

A Multifaceted Organizational Physician Assessment Program: Validity Evidence and Implications for the Use of Performance Data

Andrea N. Leep Hunderfund, MD, MHPE; Yoon Soo Park, PhD;
Frederic W. Hafferty, PhD; Kelly M. Nowicki, MA; Steven I. Altchuler, PhD, MD;
and Darcy A. Reed, MD, MPH

Abstract

Objective: To provide validity evidence for a multifaceted organizational program for assessing physician performance and evaluate the practical and psychometric consequences of 2 approaches to scoring (mean vs top box scores).

Participants and Methods: Participants included physicians with a predominantly outpatient practice in general internal medicine (n=95), neurology (n=99), and psychiatry (n=39) at Mayo Clinic from January 1, 2013, through December 31, 2014. Study measures included hire year, patient complaint and compliment rates, note-signing timeliness, cost per episode of care, and Likert-scaled surveys from patients, learners, and colleagues (scored using mean ratings and top box percentages).

Results: Physicians had a mean \pm SD of 0.32 ± 1.78 complaints and 0.12 ± 0.76 compliments per 100 outpatient visits. Most notes were signed on time (mean \pm SD, $96\% \pm 6.6\%$). Mean \pm SD cost was 0.56 ± 0.59 SDs above the institutional average. Mean \pm SD scores were 3.77 ± 0.25 on 4-point and 4.06 ± 0.31 to 4.94 ± 0.08 on 5-point Likert-scaled surveys. Mean \pm SD top box scores ranged from $18.6\% \pm 16.8\%$ to $90.7\% \pm 10.5\%$. Learner survey scores were positively associated with patient survey scores ($r=0.26$; $P=.003$) and negatively associated with years in practice ($r=-0.20$; $P=.02$).

Conclusion: This study provides validity evidence for 7 assessments commonly used by medical centers to measure physician performance and reports that top box scores amplify differences among high-performing physicians. These findings inform the most appropriate uses of physician performance data and provide practical guidance to organizations seeking to implement similar assessment programs or use existing performance data in more meaningful ways.

© 2017 Mayo Foundation for Medical Education and Research. Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>) ■ Mayo Clin Proc Inn Qual Out 2017;■(■):1-11

As a self-regulating profession, medicine is accountable for ensuring that physicians are competent in performing their clinical roles and responsibilities,^{1,2} and health care organizations play an important role in this process.^{3,4} Organizations collect physician performance data for many reasons (eg, ensuring physician competency, supporting health care choices by consumers, improving care quality, or satisfying regulatory or accreditation requirements)⁵ and can use performance data in various ways. For example, scores can be used to ensure that minimal performance expectations are met^{3,6} or to drive continuous improvement.⁷⁻⁹

Failure to meet performance expectations can lead directly to punitive consequences or can trigger additional investigations to determine whether a concern exists.¹⁰⁻¹³ Likewise, scores can be used primarily as formative feedback^{14,15} or for higher-stakes decisions (eg, promotion, employment, salary, privileging, and public transparency).^{9,10,16-19}

This panoply of purposes complicates the collection, distribution, analysis, and interpretation of physician performance data. Without a rigorous examination of the validity of their physician assessment programs, organizations risk using physician performance data in ways that are inappropriate or potentially



From the Department of Neurology (A.N.L.H.), Department of General Internal Medicine (F.W.H.), Department of Clinical Ethics (K.M.N.), Department of Psychiatry (S.I.A.), and Department of Primary Care Internal Medicine (D.A.R.), Mayo Clinic, Rochester, MN; and Medical Education, University of Illinois at Chicago (Y.S.P.).

detrimental.²⁰⁻²² Furthermore, the validity of commonly used physician performance measures may not be sufficient to support all intended purposes.

The use of physician performance data is further complicated by different approaches to scoring. For example, scores based on Likert-type ratings of performance can be reported as means (as often done for learner, multisource, or peer feedback surveys^{1,23}) or as the percentage of optimal ratings, also known as top box scores (as often done for patient satisfaction surveys²⁴⁻²⁶). The way in which scores are calculated affects their validity (eg, mean scores better represent the distribution of ratings, while top box scores may be more readily understood),²⁷⁻²⁹ yet this issue has not been extensively examined in the context of a multifaceted organizational physician performance assessment program.

For these reasons, we sought to (1) provide validity evidence for 7 different types of assessments commonly used to measure physician performance and (2) examine the practical and psychometric consequences of the 2 aforementioned approaches to scoring (mean vs top box scores).

PARTICIPANTS AND METHODS

This study was a retrospective analysis of deidentified physician clinical performance data collected via routine institutional practices and was considered exempt by the Mayo Clinic Institutional Review Board.

Study Participants and Setting

Study participants included all physicians with a predominantly outpatient practice in general internal medicine (GIM; n=95), neurology (n=99), and psychiatry (n=39) at Mayo Clinic in Rochester, Minnesota, from January 1, 2013, through December 31, 2014. Physicians within the 3 included specialties collectively completed more than 300,000 outpatient visits during the study time frame.

Measures

Physician performance measures included the following:

- Unsolicited patient complaints and compliments related to physician care, reported

as the number of complaints or compliments per 100 outpatient visits.

- Percentage of notes that were signed on time according to institutional policy (eg, clinical notes must be signed within 30 days).
- Mean internal cost per episode of care (ie, cost to the institution of providing tests and consults within a discrete period), reported as a z score relative to the institutional mean. Internal costs reflect utilization (eg, physicians who order more or more costly tests and consultations have higher internal costs) and are unrelated to prices or charges to patients/insurers. Internal costs are attributed to the physician with the highest evaluation and management billing code on the first day of a patient's evaluation. An episode of care comprises the subsequent days over which tests and consultations are performed.
- Patient satisfaction survey provided by Avatar International LLC³⁰ (9 items rated using a 5-point Likert scale ranging from 1 = strongly disagree to 5 = strongly agree, 0 = not applicable).
- Learner feedback surveys, ie, evaluation forms completed by residents and fellows (subsets of items from a total pool of 22 items rated using a 5-point Likert scale: 1 = needs improvement, 2-4 = average, 5 = top 10%, 0 = not applicable; free-text comments required for ratings of 1 or 5).
- Multisource feedback (MSF) surveys for GIM (7 items rated using a 5-point Likert scale: 1 = needs improvement, 2-4 = meets expectations, 5 = exceeds expectations, 0 = not applicable; free-text comments required for ratings of 1 or 5) and psychiatry (5 items rated using a 4-point Likert scale ranging from 1 = strongly disagree to 4 = strongly agree).
- Peer feedback survey for neurology³¹ (6 items rated using a 5-point Likert scale: 1 = never, 2 = rarely, 3 = occasionally, 4 = frequently, 5 = always, 0 = not applicable).

These data were collected for a variety of internal, accreditation, certification, and regulatory reasons, as is typical of physician performance data.³²⁻³⁴ Scores were not linked to physician reimbursement or published publicly. The GIM and psychiatry MSF surveys

Download English Version:

<https://daneshyari.com/en/article/8924765>

Download Persian Version:

<https://daneshyari.com/article/8924765>

[Daneshyari.com](https://daneshyari.com)