Pragmatic Measures
What They Are and Why We Need Them

Russell E. Glasgow, PhD, William T. Riley, PhD

Abstract: Pragmatic measures are important to facilitate implementation and dissemination, address stakeholder issues, and drive quality improvement. This paper proposes necessary and recommended criteria for pragmatic measures, provides examples of projects to develop and identify such measures, addresses potential concerns about these recommendations, and identifies areas for future research and application. Key criteria for pragmatic measures include importance to stakeholders in addition to researchers, low burden, broad applicability, sensitivity to change, and being actionable.

Examples of pragmatic measures are provided, including ones for different settings (e.g., primary care, hospital) and levels (e.g., individual, practitioner, setting) that illustrate approaches to produce broad-scale dissemination and the development of brief, standardized measures for use in pragmatic studies. There is an important need for pragmatic measures to facilitate pragmatic research, guide quality improvement, and inform progress on public health goals, but few examples are currently available. Development and evaluation of pragmatic measures and metrics would provide useful resources to advance science, policy, and practice.

(Am J Prev Med 2013;45(2):237–243)

Introduction

The gap between research and practice is well documented[1,2] and due at least in part to a perceived lack of relevance to the situations of stakeholders, who include both practitioners and policy/decision makers.[3,4] An evolving literature addresses pragmatic trials and study designs,[5–7] but there has been no parallel focus on related pragmatic measures. Pragmatic trials have been defined by the CONSORT group as "designed to answer the question of whether a program works under usual conditions (as compared to ideal conditions)," but there has been no definition of pragmatic measures. Berwick has argued that the evidence-based movement needs to adopt a pragmatic science approach,[8] and an important part of pragmatic science is having and using pragmatic measures to assess improvement. In brief, a pragmatic measure is one that has relevance to stakeholders and is feasible to use in most real-world settings to assess progress.

Perhaps the best known pragmatic measure is the Patient Health Questionnaire with nine items (PHQ-9), and the brief screening PHQ-2 (two items) derived from it. The PHQ-9 is a self-administered version of the PRIME-MD depression module for rapid assessment of depression severity in primary care.[9] An extensive literature supports the validity of the PHQ-9 in a range of populations, both for depression screening and for monitoring depression severity over time.[10] The two-item version of the PHQ also has been shown to be a valid measure of depression.[11,12] These studies support the use of the PHQ-2 as the initial screen for completing the remaining items of the PHQ-9, thereby reducing respondent burden for the vast majority of patients without substantial depressive symptomatology.

There are a number of routinely collected, pragmatic indices that can be extracted from the medical record. Weight and vital signs (temperature, heart rate, blood pressure) are almost universally collected at each office visit. For management of specific conditions, a range of indices such as lipids, hemoglobin A1c, and spirometry frequently are obtained in standard care. In contrast, indices of health behavior and health-related quality of life (e.g., physical activity, dietary intake, pain, physical and social function) are seldom collected. Even the most salient health behavior vital sign, smoking status, is screened for in less than two thirds of outpatient visits[13] and is seldom documented in a standardized manner.[14]

From the Division of Cancer Control and Population Sciences (Glasgow, Riley), National Cancer Institute, Rockville, Maryland

Address correspondence to: Russell E. Glasgow, PhD, 9609 Medical Center Drive, Room 4E440, Rockville, MD 20850. E-mail: glasgowre@mail.nih.gov.

0749-3797/$36.00
http://dx.doi.org/10.1016/j.amepre.2013.03.010

There are numerous uses for pragmatic measures, including in pragmatic studies[6,15]; comparative effectiveness research[16,17] (www.pcori.org/); public health surveillance; program monitoring and evaluation; and quality improvement.[8] The purposes of this paper are to (1) discuss the rationale for and propose criteria for pragmatic measures; (2) illustrate pragmatic measures through examples to identify, develop, and evaluate such measures; and (3) discuss implications, limitations of, and opportunities for pragmatic measures.

Recommended Criteria

There is a compelling need for relevant, pragmatic measures that can be used in real-world settings. However, there are no agreed upon criteria for what constitutes a pragmatic measure. Necessary and desired characteristics for a measure to be considered pragmatic are proposed below. Because there are currently so few pragmatic measures available, a longer list of desired characteristics was separated into those that are essential or "required" and those that are desirable but not necessary (Table 1). Within each category, all of the criteria were considered to be approximately equally important. Although the first two criteria seem especially axiomatic for pragmatic measures, all are important for a pragmatic measure to achieve its intended results.

Required or Strongly Recommended Criteria

Important to Stakeholders

Consistent with recent developments in patient-centered care, pragmatic designs, and comparative effectiveness research, the constructs assessed by pragmatic measures should be deemed important by stakeholder groups, including citizens/patients; practitioners (both public health and clinical); organizational decision makers; and policymakers. The interests of these stakeholder groups are overlapping but different, and each is important if a measure is to be adopted widely or used to make important decisions. Formative research strategies such as focus groups and survey methods can be used to obtain feedback on the outcomes of interest.

Table 1. Required and recommended criteria for pragmatic measures

Proposed criteria	Recommendations	Comments
Required criteria		
Important to stakeholders	Involve stakeholders from outset and on an ongoing basis	Different stakeholders have different priorities
Burden is low for both respondents and staff	Usually brief and inexpensive	Consider both time and cost
Actionable	Enhances use in busy, real-world settings. Can correct identified problems	Helps if easy to score, interpret and are related to evidence-based interventions
Sensitive to change (needs to be reliable over time and valid)	Can be used longitudinally to track progress and detect intervention effects	For use in intervention studies and quality improvement
Additional recommended criteria		
Broadly applicable	Feasible so almost all can complete, little missing data	Can be used to address equity issues, compare subgroups and settings
Use for benchmark; has norms to interpret or addresses public health goals	Useful across settings and populations; publicly available	Such as Healthy People 2020, IOM, HEDIS, or NCQA criteria or guidelines, such as PROMIS metrics
Unlikely to cause harm	Minimal unintended consequences	Do not want to create liability situations, should not collect measure if not prepared to address results
Psychometrically strong	Some criteria such as internal consistency not as critical for pragmatic measures	Helpful, but sometimes the "perfect is the enemy of the good"
Related to theory or model	Can advance understanding and interpretation of results	Helps to advance scientific understanding

HEDIS, Health Plan Employer Data and Information Set; NCQA, National Committee for Quality Assurance; PROMIS, Patient-Reported Outcomes Measurement Information System

Low Respondent and Staff Burden

To be incorporated into routine practice, pragmatic measures must require minimal time and effort to complete, easily integrate into local practice patterns, and be scored readily. Each measure should take no longer than 2 minutes for the typical respondent to complete, and a total of 10–15 minutes for an entire battery of pragmatic measures. Burden may be greater for low literacy or physically disabled patients who may require assistance. Modality and user-friendliness also must be considered in assessing burden.

Paper-based measures, although universally implementable, are at a distinct disadvantage for reducing burden because they must be administered, collected, scored, and entered manually. In contrast, computerized administration, scoring, and entry into the organizational or health record minimizes staff burden. Approaches—such as item response theory and computer adaptive testing (CAT), can be used to require the fewest number of items necessary to provide adequate precision (reliability) to measure a unidimensional construct. CAT simulations typically achieve precise estimates of much larger item banks in only four to six items.[18]

Actionable

Regardless of the research value of a given construct, to be pragmatic it must be of value to practitioners, who must respond to the results produced. The organization should have defined strategies or guidelines for responding to each construct measured. For example, clinicians know how to respond to blood pressure data based on widely accepted standards of practice.[19] Providing practitioners with data to which they are unsure how to respond or are unable to respond increases the likelihood of errors and liability. A pragmatic measure should have clear evidence-based interventions or policies associated with it. Consistent use is likely enhanced when the results have clear and immediate value to those administering them.

Sensitive to Change

Not all measures are intended to monitor and measure outcomes over time. Some are primarily for screening and diagnostic purposes and may not be sufficiently sensitive to change. For example, the Alcohol Use Disorders Identification Test–Hepatitis C (AUDIT-C) is a valid and brief screening measure for likely health and safety concerns from alcohol consumption, but the "past year" reporting period makes it insensitive to change over shorter time periods.[20]

Pragmatic measures generally should have relatively short reporting periods (past day, week, or month) and exhibit large effect sizes from interventions. Reliability and sensitivity to change are closely linked in this regard. A less reliable measure may show large changes over time but fail to identify the true change signal among the noise generated from its poor reliability. Greater reliability produces larger effect sizes, and is a core measurement attribute for determination of reliable change or minimally important difference.

Desirable but Not Necessary Criteria

Broadly Applicable

Pragmatic measures should be broadly available and able to be completed by almost all relevant people or settings. Data are needed from all segments of the population to assess trends and progress on key public health issues such as health inequities, and to produce generalizable findings. This produces data collection challenges, especially with survey participation rates decreasing and evolving digital-divide challenges from the types of media available to various subgroups.[21] To be broadly applicable, pragmatic measures should be easily read and understood by those with lower literacy and numeracy levels; culturally sensitive and appropriate for use regardless of gender, age, or racial/ethnic group; and applicable across disease conditions (e.g., not disease-specific).[22]

Serve As a Benchmark

Ideally, a pragmatic measure should be comparable to the effects observed in controlled studies. For traditional research, this is often achieved by utilizing the measures from published studies, but this approach often violates the respondent burden criteria above because controlled studies are less concerned about respondent burden. In pragmatic studies, benchmarking is often achieved by selecting the least burdensome measure from among the array of outcomes in a well-controlled efficacy trial. With item response theory methodology, one can extract a small number of items from a larger item bank, or allow a computer algorithm to select the items, and use these extracted items to estimate the score, thus providing direct comparisons to definitive trials with greatly reduced burden. For public health applications, benchmarking can refer to the pragmatic measure being used to measure progress toward key objectives, such as those in *Healthy People 2020* or reductions in health disparities.

Unlikely to Cause Harm

Lessons from the past decade of healthcare research are that well-meaning actions can produce unintended consequences and harm. The Action to Control Cardiovascular Risk in Diabetes (ACCORD) trial found that efforts to produce tight glucose control produced more deaths than usual care[23]; and some cancer treatment agents have been found to produce cardiotoxic effects.[24] This means that those who will be responsible for acting on the results have the capacity to address the resulting issues.

For example, the U.S. Preventive Services Task Force (USPSTF) balances the risks and benefits of various screening measures and in some cases (e.g., carotid artery stenosis) determines that the risks of some screenings outweigh their potential benefits (www.uspreventiveservicestaskforce.org/uspstf07/methods/benefit.htm). There is also potential for some items that assess sensitive subjects (e.g., sexual activity, substance use) to create animosity between consumers and health organizations or authorities. This is especially the case in situations where there is limited trust.

Psychometrically Strong

Like all measures, pragmatic measures should have good reliability and validity. Internal consistency is the most common form of reliability reported in the literature, in part because it is the easiest type of reliability to evaluate. However, it is heavily influenced by the number of items, which negatively affects respondent burden. Other forms of reliability are more central for pragmatic measures. Single-item measures, often used in survey methodology,[25] are not amenable to internal consistency reliability, but their test–retest reliability can be assessed. For observational scales, inter-rater reliability is preferred.

For pragmatic measures, content validity is important and can often be established with item development methodology that includes iterative feedback from respondents and stakeholders via focus groups, cognitive testing, and other qualitative methods.[26] Concurrent validity based on alternative measures of the same construct, especially alternative measures using a different mode (e.g., self-report versus interview), is valuable. Perhaps most important, predictive validity data often provide the basis for the actions recommended to practitioners.

Related to Theory or Model

Many have argued that nothing is so practical as a good theory. The current authors view pragmatic measures that validly capture important elements of a theory or model, such as the Chronic Care Model[27] or PRECEDE-PROCEED,[28] as an added bonus. Requiring all pragmatic measures to have a theoretic basis would unnecessarily limit development, but where possible, having metrics that relate to theories or models can aid understanding. This is consistent with U.S. Food and Drug Administration recommendations to develop and select measures based on a conceptual framework of how the measures interrelate, especially over time.[29]

Examples of Pragmatic Measures

Brief Patient Reports of Health Behaviors

An ongoing project illustrates the integration of pragmatic models, methods and measures.[30] The Evidence Integration Triangle[31] helps to conceptualize the necessary evidence-based interventions, pragmatic measures, and participatory processes needed to increase the frequency and quality with which health behavior and psychosocial issues are addressed in primary care and public health. In a recent application,[32] evidence-based interventions were taken from literature reviews and USPSTF recommendations (www.uspreventiveservicestaskforce.org) for primary care–based health behavior counseling and treatment of depression, anxiety, risky drinking, and substance abuse.

A series of 17 items assessing ten different health behaviors (e.g., tobacco use, physical activity, healthy eating, risky drinking) and psychosocial issues (e.g., depression, anxiety, stress, quality of life) were identified using iterative expert opinion and crowd-sourcing approaches. This involved integrating input from patients and patient advocates, practitioners, primary care researchers, and policy/decision makers. The resulting screening measures are very brief (generally one to two items); have national norms; and are practical to use in primary care and low-resource settings. These practical progress measures are designed to be used longitudinally to evaluate patient health status and progress on these behavioral and mental health issues (Table 2). An ongoing pragmatic implementation study is evaluating use of these items in diverse real-world settings.

The Patient-Reported Outcomes Measurement Information System

This system, known as PROMIS, is an NIH-supported project to develop item banks measuring constructs relevant across chronic diseases.[33,34] These banks, measuring various domains of physical, mental, and social health, are item response theory–developed banks that meet the criteria of low respondent burden (CAT or four- to eight-item short forms); high reliability; and increasing evidence for sensitivity to change and validity in various clinical samples.[35,36] PROMIS short forms also have been incorporated in the EPIC electronic health record (EHR) system, and outputs are HL7 compatible, making the results easily integrated into care settings with minimal burden to the clinical site. There are currently insufficient data to generate actionable clinical responses to the results of PROMIS banks, but these banks meet all of the other criteria for pragmatic measures.

The final two brief pragmatic measures below provide examples at the intervention staff and setting level, respectively. The Gawande checklist[37] of steps that surgical teams should take to prevent errors and adverse outcomes provides an innovative example of a nontraditional but very relevant pragmatic measure designed for immediate action (Table 2). Finally, a setting-level measure might be

Table 2. Example measures and their pragmatic strengths and limitations

Measure	Pragmatic strengths	Limitations
PHQ scales (depression and anxiety)	Low burden (2–4 or 9 items) Very actionable (based on DSM criteria) Sensitive to change—good data on this Broadly applicable, validated in diverse populations; has good norms	Could be liability issues, especially with suicide item if no follow-up Not related to theory (but is to DSM)
Brief Reports of Health Behaviors (My Own Health Report) Project	Strong stakeholder involvement and relevance; low burden (17 items to assess ten areas); actionable, with brief staff training; broadly applicable (English, Spanish, low literacy)	Not theory related Too brief to have strong psychometrics Some items have norms, others do not
PROMIS measures	Strong psychometrics, excellent benchmarking and norms, burden is low (if can implement computer adaptive testing) to moderate (if not)	Currently limited application if do not have strong HIT capabilities Unclear if actionable
Clinician/staff level Gawande surgical checklist	Very little burden, completely actionable, broadly applicable, very pragmatic, very relevant to stakeholders	Liability issues if do not reach 100% Not related to theory Psychometrics unknown
Setting level EHR data on percentage of eligible patients screened for colorectal cancer—and "dashboard" feedback to settings	Little additional burden after complex EHR coding algorithms created and verified; actionable at both individual and group (patient panel) levels; very relevant to stakeholders	Possible unintended consequences of "punishing" clinicians that serve hard-to-reach patients Not related to theory Psychometrics unknown or NA Moderate applicability—need strong HIT support and data-quality monitoring

EHR, electronic health record; HIT, health information technology; NA, not applicable; PHQ, Patient Health Questionnaire; PROMIS, Patient-Reported Outcomes Measurement Information System

the percentage and representativeness of all patients in a community health center or HMO who are recommended to be screened for colorectal cancer screening that have received this screening within a specified interval. A pragmatic version of this or related Health Plan Employer Data and Information Set (HEDIS)–type measures would be an automated EHR-based system that did not require added entry by either patients of staff, but that would produce actionable prompts both for individual patient encounters and also "dashboard" feedback for various teams on their performance across all their patients.

Discussion

The current paper presents a summary of the rationale for pragmatic measures, discusses a set of "required or strongly recommended" characteristics and a related set of desirable but less critical criteria, and provides brief examples of measures that meet many of these criteria. These examples demonstrate promising initial results, but there is still much to be done. The examples provided suggest that identification of a core set of pragmatic measures is possible; that it is feasible to get diverse stakeholder groups to come to consensus on pragmatic measures; and that citizens, health practitioners, policymakers, and researchers all recognize the need for such pragmatic measures to help focus appropriate action in the face of an overwhelming amount of data, but often little actionable information.[38]

These recommended criteria for pragmatic measures are not mandates that should be applied rigidly or unthinkingly, but rather a concrete starting point to initiate discussion in a complex area. They balance the need for basic core criteria, without letting "the perfect become the enemy of the good," or creating impractical standards for pragmatic measures, especially because there are so few currently available.[39] Reasonable people certainly could come to different conclusions about which criteria are necessary versus desired.

There is an important need for pragmatic measures and more than one approach to getting there. For example, Gawande[37] checklists as discussed above are a type of pragmatic measure and there is great potential for unobtrusive, automated pragmatic measures as well. There are also potential limitations, including that pragmatic measures may not have ideal psychometrics; they may be less sensitive to change or to differences between groups than longer, more burdensome measures

often used in clinical trials and efficacy research. Because pragmatic measures are often developed for broad, general use, they may not be a good fit for some local uses or specific conditions. One solution to this dilemma may be to use a core battery of validated pragmatic measures, and then add other locally or disease-specific measures as needed.

Implications and Opportunities

Pragmatic measures and pragmatic approaches focus on stakeholder perspectives. As illustrated above, stakeholders should be engaged meaningfully in the identification, development, and use of pragmatic measures. One of the high-priority areas for development of pragmatic measures is scales that differentiate meaningful engagement of patients, stakeholders, or communities from minimal surface acknowledgement of these principles.[39]

Comparative effectiveness research (CER)[17,40,41] is receiving increased attention, and the availability of good pragmatic measures is critical to CER. There is also a need for comparative research on potential pragmatic measures. Some initial pragmatic measures have been identified for patient-reported measures of population health issues in adult primary care, but many more are needed, especially in pediatrics, geriatrics, the cognitively impaired, and in specialty areas such as cancer, diabetes, asthma, and heart disease.

A particular challenge is identification of measures that assess relevant patient-centered issues for those having multiple morbidities, without requiring completion of lengthy and burdensome compilations of questions for individual diseases. Generic measures such as the PROMIS banks designed to measure outcomes across an array of chronic diseases will help address this challenge, but further research on the most efficient and effective strategies for developing pragmatic measures that are broadly applicable across age and disease groups should receive high priority.

Future personalized pragmatic assessment batteries will use methods such as computer adaptive testing. Such automated approaches allow reduction in the number of items and make item administration, scoring, and feedback to both consumers and staff immediate and actionable. Unfortunately, with the exception of a few academic centers and some integrated healthcare systems, the potential for automated and integrated pragmatic measure administration and scoring has not yet been realized. The primary challenges are not the technology, interoperability, or Health Information Portability and Accountability Act/security issues, but pragmatic and patient-centered measures receiving sufficient priority from decision makers and information technology staff. On the methodologic side, approaches such as item response theory work reasonably well to reduce the number of items for unidimensional constructs, but many of the pressing issues in health care and population health are complex and multidimensional, and multidimensional item response theory approaches need to be further advanced.[42]

Although there are at present only a limited number of measures that meet the current proposed pragmatic criteria, there are a sufficient number of eligible measures for many issues to test and refine these measures now while developing even better ones. The authors hope that these criteria serve as guidelines for selecting the most appropriate measure for any given pragmatic purpose, and help to identify additional research gaps that need to be filled for a measure to be optimized for pragmatic research and practice.

The opinions expressed are those of the authors and do not necessarily reflect those of the National Cancer Institute.

No financial disclosures were reported by the authors of this paper.

References

1. Glasgow RE, Vinson C, Chambers D, Khoury MJ, Kaplan RM, Hunter C. National Institutes of Health approaches to dissemination and implementation science: current and future directions. Am J Public Health 2012;102(7):1274–81.
2. Brownson RC, Colditz GA, Proctor EK. Dissemination and implementation research in health: translating science to practice, 1st ed. New York: Oxford University Press, 2012.
3. Rothwell PM. External validity of randomised controlled trials: to whom do the results of this trial apply? Lancet 2005;365:82–93.
4. Kessler R, Glasgow RE. A proposal to speed translation of healthcare research into practice: dramatic change is needed. Am J Prev Med 2011;40(6):637–44.
5. Thorpe KE, Zwarenstein M, Oxman AD, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. CMAJ 2009;180(10):E47–E57.
6. Tunis SR, Stryer DB, Clancey CM. Practical clinical trials: Increasing the value of clinical research for decision making in clinical and health policy. JAMA 2003;290:1624–32.
7. Zwarenstein M, Treweek S. What kind of randomised trials do patients and clinicians need? Evid Based Med 2009;14(4):101–3.
8. Berwick DM. Broadening the view of evidence-based medicine. Qual Saf Health Care 2005;14(5):315–6.
9. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001;16(9):606–13.
10. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. CMAJ 2012;184(3):E191–E196.
11. Kroenke K, Spitzer RL, Williams JB. The Patient Health Questionnaire-2: validity of a two-item depression screener. Med Care 2003;41(11):1284–92.
12. Arroll B, Goodyear-Smith F, Crengle S, et al. Validation of PHQ-2 and PHQ-9 to screen for major depression in the primary care population. Ann Fam Med 2010;8(4):348–53.

13. Jamal A, Dube SR, Malarcher AM, Shaw L, Engstrom MC. Tobacco use screening and counseling during physician office visits among adults—National Ambulatory Medical Care Survey and National Health Interview Survey, U.S., 2005-2009. MMWR Morb Mortal Wkly Rep 2012;61(S):38–45.
14. Glasgow RE, Kaplan RM, Ockene JK, Fisher EB, Emmons KM. Patient-reported measures of psychosocial issues and health behavior should be added to electronic health records. Health Aff (Millwood) 2012; 31(3):497–504.
15. Glasgow RE, Magid DJ, Beck A, Ritzwoller D, Estabrooks PA. Practical clinical trials for translating research to practice: design and measurement recommendations. Med Care 2005;43(6):551–7; PMID 15908849.
16. IOM. Learning what works: infrastructure required for comparative effectiveness research: workshop summary. www.iom.edu/Reports/2011/Learning-What-Works-Infrastructure-Required-for-Comparative-Effectiveness-Research.aspx. 2011.
17. Selby JV, Beal AC, Frank L. The Patient-Centered Outcomes Research Institute (PCORI) national priorities for research and initial research agenda. JAMA 2012;307(15):1583–4.
18. Cook KF, O'Malley KJ, Roddey TS. Dynamic assessment of health outcomes: time to let the CAT out of the bag? Health Serv Res 2005;40 (5 Pt 2):1694–711.
19. DHHS. Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure (JNC 7). www.nhlbi.nih.gov/guidelines/hypertension/. 2004. NIH publication no. 04-5230.
20. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol Use Disorders Identification Test. Arch Intern Med 1998;158(16):1789–95.
21. Fortney JC, Burgess JF Jr, Bosworth HB, Booth BM, Kaboli PJ. A re-conceptualization of access for 21st century healthcare. J Gen Intern Med 2011;26(S2):639–47.
22. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61(2):102–9.
23. Ismail-Beigi F, Craven T, Banerji MA, et al. Effect of intensive treatment of hyperglycaemia on microvascular outcomes in type 2 diabetes: an analysis of the ACCORD randomised trial. Lancet 2010;376(9739):419–30.
24. Yusuf SW, Ilias-Khan NA, Durand JB. Chemotherapy-induced cardiomyopathy. Expert Rev Cardiovasc Ther 2011;9(2):231–43.
25. Stover PJ, Harlan WR, Hammond JA, Hendershot T, Hamilton CM. PhenX: a toolkit for interdisciplinary genetics research. Curr Opin Lipidol 2010;21(2):136–40.
26. Brod M, Tesler LE, Christensen TL. Qualitative research and content validity: developing best practices based on science and experience. Qual Life Res 2009;18(9):1263–78.
27. Wagner EH, Glasgow RE, Davis C, et al. Quality improvement in chronic illness care: a collaborative approach. Jt Comm J Qual Improv 2001;27(2):63–80.
28. Green LW, Kreuter MW. Health program planning: an educational and ecological approach. 4th ed. New York: Mayfield Publishing Company, 2005.
29. Rothman ML, Beltran P, Cappelleri JC, Lipscomb J, Teschendorf B. Patient-reported outcomes: conceptual issues. Value Health 2007; 10(S2):S66–S75.
30. Glasgow RE. What does it mean to be pragmatic? Pragmatic methods, measures and models to facilitate research translation. Health Educ Behav 2013:In press.
31. Glasgow RE, Green LW, Taylor MV, Stange KC. An evidence integration triangle for aligning science with policy and practice. Am J Prev Med 2012;42(6):646–54.
32. Estabrooks PA, Boyle M, Emmons KM, et al. Harmonized patient-reported data elements in the electronic health record: supporting meaningful use by primary care action on health behaviors and key psychosocial factors. J Am Med Inform Assoc 2012;19(4): 575–82.
33. Cella D, Yount S, Rothrock N, et al. on behalf of the PROMIS Cooperative Group. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. Med Care 2007;45(5S1): S3–S11.
34. Cella D, Riley W, Stone A, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. J Clin Epidemiol 2010;63(11):1179–94.
35. Fries J, Rose M, Krishnan E. The PROMIS of better outcome assessment: responsiveness, floor and ceiling effects, and Internet administration. J Rheumatol 2011;38(8):1759–64.
36. Cook KF, Bamer AM, Roddey TS, Kraft GH, Kim J, Amtmann DA. PROMIS fatigue short form for use by individuals who have multiple sclerosis. Qual Life Res 2012;21(6):1021–30.
37. Gawande A. The checklist manifesto: how to get things right. New York: Metropolitan Books. Henry Holt & Company, 2009.
38. Rebitzer JB, Rege M, Shepard C. Influence, information overload, and information technology in health care. Adv Health Econ Health Serv Res 2008;19:43–69.
39. Rabin BA, Purcell P, Naveed S, et al. Advancing the application, quality and harmonization of implementation science measures. Implement Sci 2012;7(1):119.
40. Glasgow RE, Steiner JF. Comparative effectiveness research to accelerate translation: recommendations for an emerging field of science. Brownson RC, Colditz G, Proctor E, eds. Dissemination and implementation research in health: translating science and practice. New York: Oxford University Press, 2012:72–93.
41. IOM. Redesigning the clinical effectiveness research paradigm: innovation and practice-based approaches: workshop summary. Washington DC: National Academies Press, 2010.
42. Gibbons RD, Rush AJ, Immekus JC. On the psychometric validity of the domains of the PDSQ: an illustration of the bi-factor item response theory model. J Psychiatr Res 2009;43(4):401–10.