

## The Clinimetrics of HowsYourHealth Survey Items

HowsYourHealth is rooted in patient-reported measurement of “what matters” to them. This paper describes what we know about patient-reported measurement and HowsYourHealth measures. A companion paper describes emerging methods to mitigate problems of patient-reported measures for clinical practice comparisons.

### **BACKGROUND**

Under the auspices of the Institute of Medicine our group and A.R. Feinstein, (who literally wrote the book on “Clinimetrics”), laid out the principles for patient reported measurement. (The specific topic was the measurement of health status in clinical practice).

*Feinstein AR. Benefits and Obstacles for Development of Health Status Assessment Measures in Clinical Settings. Medical Care. 30(5):MS50-MS56, May 1992*

Health status measures offer scientific, humanistic, and economic benefits for clinical medicine. The main problem is the many intellectual and pragmatic obstacles that block successful development of these measures. The inventory of such problems includes the following: definition of health; medical components of health status; who makes the choice about what to include and emphasize; attributes to be rated by patients or clinicians; indexes to be created from those attributes (including mega-variable indexes, global indexes, and oligo-category indexes); different measurements of the same entity; and clinimetric problems in nonclinimetric models. Several solutions to these multiple, complex difficulties can be offered: 1) ensure that a specific purpose, focus, and setting are clearly identified for every health status index; 2) recognize that an off-the-shelf index with high statistical scores for so-called reliability and validity may not be pertinent for a given current situation in which it is to be used; 3) avoid indexes involving combinations of excessive numbers of variables; 4) let patients choose the most significant foci and components of the indexes; 5) seek greater communication and understanding among multidisciplinary collaborators, who may have many differences in the ethos and goals with which they approach the construction of health status indexes; and 6) recognize that the construction of suitable health status indexes is an outstanding challenge in basic scientific inquiry, and, in this spirit, support major alterations in the current ideology for conceptualization and funding of what is basic science in clinical medicine.

Our report was the following:

*Wasson JH, Keller A, Rubenstein LV, Hays RD, Nelson EC, Johnson D, and the Dartmouth Primary Care COOP. Benefits and obstacles of health status assessment in ambulatory settings: The clinician's point of view. Medical Care 1992;30(5 suppl):MS42-MS49.*

In the past decade physicians have identified the need to expand patient assessment to include global function and quality of life. During the same period, the busy clinic has evolved into the location where this assessment seems most appropriate. Integrating functional health assessment into a busy clinical practice is difficult because the necessary steps require time, thought, recording, and follow-up. Attention to the office ecosystem is very important before any patient care management method is introduced. The clinician must transform the results of health status screening into a specific functional diagnosis. The clinician has to understand the sensitivity, specificity, and predictive value of the measure for a preliminary diagnosis to be made. Often, additional measurements must be taken to establish a specific diagnosis. These steps encompass assessment linkage. Once the specific cause for the dysfunction is recognized, the clinician then has to determine the need for special resources. This is called the resource linkage. By following the steps outlined in this paper, the clinician should be able to overcome many obstacles for functional health status assessment in busy ambulatory settings.

## **PUBLISHED ARTICLES FOR ADULT HOWSYOURHEALTH**

Over several decades our group has published articles about the clinimetric properties of many of the measures used in HowsYourHealth. The following examples refer to items included in the Adult HowsYourHealth surveys.

*Nelson EC, Landgraf JM, Hays RD, Wasson JH, Kirk JW. The functional status of patients: How can it be measured in physicians' offices? Med Care 1990;28(12):1111-1126*

Physicians wishing to maintain the functional capacity of their patients often need, but usually do not have, practical measures of function. The Dartmouth COOP, a primary care research network, developed nine pictorial Charts to efficiently measure patient function in busy office practice. Each Chart has a five-point scale, is illustrated, and can be self-administered or administered by office staff. The Charts are used to measure the patients' overall functional health just as Snellen Charts are used to measure vision. Studies to assess the Charts' reliability, validity, acceptability and clinical utility were conducted on over 2,000 patients in four diverse clinical settings. Results show that the Charts are both reliable and valid. One-hour test-retest intraclass correlations for elderly patients ranged from 0.78 to 0.98 and from 0.73 to 0.98 for low income patients. The average Pearson product-moment correlation between Charts and previously validated measures of function was 0.61 and the Charts were as capable of detecting the association between disease and functioning as were longer, standard measures. Most clinicians and patients report that the Charts are easy to use and provide a valuable tool to measure overall function in busy office practice. For the 25% of patients in which the Charts uncovered new information, changes in clinical management were initiated for 40% of them. We conclude that the COOP Charts are practical, reliable, valid, sensitive to the effects of disease and useful for quickly measuring patient function.

Wasson JH, Jette AM, Anderson J, Johnson DJ, Nelson EC, Kilo CM. Routine, single-item screening to identify abusive relationships in women. *J Fam Practice* 2000; 49(11):1017-1024.

To develop the single-item screening tool primary care physicians and nurses examined several methods for detection of abuse identified in the published literature we developed 2 word-and-picture charts based on previously tested methods. Two groups of women who had recently experienced domestic abuse (total N=13) reviewed these charts, and their suggestions were incorporated. For our investigation, the word-and-picture Relationships Chart was used (Figure 1). To validate the Relationships Chart we administered it to 51 women volunteers in urban and rural domestic abuse support groups. The control group consisted of 48 randomly selected patients in 2 obstetric and gynecology clinics. These women were also asked to complete 41 items about spouse or partner abuse. The items were scored on a 5-point scale so the maximum score would be 205 and the minimum score 41. Women seeking help from a support group because of their current or previous involvement in an abusive relationship scored much worse than women in the control group on both the single-item chart and the multi-item score ( $P < .001$ ). Based on the chart, when women reported potential abuse “none or a little of the time,” their average score on the multi-item questionnaire was 62; when their response was “some of the time” their average score was 86; and when they reported “often or always” experiencing potential abuse during the previous 4 weeks their average score was 114. Eighty-eight of the women completed the Relationship Chart a second time within 10 days after the initial administration. The average test-retest correlation was 0.60; on the 5-point scale 88.4% of the responses stayed the same or shifted by only 1 scale point. Thus, the Relationship Chart had reasonable face and criterion validity and more than met minimum standards for reliability.

Wasson JH, Mackenzie TA, Hall M. Patients Use an Internet Technology to Report When Things Go Wrong. *2007 Quality and Safety in Health Care*; 2007;16:213-217

Background: As patients directly experience harm from adverse events, investigators have proposed patient-report to complement professional reporting of adverse events. Objective: To investigate how an automated health assessment system can be used to identify adverse events. Design and setting: Internet survey responses from April 2003 to April 2005 involving communities and clinical practices across the USA. Patients: 44 860 adults aged 19–69 years. Outcome: Patient perceptions of adverse events experienced during the previous year. Independent legal review was also used to estimate how many patient-reports were serious enough to be potentially compensable. Results: Although patient reports of possible adverse events was low (1.4%), the percentage of adverse events was eight times higher for patients with the greatest burden of illness than for those with the least (3.4% vs 0.4%). Two expert malpractice attorneys agreed that 9% of the adverse events seemed to be serious. Conclusions: Patients will use internet technology to report their perceptions of health-related adverse events. Some of the patient-reported events reported will be serious.

Wasson JH, Benjamin R, Johnson D, Moore LG, and Mackenzie T. Patient Use the Internet to Enter the Medical Home. *J.Amb.Care.Mgmt.* 2011; 34:38-46 (From the *Methods Section only*)

We nest patient-reported data into 4 constructs consistent with attributes of the PCMH described previously.

**PCMH Processes.** Single item measures for:

- personal **continuity**—a patient has a personal doctor or nurse,
- very easy **access** to medical care when needed,
- **efficiency**—the practice appears well organized and does not waste a patient's time,
- **coordination** of care if a patient has 2 or more doctors.

**The Style of PCMH Interaction** between a patient and the practice's physician and office staff. Style is based on the average of 4 measures of physician respectfulness, physician willingness to listen, staff courtesy, and staff helpfulness (Cronbach's alpha 0.88). ( Consumer Assessment of Health Providers and Systems).

**A PCMH Communication** composite based on the average of 3 measures for physician knowledge of the patient history, clarity of explanations, and quality of information about chronic diseases (Cronbach's alpha 0.81).

**Practice Awareness of Functional Limits.** Patient well-being has many determinants; an example is functional health. If truly "whole-person oriented" the medical home should be aware of more than disease and bioclinical measures. The Internet assessment includes patient function measures designed for the primary care practice setting. For each of 6 functional domains (limits in daily living or social activities or social support; impediments from pain or emotional problems; reduced physical function) patients note if their doctor or nurse is aware of the limit. The ratio of awareness is averaged into a score for patients having functional limits (Average correlation among measures:  $r = 0.64$ ; Cronbach's alpha = 0.93 for a combination of the 6 measures.).

We also provide a **Single Overarching Measure for PCMH** proposed by Berwick: the percentage of patients who strongly agree "they receive exactly the care they want and need exactly when and how they want and need it". Patient report using this measure is associated with the independent ratings of office function by the practice staff.

We examine the association of the overarching measure and constructs of PCMH with patient reports about PCMH desirable consequences that include:

**Patient Confidence with Self-management.** Although its conceptual value has been recognized for decades, patient confidence is now clinically accepted as a critical path for improved patient outcome. The single question used here—"how confident are you that you can manage and control your health problems or concerns?"—elicits a response to similar other measures for this concept.

**Benchmarks for Prevention** based on completion of mammography, bowel cancer and lipid screening in patients 50 years or older and Benchmarks for condition management that include control of blood pressure, cholesterol, and blood glucose when patients report a diagnoses of hypertension, cardiovascular disease, or diabetes. Patient Internet-based reports of benchmarks for condition management have been previously validated. For this study, the office staff in 6 practices audited 451 (84%) of 541 eligible medical records. Accuracy of patient self-report for the presence or absence of benchmark attainment was 96% for breast cancer screening (within 2 years), 94% for blood pressure control (within 6 months), 85% for lipid control (within 6 months), 72% to 92% for diabetic control (depending on method— within 6 months), and screening for bowel cancer was 76% (within 2 years for hemoccult test and 9 years for colonoscopy). Neither patient educational attainment nor financial status had a consistent effect on accuracy.

**Wellness Activities** based on self-report for healthy eating and risk avoidance, regular exercise, and not smoking; and

**Non-Use of an Emergency Department or Hospital** in the past year.

## **REAL-TIME, ONLINE FIELD TESTING OF HOWSYOURHEALTH ITEMS**

While publication has historically been academically important our group has found it even more important to be able to rapidly test items on-line for unforeseen situations such as: i) when literacy and language differ, ii) sponsors of HowsYourHealth need to customize questions for their particular circumstances, or iii) researchers/public policy experts/payers suggest that new measures must be included. The following examples illustrate how HowsYourHealth is used for On-Line testing.

**On-Line Testing for Reliability.** If the item is poorly worded or refers to a very vague concept, the same person may answer it differently at different times. This is called test-retest reliability. HowsYourHealth enables estimates of test-retest reliability. As an example, the concept of patients' report that "their medications may be making them ill" may initially seem a vague query. Therefore, at beginning of survey we asked patients if their medications might be making them ill. Many minutes later in the HowsYourHealth survey we asked again. Of those who said that their medication were not causing illness, test-retest agreement was 95% (of 2276 responses); for those who reported "yes" or "maybe" that their medications were causing illness agreement was 87% (or 549). A kappa statistic of agreement was 0.80 (very good).

**On-Line Testing for Validity.** Reliability does not imply validity. Even though patients consistently report that their medications may be causing illness, is their report valid? When an geriatrician and a clinical pharmacologist independently and in a blinded way reviewed medical record information, the reviewers agreed that the risk of an adverse clinical side-effects from the medications were several

times higher in those patients who said “yes” and “maybe” than in those who reported “no.” (Annual Review of Gerontology and Geriatrics, 1992;(12)109-125) In this case, the reviewers served as the referees of the “standard” against which patient report for medication problems was compared. We have also used medical record reports of procedures performed and test results to validate patient report of procedures and tests. (J.Amb.Care.Mgmt. 2011; 34:38-46)

Validation based on medical record review is laborious and is only applicable in a few circumstances. Therefore, On-Line HowsYourHealth testing often uses three easier methods to validate items.

i) The first and easiest is simply “face” validity...the item “looks like” it will measure what the clinician or patient cares about. Using the “custom options” section of HowsYourHealth any sponsor of HowsYourHealth can post an item and quickly obtain hundreds of responses. The results can be examined immediately against the respondent’s demographics, burden of illness, habits, etc. Although psychometrically oriented researchers may disparage face validation, for many clinicians who want to have specific items suitable for action in the ecology of their practices, “face” validity is often “good enough.”

ii) Another and only slightly more difficult validation approach is to place items in the HowsYourHealth survey that refer to external “standards” (that themselves may be internally or externally referential). An example included in HowsYourHealth is the Consumer Assessment of Health Providers and Systems (CAHPS). During the past three years items from CAHPS have been randomly distributed to respondents in four “bundles” (to reduce respondent burden). There are more than 6000 respondents per CAHPS bundle.

The following Table illustrates those CAHPS items most closely related to quality of care items in HowsYourHealth.

The CAHPS items that look most like those in HowsYourHealth seem to define attributes consistent with the intent of each listed HowsYourHealth domain: access, efficiency, continuity and information. In clinimetric terms, the correlations between the HowsYourHealth domains and the “look like” CAHPS items are higher ( $r = 0.4-0.5$ : convergent validity) than those in the “least look like” group ( $r = 0.03-0.21$ ; discriminant validity).

| <b>HowsYourHealth Items</b>   | <b>Looks Most Like in CAHPS</b>  | <b>Looks Least Like in CAHPS</b>   |
|---|--|--|
| <b>ACCESS:</b><br>How easy is it to get medical care when you need it?  | ... care you needed right away....   | ... doctor ask you ... sad, empty or depressed?  |
|   | ...get appointment as soon as you thought you needed...  | ... doctor talk about...things that worry your or cause stress?  |
|   | ...phone answer to medical question as soon as needed...   | ...doctor talk about healthy diet and eating habits?   |
| <b>EFFICIENCY:</b><br>When you visit your doctor's office, how often is it well organized and does not waste your time?                 | ... how often did you see your regular doctor within 15 minutes of your appointment time?                | ... doctor ask you ... sad, empty or depressed?  |
|   | ... how often did your doctor spend enough time with you?  | ... doctor talk about...things that worry your or cause stress?  |
|   | ... were clerks and receptionists as helpful as you thought they should be?                              | ...doctor talk about healthy diet and eating habits?   |
| <b>CONTINUITY:</b><br>Do you have one person you think of as your personal doctor or nurse?   | ... how often did your regular doctor seem to know the important information about your medical history? | ... did your regular doctor tell you there was more than one choice for your treatment or health care? |
|   | ... how often did your doctor spend enough time with you?  | ... did your regular doctor ask you which choice was best for you?                                     |
|   | ... how often did your regular doctor show respect for what you had to say?                              | ... doctor ask you ... sad, empty or depressed?  |
| <b>CHRONIC DISEASE INFORMATION:</b><br>In general, how would you rate the information or doctor or nurse gave you about these problems? | ... how often did your regular doctor explain things in a way that was easy to understand?               | ... doctor talk about...things that worry your or cause stress?  |
|   | ... how often did your doctor spend enough time with you?  | ... doctor ask you ... sad, empty or depressed?  |
|   | ... how often did your regular doctor seem to know the important information about your medical history? | ... doctor give help you needed to make changes to prevent illness?                                    |

[The branching logic enables many levels of validation. As an example, although the CAHPS questions about stress and depression are unrelated to the listed quality domains their convergent correlations are actually quite strong ( $r = 0.50$ ) when compared with a HowsYourHealth question that asks of persons with self-rated emotional problem if they believe that their doctor or nurse is aware of the problem.]

iii) A third validation approach within HowsYourHealth is based on the addition of open-ended questions. (We used this approach to validate questions in HowsYourHealth about patient reported harms from medical treatment described previously - Quality and Safety in Health Care; 2007:16:213-217).

Clinicians also commonly use open-ended items to ask patients why the practice is being rated poorly on particular items and how it might improve. For example, the clinician posts in the survey: “ some of our patients have reported that they have difficulty getting medical care when they need it. Do you have any suggestions about how we can make it easier for our patients to get medical care when they need it?”

## ***On-Line Testing for Resolution (Power) Of Measures***

Depending on the purpose of an assessment—from “just good enough” to “research worthy”—the number of items requested of patients can change. Office practices and patients generally appreciate assessment brevity linked to reports for action; researchers may favor multiple questions for a domain to decrease variance of the summated score and increase resolution (power) of the measure to detect differences.

Which level of resolution is “best?” Feinstein emphasized that the creators of measures must carefully articulate their measurement objectives, they must make clear how much they favor statistical psychometric properties of the instrument versus “action orientation”, and they must determine how comfortable the end users (patients and clinicians) are with combinations of variables in a measure versus responding to a single measure.

At least in the situation of CAHPS and HowsYourHealth, the good news is that the internet enables the inclusion of both approaches so that, to re-quote Feinstein, there can be “greater communication and understanding among multidisciplinary collaborators, who may have many differences in the ethos and goals with which they approach the construction of (health status) indexes”. As an example, for measures of access both a HowsYourHealth single question and a CAHPS multi-item index are highly correlated ( $r = 0.80$ ). With at least 30 responding patients none of 29 practices in the lower or upper tertile were misclassified to the opposite tertile by either measure. In this example a perceived difference in measures is not practically relevant.

## **“ACTION GROUPINGS” OF RESPONSES FOR PRACTICE IMPROVEMENT**

When HowsYourHealth is used, its primary emphasis is to serve the needs of a particular patient and a particular clinician by placing them both “on the same page of communication.” Many measures are reported by patients for this purpose and a few additional measures are collected to help clinicians “diagnose” the quality of their practice – (an hopefully take action to improve!) The patient reports about practice quality remain anonymous.

Over the years we have empirically discovered that clinicians prefer “action groupings” of quality metrics that they do not have to sort through many variables. Action Groupings have been greatly facilitated by analyses that show which measures are “drivers” of quality. While the clinician receives information about what is the matter and what matters to each patient, the clinical setting in which that clinician works has instantaneously available to it patient-reported performance for all of the measures listed below. The Figure below illustrates the strongest linkage (**in bold**) among HowsYourHealth measures within three

domains of patient report: experience, engagement, and desirable consequences.



Practice targets for improvement are easily identified based on these "action groupings", the individual items within the groupings, and additional custom questions that can be added to HowsYourHealth by a practice.

## Emerging Methods to Mitigate Problems of Patient-Reported Measures For Clinical Practice Evaluation

### **BACKGROUND**

Clinicians use patient-reported information every day to serve their patients' needs. When fluid, unstandardized inquiries by a clinician becomes standardized and documented, they can also be used to evaluate performance of clinicians and practices. The Table below, (taken from *Balanced Measures for Patient-Centered Care*. Jamb Care Mngmnt. 2009;32: 44-51) provides an illustrative categorization of standardized measures.

**Table 1.** Differences between patient-reported action and comparison measures

|                       | <b>Patient-reported measures leading to specific actions for individual patients</b>                            | <b>Patient-reported measures used to quantify specific actions for many patients (the "population")</b>  |
|-----------------------|---|--|
| Patient role          | Patient report is part of self-assessment and feedback  | Supplies data for research variables or "satisfaction" for payment (does not include patient self-assessment and feedback)                     |
| Information           | Limited to a few measures that are "just good enough" so that response burden is minimized for any patient      | Often includes many measures for precision of estimates; sampling strategies are designed to ensure comparable estimates                       |
| Information cycle     | Information for action by patient and clinician "today"   | Usually delayed aggregate summaries (often requires adjustment for confounders); payment and reputation may depend on comparisons              |
| Behavioral assumption | Screening for "what matters" to individual patients enhances collaborative care between patients and clinicians | Clinicians will "respond" to the data (but, by definition, summaries and report cards miss opportunities for individual patient interventions) |

Since HowsYourHealth has for many years used the internet to routinely obtain both types of measures and make the results available to clinicians and patients in real time, this paper is neither about the relative value of these two categories of measures nor how to gather and report these measures in clinical practice. Rather, this paper outlines three problems that result from the increasing availability of clinical evaluative measures. We also describe and methods offered by HowsYourHealth that can mitigate the problems. The three problems are:

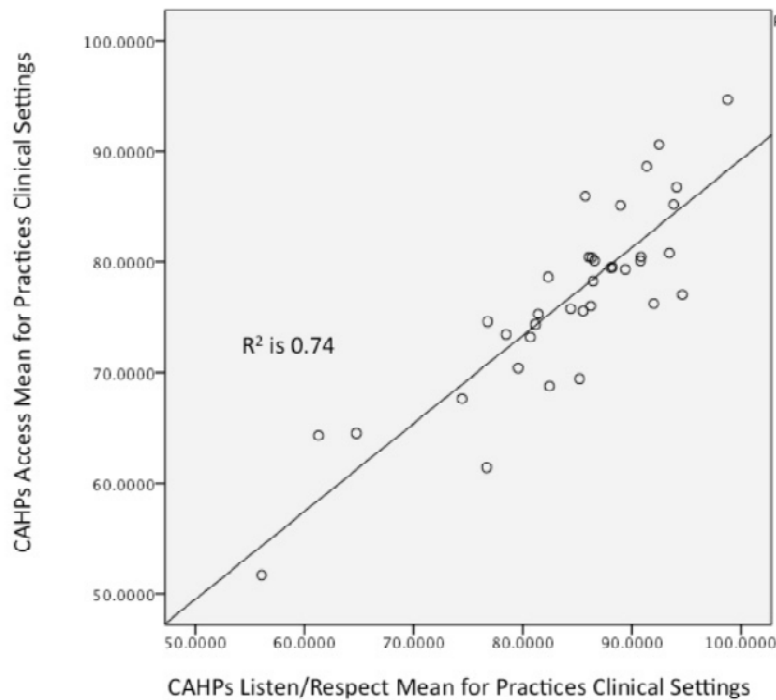
1. Correlation of measures
2. Known biases of sampling strategies
3. Cheating to gain pay for performance

## THREE PROBLEMS USING PATIENT-REPORTED EVALUATIVE MEASURES

### 1. Correlation

Correlation is a problem when a specific diagnosis is needed. In clinical practice, patients' report of their emotional state is correlated (about 0.5) with their social activities. Is the diagnosis "isolation" and the prescription "go out with your friends" or is the diagnosis "depression" and the prescription "depression management"? Similarly, in clinical practice patients' report of their access to care is correlated (about 0.5) with the degree they report their doctor listens to and respects them. Will better listening and respect compensate for poor access or vice versa?

Even modest correlations can frustrate clinical settings when the reports of many patients are aggregated into a "report card" of performance. The following Figure takes two rather different constructs of care that had enough clinimetric attributes to be included in CAHPS. The practical result of the correlation is that a setting's performance in one measure is usually quite similar to the rating on the other.



In summary, even though measures may have adequate "clinimetric" properties the correlations among many patient-reported measures can vitiate diagnostic value.

## 2. Biases of Sampling Methods

While having a “random sample” is a laudable objective for making sure that the results are representative of the population sampled, random sampling of a population in which only 40% will respond is, on its face, an effort to sample a biased sample. Yet “random sampling” of low responding populations is usually proposed for evaluating practices.

Moreover, one national organization has suggested that to obtain comparable samples – and adequate patient experiences of care - practices should select patients who have had multiple visits (e.g. at least 2 visits in the past 12 months). Yet, revisit interval has been shown to be 70% dependent on doctor preference and not underlying patient need. (JGIM 1999;14: 230-236).

Another biased sampling strategy is based on the identification of patients with “chronic diseases.” Again, it is well known that there are considerable practice style variations in diagnostic label that may be poorly related to the patient’s actual situation. (JAMA, 2011; 306: 461-568)

In summary, to impose sampling strategies on clinical practices might be justified if we did not already know that the strategies are biased and the population sampled is going to have a very low response rate. Furthermore, sampling imposes a very significant administrative burden on most clinical settings.

## 3. Cheating

When payment is linked to patient-reported performance, there is a possibility of cheating. Having an uncorruptible third party manage the measurement and reporting is the infallible option. However, third party evaluation of practices is very expensive and, at least in its current form, often fosters resentment.

### **EMERGING METHODS TO MITIGATE PROBLEMS OF PATIENT-REPORTED MEASURES FOR CLINICAL PRACTICE EVALUATION**

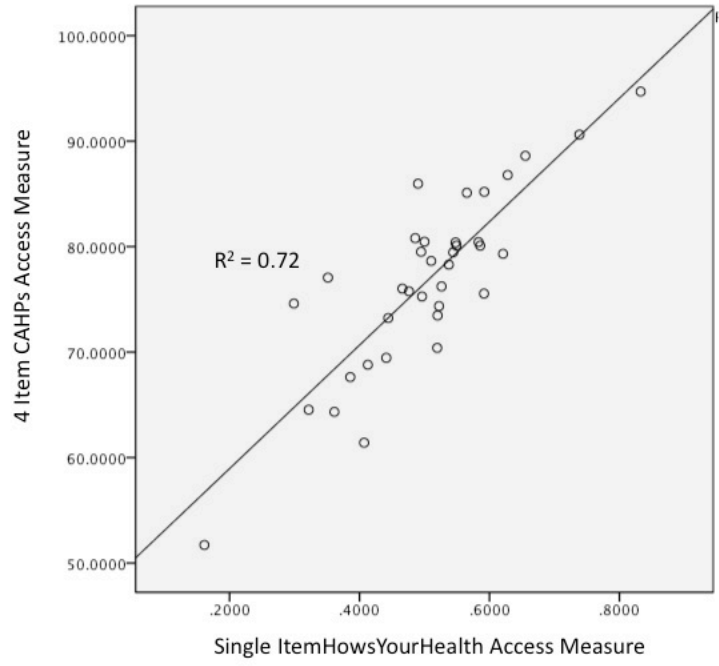
*“If it cannot be measured, it cannot be controlled” Lord Kelvin*

*“Not everything that counts, can be counted,  
and not everything that can be counted, counts” Albert Einstein*

These quotations illustrate a spectrum of opinion that underlies researchers’, clinicians’ and policymakers’ solutions to the three problems. There is no “right” choice for these problems, only partial solutions. The following represent our pragmatic suggestions based on more than a decade of use of HowsYourHealth in clinical practices, communities and employer settings.

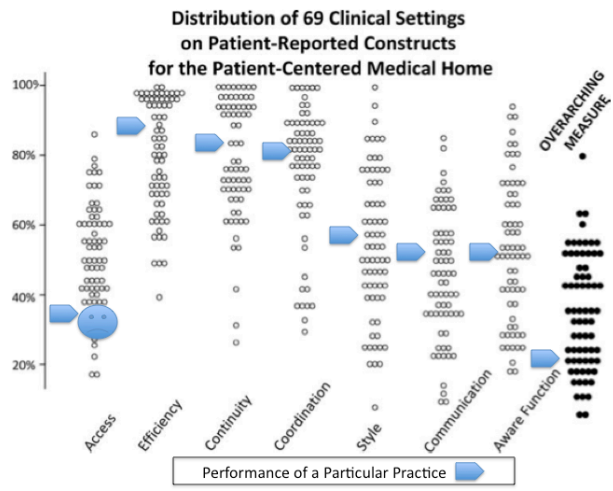
# 1. Correlation

Correlation offers both a problem and an opportunity. As an example of the opportunity consider the following Figure.



When at the level of a clinical setting a single item measure is so correlated with a complex measure, the single measure may be used and thereby reduce patient response burden. In fact, for paper-based measurement, only 10 items from HowsYourHealth can have great utility for practices that lack internet access. (Jamb Care Mngmnt. 2009;32: 56-71)

Neither a single or complex access is guaranteed to be diagnostically specific. However, as shown in the Figure on the right, when relative to other correlated measures the measure for access, as an example, is disproportionately worse than the other measures, the clinicians may want to focus further inquiries on access.



In a companion paper, “The Clinimetrics of HowsYourHealth Survey Items,” we describe that HowsYourHealth offers clinicians an optional use of open-ended items. For example, the clinician can post in the survey: “some of our patients have reported that they have difficulty getting medical care when they need it. Do you have any suggestions about how we can make it easier for our patients to get medical care when they need it?” In other words, the practice can use an automated specific method to determine what might be wrong and how to fix it. Diagnostic certainty is attained by additional inquiry; it is not at the mercy of additional preset measures that may not be diagnostically specific.

Correlation does not prove causation. Nonetheless, careful analyses of correlations across practices can provide insights about what measures seem to be linked more closely to other measures. As an example, the Figure below illustrates the strongest linkage (**in bold**) among HowsYourHealth measures within three domains of patient report: experience, engagement, and desirable consequences. While the clinician receives information about what is the matter and what matters to each patient, the clinical setting in which that clinician works has instantaneously available to it patient-reported performance for all of the measures listed below.



## 2. Biases

Regardless of whether a practice is managing sick, healthy, rich, poor, young old, male, or female patients the most desirable score on any HowsYourHealth quality measure is 100%. Only when a practice wants or needs to compare itself to other practices might it be of interested if there are large case mix differences or biases in the sample population.

With a push of the button HowsYourHealth is designed to automatically generate all patient-reported measures across three levels of patient burden of illness, financial status, gender and age.

## 3. Cheating

To maximize response rates patients do not have to remember passwords or passcodes; they do not have to surrender identifying information before completing HowsYourHealth. However, after completing the survey a practice can automatically ask patients to affix personal identifiers and send the summary to a clinician.

When the practice has made it clear that the information will be used to guide care, most patients (90%+) consent to send the information. At the same time the patient identifier and selected responses are placed in a registry. Once patient-identifying information is in the registry practices can provide a third party documentation about who completed the survey and the respondents relationship to all patients in the practice panel. As a specific example, a typical primary care clinician with a panel of 750 patients aged 50+ could attest to the authenticity of the respondents and document that during over each of three years  $\geq 25\%$  of the 750 unique patients used the technology - (automatically documented, with patient consent, by HowsYourHealth).

While not a fool-proof method to guard against cheating, a practice that uses HowsYourHealth as part of everyday work can soon have data for hundreds of patients. Collection of the data requires neither administrative machinations for patient selection, parallel documentation, and low-response survey distributions.