VBP program to ensure the weights appropriately reflect the level of importance (51). Currently, PS/PE has the same weight as clinical outcomes and patient safety. PROMIS® can provide an accurate and evidence-based metric for clinicians and hospitals to improve performance in response to patient feedback. Care would need to be taken, however, to avoid the types of unintended consequences seen with implementation of PS/PE surveys, such as strains of time and resources; alteration of provider and system behaviors (such as declining to take sicker patients likely to have lower PRO scores); and a decrease in primary care continuity due to quick access to any provider to improve scores (52).

Our study had some limitations. By limiting the data to randomized controlled trials, systematic reviews, and meta-analyses, we may have excluded quality studies. Many PS/PE studies are missed because it is difficult to randomize the collection of quality metrics. In addition,

11

for PS/PE, we only assessed CAHPS® and PG®, and PROMIS® for PRO. Including other surveys

may have altered the results, although this effect would likely be modest for PS/PE given the large

market share of the two surveys.

1.1.3 Conclusions

Seeking the feedback of patients is critical in the pursuit of healthcare quality. This

ensures that patients have a voice, serve as protection against system failures, and allow

physicians and hospitals to identify areas for improvement (29). The most-utilized PS/PE

surveys, CAHPS® and PG®, have questionable relationships with desired health outcomes, and

our literature review found very little high-level data showing that CAHPS® and PG® scores

correlate with improved healthcare quality. PROMIS® is a measurement system of highly

reliable, precise measures of patient-reported health status (11). CMS should consider a policy

change to decrease the weight of PS/PE in the VBP program. At the same time, CMS should

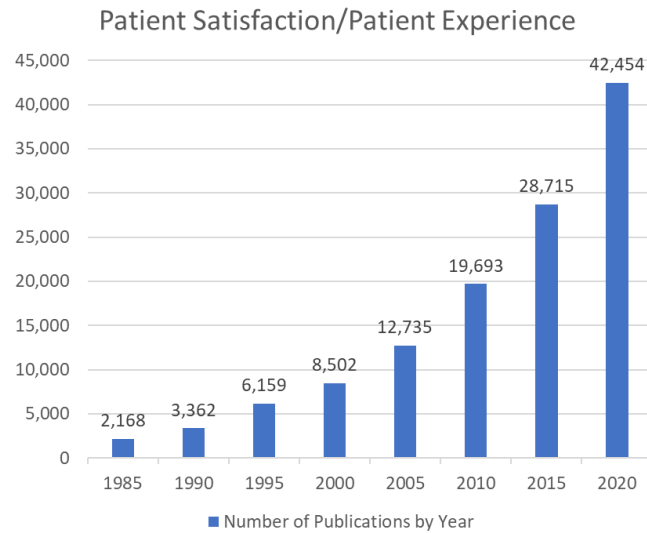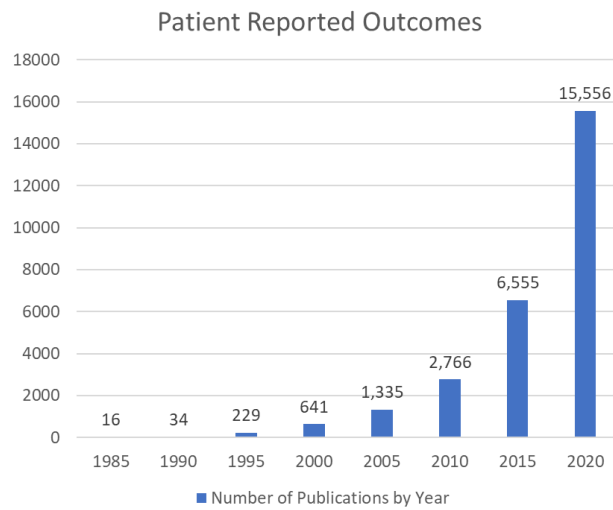consider incorporating PRO data into reimbursement.

2.0 Figures

Patient Satisfaction/Patient Experience

Figure 1. PS/PE Publications 1985 – 2020

Patient Reported Outcomes

Figure 2. PRO Publications 1985 – 2020

2.1 Tables

Table 1. List of Studies

Research Study	Design Type	Data/Methods	Findings
Navarro et al. (2021)	Systematic Review	Included studies investigating associations between CAHPS® composite ratings and health care clinical outcomes or quality measures of care	The use of CAHPS® was substantiated for certain outcomes and independent measures of patient-reported quality; however, inconsistent findings point to the need for more research
Hadlandsmyth et al (2020)	Randomized control trial	Evaluated the reliability and validity of PROMIS measures with 67 US Military Veterans following orthopedic surgery	PROMIS Anxiety, PROMIS Depression, and PROMIS Pain Interference demonstrated validity and reliability
Kroenke et al (2018)	Randomized Clinical Trial	After completing the PROMIS symptom measures electronically immediately prior to their visit, 300 participants were randomized to a feedback group in which their clinician received a visual display of symptom scores or a control group in which scores were not provided to clinicians	Simple feedback of symptom scores to primary care clinicians in the absence of additional systems support or incentives is not superior to usual care in improving symptom outcomes.
Wilford et al (2018)	Randomized clinical trial	A 15-item questionnaire was used in a cervical cancer biobehavioral randomized clinical trial, testing psychosocial telephone counseling (PTC) against usual care (UC). It was administered to 204 patients prior to	The PROMIS depression and anxiety short forms reliably and validly assess cervical cancer-specific emotional distress, capture salient features of distress in this population, and perform as well or

		randomization, four months post-enrollment, and nine months post-enrollment, together with legacy measures of depression.	better than legacy measures.
Merriwether et al (2016)	Analysis of baseline data from the randomized FAST trial	Participants completed the Revised Fibromyalgia Impact Questionnaire (FIQR) and 10 PROMIS static SF instruments. Internal consistency was calculated using Cronbach alpha. Convergent validity was examined against the FIQR using Pearson correlation and multiple regression analysis.	Select PROMIS instruments demonstrate convergent validity with the FIQR in measuring fibromyalgia disease severity
Lazor et al (2016)	Systematic Review	Searches were conducted of MEDLINE, Embase, PsycINFO, HAPI, and CINAHL, including studies that used at least one instrument to measure anxiety quantitatively in children or adolescents with cancer or undergoing HSCT.	Only three multi-item (including PROMIS Pediatric Anxiety Short Form) and two single-item instruments were identified as being reliable and valid among pediatric cancer or HSCT patients and would therefore be appropriate to measure anxiety in this population.

Bibliography

1. Institute of Medicine (IOM). Crossing the Quality Chasm: A New Health System for the 21st Century. *National Academy Press*, 2001.

2. Epstein R, et al. Measuring patient-centered communication in patient-physician consultations; theoretical and practical issues. *Social Science & Medicine* 2005; Vol 61(7): 1516-1528.

3. Junewicz A and Youngner S. Patient Satisfaction Surveys on a Scale of 1 to 10: Improving Health Care, or Leading it Astray? *Hastings Center Report* 2015; no 3: 43-51.

4. Agency for Healthcare Research and Quality. *What is Patient Experience?* www.ahrq.org (2021, accessed 8 March 2022).

5. Massachusetts Medical Society. *Patient Satisfaction Surveys.* https://catalyst.nejm.org/doi/full/10.1056/CAT.18.0288 (2018, accessed 8 March 2022).

6. Berkowitz B. The Patient Experience and Patient Satisfaction: Measurement of a Complex Dynamic. *The Online Journal of Issues in Nursing* 2016; Vol 21, No. 1, Manuscript 1.

7. Centers for Medicare & Medicaid Services. *Hospital-Based Purchasing Programs*, www.cms.gov (2021, accessed 8 February 2022).

8. Centers for Medicare & Medicaid Services. *Consumer Assessment of Healthcare Providers & Systems*, www.cms.gov (2022, accessed 8 February 2022).

9. Press Ganey®. *Press Ganey Receives CMS Approval to Administer Primary Care First Patient Experience of Care Surveys,* www.pressganey.com (2021, accessed 8 February 2022).

10. National Quality Forum. *Patient Reported Outcomes,* www.qualityforum.org (2022, accessed 8 February 2022).

11. National Institutes of Health. *Patient-Reported Outcomes Measurement Information System (PROMIS)*, www.commonfund.nih.gov/promis/index (2019, accessed 8 February 2022).

12. Hsiao C, Dymek C, Kim B, and Russell B. Advancing the use of patient-reported outcomes in practice: understanding challenges, opportunities, and the potential of health information technology. *Qual Life Res* 2019; Jun;28(6): 1575-1583.

13. Squitieri L, Bozic K, and Pusic A. The Role of Patient-Reported Outcomes Measures in Value-Based Payment Reform. *Value Health* 2017; Jun;20(6):834-836.

14. Guideway Care. *Patient Satisfaction: What It Is, Why it Matters & How to Improve It,* www.guidewaycare.com (2019, accessed 8 March 2022).

15. Murad M, Asi N, Alsawas M, Alahdab F. New Evidence Pyramid. *Evid Based Med* 2016; Aug;21(4) 125-127.

16. Lehrich B et al. Predictors of Patient Satisfaction in Spine Surgery: A Systematic Review. *World Neurosurg* 2021;Feb(146):e1160-1170.

17. National Library of Medicine. *Results by Year,* www.pubmed.ncbi.nlm.nih.gov (2022, accessed 12 February 2022).

18. Navarro S, Ochoa C, Chan E, Du S, Farias, A. Will Improvements in Patient Experience With Care Impact Clinical and Quality of Care Outcomes? A Systematic Review. *Medical Care* 2021;Sept;59(9) 843-856.

19. Hadlandsmyth K, Dindo L, St. Marie B, Roohina W, Embree J, Noiseux N, Tripp-Reimer T, Zimmerman B, Rakel B. PROMIS Instruments: Reliability and Validity in Veterans Following Orthopedic Surgery. *Eval Health Prof* 2020;Dec;43(4) 207-212.

20. Kroenke K et al. Incorporating PROMIS Symptom Measures into Primary Care Practice – a Randomized Controlled Trial. *J Gen Ingern Med* 2018;April33(8) 1245-52.

21. Wilford J, Osann K, Hseih S, Mond B, Nelson E, Wenzel L. Validation of PROMIS Emotional Distress Short Form Scales for Cervical Cancer. *Gynecol Oncol* 2018;Oct151(1) 111-116.

22. Merriwether E et al. Reliability and Construct Validity of the PROMIS Instruments in Women with Fibromyalgia. *Pain Medicine* 2017;18:1485-1495.

23. Lazor T, Tigelaar L, Pole J, De Souza C, Tomlinson D, Sung L. Instruments to Measure Anxiety in Children, Adolescents, and Young Adults with Cancer: A Systematic Review. *Support Care Cancer* 2017; June25: 2921-2931.

24. Institute of Medicine. *To Err is Human: Building a Safer Health System* https://pubmed.ncbi.nlm.nih.gov/25077248/ (2000, accessed 11 March 2022).

25. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st Century https://pubmed.ncbi.nlm.nih.gov/25057539/ (2001, accessed 11 March 2022).

26. Allen-Duck et al. Healthcare Quality: A Concept Analysis. *Nurs Forum* 2017 Oct:54(4): 377-386.

27. Centers for Medicare & Medicaid Services. *What is a Quality Measure?* www.cms.gov (2021, accessed 11 March 2022).

28. Virginia B et al. Psychometric properties of health related quality of life measures in acute coronary syndrome patients: a systematic review protocol. *JBI Database of Systematic Review and Implementation Reports* 2018 February:16(2): 316-323.

29. Millenson M, Macri J. *Timely Analysis of Immediate Health Policy Issues*. www.rwjf.org (2012, accessed 13 February 2022).

30. Centers for Medicare & Medicaid Services. BPCI Advanced Model Overview.www.cms.gov (2021, accessed 15 April 2022).

31. Edgman-Levitan S, Cleary PD. What Information Do Consumers Want and Need? *Health Affairs* 1996; 15(4): 42–56.

32. Chong K, Damiano P, Hays R. Psychometric Performance of the Consumer Assessment of Healthcare Providers and Systems (CAHPS) 4.0 Adult Health Plan Survey. *Primary Health Care* 2012; 2(4): 1-7.

33. Press Ganey Associates, Inc. *Medical Practice Survey Psychometrics Report* www.pressganey.com (2010, accessed 11 March 2022).

34. War, J, Sherbourne CD. The MOS 36-Item Short-Form Survey (SF-36). *Medical Care* 1992; 30(6): 473–83.

35. Centers for Medicare & Medicaid Services. *NHE Fact Sheet*, www.cms.gov (2022, accessed 13 February 2022).

36. Mehta SJ. Patient Satisfaction Reporting and Its Implications for Patient Care. *AMA J Ethics* 2015; 17(7): 616-621.

37. Casalino L, Elster A, Eisenberg A, Lewis E, Montgomery J, Ramos D. Will pay-for-performance and quality reporting affect health care disparities: *Health Affairs* 2007;26(3) 405-414.

38. Jerant A, Agnoli A, Franks P. Satisfaction with Health Care Among Prescription Opioid Recipients. *Journal of the American Board of Family Medicine* 2020;Jan;33(1) 34-41.

39. Liu L, Gauri D, Jindal R. The Role of Patient Satisfaction in Hospitals' Medicare Reimbursements. *Journal of Public Policy & Marketing* 2021;April;40(4):558-570.

40. Chang et al. Patients' global ratings of their health care are not associated with the technical quality of their care. *Ann Intern Med* 2006;May;144(9): 665-72.

41. Chen Q, Beal E, Okunrintemi V, Cerier E, Paredes A, Sun S, Olsen G, Pawlik T. The Association Between Patient Satisfaction and Patient-Reported Health Outcomes. Journal of Patient Experience 2019;Sept;201-209.

42. Mazor K, Clauser B, Field T, YoodR, Gurwitz J. A Demonstration of the Impact of Response Bias on the Results of Patient Satisfaction Surveys. *Health Sciences Research* 2002;November;37(5): 1403-1417.

43. Sotto-Santiago S, Slaven J, Rohr-Kirchgraber. (Dis)Incentivizing Patient Satisfaction Metrics: The Unintended Consequences of Institutional Bias. *Health Equity* 2019; April (3)1: 13-18.

44. Presson et al. Psychometric properties of the Press Ganey® outpatient medical practice survey. *Health and Quality of Life Outcomes* 2017 32(15)610-613.

45. Emergency Physicians Monthly. *Are Press Ganey Statistics Reliable?* www.epmonthly.com (2019, accessed 13 February 2022).

46. HCAHPS® Online. *CAHPS® Hospital Survey* www.hcahpsonline.org (2022, accessed 15 April 2022).

47. Fenton J, Jerant A, Bertakis K. The Cost of Satisfaction: A National Study of Patient Satisfaction, Health Care Utilization, Expenditures, and Mortality. *Arch Intern Med* 2012; 172(5): 405-411.

48. Manary M, Boulding W, Staelin R, Glickman S. The Patient Experience and Health Outcomes. *N Engl J Med* 2013; 368(3): 201-203.

49. PROMIS Health Organization. *What is PROMIS?* www.promishealth.org (2022; accessed 13 February 2022).

50. Geiger N. On Tying Medicare Reimbursement to Patient Satisfaction Surveys. *American Journal of Nursing* 2012;July;112(7):11.

51. Revere L, Langland-Orban B, Large J, Yang Y. Evaluating the robustness of the CMS Hospital Valute-Based Purchasing measurement system. Health Sciences Research 2021;January;56(3):464-473.

52. Massachusetts Medical Society. *Patient-Reported Outcome Measures: Current State and MMS Principles.* www.massmed.org/promis (2018; accessed 15 April 2022).