Development of Abbreviated Measures to Assess Patient Trust in a Physician, the Medical Profession, and Health Insurers.

Elizabeth Dugan PhD\(^1\), Felicia Trachtenberg PhD; and Mark Hall JD\(^1\)

\(^1\)The University of Massachusetts Medical School, Division of Geriatric Medicine, Worcester MA 01605

\(^2\)The New England Research Institutes, 9 Galen Street, Watertown MA 02472; and

\(^3\)Wake Forest University School of Medicine, Winston-Salem, NC 27157.

Key words: trust, satisfaction, patient-centered outcome measure

Corresponding Author:

Elizabeth Dugan, Ph.D.
Division of Geriatric Medicine
University of Massachusetts Medical School
377 Plantation Street, Suite 315
Worcester, MA 01605 USA

Elizabeth.dugan@umassmed.edu

Tel (508) 856-3493
ABSTRACT

Background: The objective of this study was to develop abbreviated scales to measure trust in a physician, an insurance provider, and in the medical profession.

Methods: Data from two samples were used. The first was a telephone survey of a national sample of English-speaking adults (N=1117) and the second was a telephone survey of English-speaking adults North Carolina residents (N=1024) who were members of an integrated healthcare delivery system. Data were analyzed to examine data completeness, scaling assumptions, internal consistency properties, and factor structure. Analyses were performed on the total sample. Additional analyses were performed on three key subgroups (determined by gender, race, and age) to further explore the performance of the abbreviated scales.

Results: Abbreviated measures (5-items) were developed for each of the three scales. Cronbach’s alpha of reliability was 0.87 for trust in a physician (test-retest reliability=0.71), 0.84 for trust in an insurer (test-retest reliability=0.73), and 0.77 for trust in the medical profession.

Conclusions: Assessment of data completeness, scale score dispersion characteristics, reliability and validity test results all provide evidence for the soundness of the abbreviated scales.
BACKGROUND

Trust is a key element of therapeutic relationships. Patient trust influences health status indirectly through continuity of care, adherence to treatment regimens, and the willingness to seek care [1-2]. Trust is hypothesized to influence health outcomes directly through the not yet well-understood mind-body pathway [3]. Biomedical researchers have paid increasing attention to trust as theoretical and measurement developments have occurred [4-24], driven in part by a concern about the potential negative influence of managed care on the doctor-patient relationship. Additional causes for concern about the doctor-patient relationship include the near daily release of conflicting health information about diet or lifestyle or medications, and well-publicized, yet rarely occurring, outrageous examples of malpractice and medical errors.

Our previous research reported on the development and validation of three instruments to measure trust in: 1) a specific doctor or healthcare provider, 2) an insurance provider, and 3) in the general medical system [16-18]. Interested readers are directed to a detailed description of a conceptual framework of trust that guided this work [3]. Briefly, the framework posits that patient trust involves patients’ vulnerability and their resulting reliance on and confidence in their physicians’ competence, motivation, honesty, and confidentiality. Yet despite the recent proliferation in research on patient trust, it is seldom a primary outcome, and is often just one of several peripheral areas of interest. Thus, the length of the original scales (10 and 11 items) limits the likelihood the scales will be widely used because of the frequent needs to minimize both respondent burden and research cost. Because of these concerns the present paper reports on the feasibility, factor structure, reliability and validity of abbreviated versions for the three instruments. Also, unlike the primary scales, which were constructed for use in general
populations, these are constructed with special attention to the performance of the scales in key subgroups (e.g., age; race; gender).

METHODS

Samples

Sample 1: National Sample. The first sample was selected by random-digit dialing. Inclusion criteria were age >21, with health insurance (n=151 excluded), visited a healthcare provider at least twice in the past two years (n=248 excluded), speak and understand English, and able to complete a telephone survey. Contacts with the 2172 potentially eligible adults resulted in the following dispositions: 1117 (51%) provided verbal informed consent and were interviewed; 571 (26%) refused; 484 (22%) were unable to participate (e.g., no answer after 15 callback attempts, too ill, or not able to speak and understand English). Complete data were obtained from 1064 and were used in analyses.

Sample 2: North Carolina HMO Sample. The second dataset was data from a random sample of enrollees in a managed care plan residing in North Carolina [19]. This sample included English-speaking adults aged 21 or older, who had been with the HMO for at least 2 years, and had made at least 2 visits to a primary care provider. Telephone contact was made with 1,908 (94.4%) resulting in the following dispositions: 319 (17%) were ineligible, 378 (20%) refused, and 1,211 (76%) provided verbal informed consent and agreed to participate. Complete data were obtained from 1,045 and were used in analyses. Two months later, a random subsample of 306 of these participants was resurveyed to assess test-retest reliability.

The telephone interviews averaged 35 minutes and were conducted by trained interviewers at the Survey Research Center of the University of South Carolina using computer assisted telephone interviewing. Verbal informed consent was obtained at the start of the
telephone interviews and the study protocols were reviewed and approved by the Wake Forest University Medical Center Institutional Review Board.

**Measurement**

The interviews collected information about patient and physician demographic characteristics; numerous items relating to trust in the subject’s personal physician, including Kao’s scale to measure physician trust in managed care members [10]; patient satisfaction with care [25]; single items to assess satisfaction in the doctor, insurer of interest and doctors in general; self-reported adherence to doctor recommendations; whether one would recommend the doctor or insurer to family and friends (Strongly agree to strongly disagree); past disputes with doctor [or, insurer] (yes, no); whether one had enough choice of doctor and insurer (yes, no); desire to switch doctor or insurer (yes, no); length of relationship with doctor and insurer of interest (number of years); self-reported physical and mental health (excellent, very good, good, fair or poor).

To reduce respondent burden the national sample was randomly divided; half were asked a battery of questions about insurance trust, and the other half were asked a battery of questions about trust in doctors in general. Complete data were obtained from 410 adults on the insurance trust items, and 502 adults for the medical system trust items. There were no statistically significant differences between the two samples on age, race, gender, or health status. All of the above measures were collected in the National sample, while the North Carolina sample did not collect information about trust in doctors in general, overall satisfaction, the willingness to recommend to family or friends, and the Kao trust scale.

**Statistical Approach**
The abbreviated scales were developed using data from the national sample and then validated with data from both samples. The items were drawn from the original 10 or 11 item scales, which were constructed using psychometric analyses focused on feasibility, factor structure, validity, and reliability. This same approach was used to develop the abbreviated scales. Feasibility analyses examined data completeness, floor and ceiling effects, and the dispersion of scores. The item response distributions were examined in the full sample and then by subgroup. Items were deemed not feasible and dropped from the scale if there was a high rate of missing data or responses were concentrated in one or two categories indicating a lack of power to discriminate.

The objective was to develop a 5-item scale. Items were selected so that the abbreviated form: reflected the content of the conceptual model (competence, honesty, fidelity, and global trust); contained both positive and negatively worded items; and performed well in the general population and key subgroups (African Americans, older adults, men and women). When two questions were otherwise equivalent, the one with the higher factor loading was chosen.

To ensure that items performed well in subgroups, scales for each subgroup were constructed and so that it could be determined to what extent there was overlap with the original scale. For each subgroup the item responses were examined and any items with low feasibility were dropped. Exploratory iterated principal components factor analysis with squared multiple correlations as initial communality estimates was performed to examine dimensionality and differences from the general population. Items with absolute factor loadings of <0.60 were identified, and subsequent items were dropped until 100% of the variance was explained.

Once the 5-items were selected, the shorter scale was then tested in the total sample and the subgroups as well. Correlations between the 5-item scale and the original 10 or 11 item scale
was considered, as well as correlations between the 5-item scale and key theoretically
determined concepts (e.g., Kao’s trust scale, general satisfaction with care, and number of years
with doctor or insurance; with desire to switch doctor or insurer, number of visits, satisfaction
with physician or insurance, willingness to recommend to friends, and whether doctors’
recommendations are always followed). A two-sample t-test was used for those variables with a
binary response format – prior dispute with doctor or insurance, having changed doctors, any or
enough choice is selecting doctor or insurance, having sought a 2nd opinion, and membership in
managed care.

Internal consistency was determined by Cronbach’s alpha of reliability. Test-retest
reliability could only be calculated using data from the North Carolina sample for the physician
trust and insurance trust 5-item scales.
RESULTS

A description of the two samples is reported in Table 1. The samples were similar in demographics; however the mean level of trust in insurer and physician in the North Carolina HMO sample was higher than that of the National sample.

5-item Scale to Measure Patient Trust in a Physician

Validity. Construct and concurrent validity were examined by correlation analyses and two-sample t-tests for items with binary responses. Table 2. reports the correlations for the 5-item scale in the National dataset and the North Carolina dataset. Trust in a physician was correlated with: satisfaction with the physician; would recommend to friends and family; general satisfaction with care; no desire to switch to another doctor; number of years under physician’s care; number of visits to physician. All correlations were significant at the p<0.001 level. Binary validation analyses showed that trust was associated with having enough choice in the selection of the physician, not having had a dispute with the physician, and not having sought a second opinion due to concerns about care. Similar associations were found in the subgroup analyses; however because the North Carolina sample had only a small percentage of elderly participants (3%), caution should be observed before drawing any conclusions about age or cohort differences. Trust generally decreased with poorer physical health (Wilcoxon tests, p=0.004). Trust also generally decreased with poorer mental health (Wilcoxon tests, p=0.012). Trust did not vary by education level or income.

Reliability. The 5-item scale had a Cronbach’s alpha of reliability in the National sample of 0.87, and in the North Carolina HMO sample it was 0.86. The 5-item scale had a lower internal consistency than either the original 10-item Wake Forest Scale (0.92) or the Kao scale (0.93). The internal consistency (Cronbach’s alpha of reliability) of the 5-item scale by subgroup
(National Sample; NC HMO Sample) was: African American or Black (0.83; 0.82); Men (0.84; 0.84), Women (0.88; 0.87); Older adult (aged 65 years or older) (0.82; 0.85). Test-retest reliability was 0.71 in the total sample. The test-retest reliability of the 5-item scale in the subgroups was as follows: Non-Hispanic White (0.687; n=320), African American or Black (0.849; n=41); Men (0.639; n=166), Women (0.754; n=205); Not elderly (0.724; n=364), Elderly (aged 65 years or older) (0.454; n=6).

**Subgroup findings.** African American adults were slightly less trusting of their doctor, but not significantly so. Women were more trusting on every single item; however the original scale showed only a marginally significant gender difference, and the 5-item scale did not show any gender difference. The very small gender difference seems to get obscured in the 5-item abbreviated scale.

**Summary:** There is evidence that a 5-item scale can be used to assess a patient’s trust in her/his doctor. The abbreviated scale is unidimensional. Responses are summed and scores are on a 5-25 scale, with higher values indicating more trust. Ceiling and floor effects were acceptable. Flesch-Kincaid reading level of the scale was 4.3 grade. The mean of the scale was 20.43, with a standard deviation of 3.13. The skewness was −1.05, and the reported kurtosis was 2.52.

**5-item Scale to Measure Trust in the Medical Profession**

**Validity.** Construct and concurrent validity were examined by correlation analyses and two-sample t-tests. Trust in the medical profession was correlated with the Kao scale (r=0.313), general satisfaction with care (r=0.482), and following doctor’s recommendations (r=0.440). The binary response validations showed that lower trust was related to having had a dispute with a physician, having changed doctors, and having sought a second opinion. Reported trust was
lower for those with poorer mental health (Wilcoxon tests, p=0.012). Trust is higher for people without a long-term medical illness, but not significantly so. Trust does not vary by education level or income.

**Reliability.** Cronbach’s alpha of reliability for the 5-item scale was 0.77. The abbreviated scale has a lower internal consistency than the original 10-item scale, but is acceptable. No test-retest reliability data were available because the questions about trust in the medical profession were not included in the NC HMO survey. The National data were analyzed to determine internal consistency for the 5-item scale in the subgroups. Cronbach’s alpha of reliability for each subgroup was: African American or Black (0.82); Men (0.80), Women (0.76); Elderly (0.72).

**Subgroup findings.** On average, older adults had more trust in the medical profession compared to younger adults (16.32 vs. 14.52, p<0.001). There was not a statistically significant difference in mean trust scores when race or gender was examined.

**Summary:** There is adequate evidence that the 5-item scale can be used to assess a patient’s trust in the medical profession. The 5-item scale is unidimensional in both the general population and in the defined subgroups. Responses are summed and scores are on a 5-25 scale, with higher values indicating more trust. Flesch-Kincaid reading grade level is 5.5. The mean of the Abbreviated Form was 14.97, with a standard deviation of 3.38. The skewness was –1.149, and the reported kurtosis was –0.330.

**5-Item Scale to Measure Trust in An Insurance Provider**

**Validity.** Construct and concurrent validity were examined by correlation analyses and two-sample t-tests. Trust was correlated with Kao’s trust scale (r=0.279), general satisfaction with care (r=0.465), satisfaction with insurance (r=0.646), desire to find another insurer (r=--
Binary response validations showed that trust was related to having any choice in selecting insurer, having enough choice in selecting insurers, having a past dispute with the insurer, and being in managed care. Adults with poorer mental health had significantly lower trust in their insurance provider than adults in better mental health. A similar, but not statistically significant trend was observed in physical health.

**Reliability.** The 5-item scale had a Cronbach’s alpha of reliability of 0.84 in the National sample, and 0.83 in the North Carolina HMO sample. Internal consistency as measured by Cronbach’s alpha for the subgroups (National Sample; NC HMO Sample) was: African American or Black (0.87; 0.77); Men (0.81; 0.83), Women (0.85; 0.83); Elderly (0.75; 0.86). Test-retest reliability of the trust in insurance provider 5-item scale was 0.729 in the general population. The abbreviated scale test-retest reliability in the subgroups was: Non-Hispanic White (0.730; n=301), African American or Black (0.759; n=44); Men (0.741; n=157), Women (0.713; n=198); Not elderly (0.728; n=349), Elderly (aged 65 years or older) (0.946; n=5).

**Subgroup findings.** Older adults (Aged 65+ years) trusted their insurance provider more than younger adults (mean trust score 18.11 vs. 16.13; p<0.001). There was not a statistically significant difference, but African American adults were slightly less trusting of their insurance provider than Non-Hispanic White adults (mean trust score: 16.10 vs. 16.64; p=0.374). Men were slightly more trusting than women, but not significantly so.

**Summary:** There is evidence that the 5-item scale can be used to assess a patient’s trust in an insurance provider. Responses are summed and scores are on a 5-25 scale, with higher values indicating more trust. The mean score was 16.57 with a 3.94 standard deviation. The skewness was −0.729, kurtosis 0.339. Flesch-Kincaid reading grade level is 7.7. The scale is one-dimensional in the general population and all subgroups and explains 100% of the variance.
DISCUSSION

The trust that a patient has for a provider or insurance provider is likely to be influenced by many factors. As healthcare financing pressures force the rapid evolution of the healthcare environment, research to understand these factors will grow in importance. The 5-item scales developed in this study will provide tools that may facilitate this research.

Development of the abbreviated scales was informed by our theoretical model and data driven. We sought to develop scales that provide sufficient measurement precision and breadth, while minimize burden and cost. The scales are brief, comprehensive and empirically validated tools. The scales require reading levels of grades 4.3 (physician trust), 5.5 (medical profession trust), and 7.5 (health insurer trust). Each 5-item scale had acceptable psychometric properties.

Several limitations of the current research should be noted. First, the results reported here are for telephone administration of the scales. The performance of the scales in other settings is yet unknown. Second, the interviews were only conducted with English-speaking adults, although subsequent research is currently in press reporting on Spanish translations of some of these items. Third, the small numbers of older adults limits the confidence that can be placed in the age-based comparisons. Further research on a larger sample of older adults, the most frequent users of healthcare is urgently needed. Research is also needed to determine the effectiveness of interventions to enhance the doctor-patient relationship, and whether such enhancements, by extension, will improve important patient outcomes.

CONCLUSIONS

Because they meet the demands of brevity, reliability, and validity and they may be widely used as measures of trust in a physician, the medical profession, and an insurance
provider. However, the long forms are recommended for research where trust is the main outcome of interest or for studies focusing on gender differences.

COMPETING INTERESTS

The authors have no competing interests.

AUTHORS CONTRIBUTIONS

Obtained research funding: MH, ED.

Research idea: ED, MH.

Statistical Analysis: FT.

Writing, revising, and final approval of manuscript: ED, MH, FT.

ACKNOWLEDGEMENTS

Research supported by the Robert Wood Johnson Foundation and the National Institute on Aging.
REFERENCES


22. Safran DG, Kosinski M, Tarlov AR, Rogers WH, Taira DH, Lieberman N, Ware JE. The


<table>
<thead>
<tr>
<th>Demographics</th>
<th>National Sample (N=1064)</th>
<th>North Carolina HMO Sample (N=1024)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>49.75 years</td>
<td>46.56 years</td>
</tr>
<tr>
<td>Female</td>
<td>68%</td>
<td>55%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>5%</td>
<td>2%</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>African American</td>
<td>10%</td>
<td>12%</td>
</tr>
<tr>
<td>American Indian or Alaska Native</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>2%</td>
<td>.5%</td>
</tr>
<tr>
<td>White</td>
<td>84%</td>
<td>86%</td>
</tr>
<tr>
<td>Other</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; High School</td>
<td>8%</td>
<td>6%</td>
</tr>
<tr>
<td>High School Graduate</td>
<td>28%</td>
<td>28%</td>
</tr>
<tr>
<td>&gt;High School</td>
<td>64%</td>
<td>66%</td>
</tr>
<tr>
<td>Physical Health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent or Very Good</td>
<td>57%</td>
<td>58%</td>
</tr>
<tr>
<td>Good</td>
<td>28%</td>
<td>33%</td>
</tr>
<tr>
<td>Fair or Poor</td>
<td>14%</td>
<td>10%</td>
</tr>
<tr>
<td>Mental Health</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Excellent or Very Good</td>
<td>72%</td>
<td>72%</td>
</tr>
<tr>
<td>Good</td>
<td>24%</td>
<td>25%</td>
</tr>
<tr>
<td>Fair or Poor</td>
<td>5%</td>
<td>3%</td>
</tr>
</tbody>
</table>

*Percentages may not equal 100% due to rounding.
Table 2. Comparison of Associations of Trust and Key Variables.

<table>
<thead>
<tr>
<th></th>
<th>Abbreviated Form National Sample</th>
<th>Abbreviated Form NC Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction with the physician</td>
<td>0.729</td>
<td>0.778</td>
</tr>
<tr>
<td>Would recommend physician</td>
<td>0.726</td>
<td>Na</td>
</tr>
<tr>
<td>General satisfaction with care</td>
<td>0.478</td>
<td>Na</td>
</tr>
<tr>
<td>Desire to change physicians</td>
<td>-0.660</td>
<td>-0.686</td>
</tr>
<tr>
<td>Number of years under dr.’s care</td>
<td>0.120</td>
<td>0.093</td>
</tr>
<tr>
<td>Number of visits to physician</td>
<td>0.127</td>
<td>0.150</td>
</tr>
</tbody>
</table>

Pearson correlation coefficients for continuous variables, Spearman correlation coefficients for categorical variables. All correlations significant at p<0.001.

Table 3. The Association of Trust in the Medical Profession Abbreviated Form and Key Variables.

<table>
<thead>
<tr>
<th></th>
<th>WFU 11-item Scale National Sample</th>
<th>Abbreviated Form National Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kao’s trust scale</td>
<td>0.306</td>
<td>0.313</td>
</tr>
<tr>
<td>General satisfaction</td>
<td>0.498</td>
<td>0.482</td>
</tr>
<tr>
<td>Follow doctor’s recommendations</td>
<td>0.449</td>
<td>0.440</td>
</tr>
<tr>
<td>Original WFU 11 item scale</td>
<td></td>
<td>0.957</td>
</tr>
</tbody>
</table>

Pearson correlation coefficients for continuous variables, Spearman correlation coefficients for categorical variables. All correlations significant at p<0.001.

Table 4. The Association of Insurance Trust Abbreviated Form and Key Variables.

<table>
<thead>
<tr>
<th></th>
<th>Abbreviated Form National Sample</th>
<th>Abbreviated Form NC Sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kao’s trust scale</td>
<td>0.279</td>
<td></td>
</tr>
<tr>
<td>General satisfaction</td>
<td>0.465</td>
<td></td>
</tr>
<tr>
<td>Satisfaction with insurer</td>
<td>0.646</td>
<td></td>
</tr>
<tr>
<td>Desire to switch insurers</td>
<td>-0.753</td>
<td>-0.589</td>
</tr>
<tr>
<td>Original WFU 10 item scale</td>
<td>0.952</td>
<td>0.948</td>
</tr>
</tbody>
</table>

Pearson correlation coefficients for continuous variables, Spearman correlation coefficients for dichotomous variables. All correlations significant at p<0.001.
Appendix

**Abbreviated measure of patient trust in a physician.**

*1. Sometimes Dr._[INSERT NAME OF DR.]__ cares more about what is convenient for (him/her) than about your medical needs.
2. Dr. _[INSERT NAME OF DR.]_ is extremely thorough and careful.
3. You completely trust Dr._[INSERT NAME OF DR.]’s decisions about which medical treatments are best for you.
4. Dr._[INSERT NAME OF DR.]__ is totally honest in telling you about all of the different treatment options available for your condition.
5. All in all, you have complete trust in Dr._[INSERT NAME OF DR.]_.

Response choices (coding) are: Strongly Agree (5), Agree (4), Neutral (3), Disagree (2), Strongly Disagree (1). Responses are summed (range 5-25) with higher scores indicating more trust.

*Negatively worded item is reverse coded.

**Abbreviated measure of patient trust in the medical profession.**

*1. Sometimes doctors care more about what is convenient for them than about their patients medical needs.
2. Doctors are extremely thorough and careful.
3. You completely trust doctors’ decisions about which medical treatments are best.
4. A doctor would never mislead you about anything [INTENTIONALLY].
5. All in all, you trust doctors completely.

Response choices (coding) are: Strongly Agree (5), Agree (4), Neutral (3), Disagree (2), Strongly Disagree (1). Responses are summed (range 5-25) with higher scores indicating more trust. *Negatively worded item is reverse coded.

**Abbreviated measure of trust in an insurer.**

*1. [INSERT INSURER’S NAME] Cares more about saving money than about getting you the treatment you need.
*2. You feel like you need to double check everything [INSERT INSURER’S NAME] does.
3. You believe [INSERT INSURER’S NAME] will pay for everything it is supposed to, even really expensive treatments.
4. If you have a question, you think [INSERT INSURER’S NAME] will give you a straight answer.
5. All in all, you have complete trust in [INSERT INSURER’S NAME].

Response choices (coding) are: Strongly Agree (5), Agree (4), Neutral (3), Disagree (2), Strongly Disagree (1). Responses are summed (range 5-25) with higher scores indicating more trust. *Negatively worded item is reverse coded.