UCLA

UCLA Previously Published Works

Title

Adjustment of Patient Experience Surveys for How People Respond.

Permalink

<https://escholarship.org/uc/item/1nb6f6h2>

Journal

Medical care, 59(3)

ISSN

0025-7079

Authors

Cefalu, Matthew

Elliott, Marc N

Hays, Ron D

Publication Date

2021-03-01

DOI

10.1097/mlr.0000000000001489

Peer reviewed

Adjustment of Patient Experience Surveys for How People Respond

Matthew Cefalu, PhD,* Marc N. Elliott, PhD,* and Ron D. Hays, PhD†

Background: Patient surveys are the primary tool to measure patient experiences of care. Caution must be taken when analyzing these data, as responses can be influenced by factors that do not reflect the quality of care received.

Objectives: To provide a practical overview of adjusting patient experience survey results to address bias related to patient case-mix, extreme response tendency, and mode of survey administration.

Research Design: We discuss options for adjustment for biases in how people respond to patient experience surveys.

Results: Case-mix adjustment (CMA) aims to compare provider performance that would have been observed if all providers had treated the same set of patients by removing the effects of patient characteristics that vary across providers. Extreme response tendency can bias the measurement of the disparities in patient experiences even after typical CMAs, since differences in patients' use of extreme response options may affect patient experience scores when they have a skewed distribution. Survey mode may affect scores for the provider entity being evaluated (eg, hospital) more than CMA if survey mode differs at the provider level.

Conclusions: It is best practice to evaluate known source of bias when analyzing patient experience surveys. Failure to adjust for patient case-mix, extreme response tendency, and survey mode in patient experience surveys may lead to erroneous comparisons of providers.

Key Words: patient experience, case mix, CAHPS, patient surveys (*Med Care* 2021;00: 000–000)

Patient surveys are important for gathering information about health care quality and are the primary way to measure patient experiences of care. These surveys are used to target quality improvement efforts and provide information about the performance of different health care providers and plans.^{1–6} In addition, the surveys are used as components of pay-for-performance programs.⁷ Patient experience surveys

have gained widespread use and have been adopted across a variety of organizations and care settings.⁸

Survey responses can be influenced by factors that do not reflect the quality of care received, such as response tendencies of the people who complete them. For example, some survey respondents may be more likely than others to report positive experiences for the same health care.^{9–13} Response tendencies may be specific to a survey mode—for example, socially desirable responses may be more likely for to face-to-face interviews than for mail surveys.¹⁴ Systematic differences in these factors jeopardize the comparability of patient experience survey results between different providers.

In this article, we review options for addressing bias in patient experience survey results due to how people respond to survey items: patient case-mix, extreme response tendencies, and mode of survey administration. We do not discuss approaches for accounting for *who* responds (eg, adjustment for the sampling design and nonresponse) because this has a rich history in survey analysis.¹⁵ Further, this article is not intended as an overview of decisions related to the collection of survey data, such as costs, sample size, sampling design, and power.

ADJUSTMENT OF PATIENT EXPERIENCE SURVEYS FOR HOW PEOPLE RESPOND

Total survey error combines variance, which is largely driven by sample size, sampling design, and implementation, and bias.¹⁶ Here we focus on bias of the survey results due to how people respond to survey items. Compared with unadjusted results, the adjustments discussed in this article add little variance, but can substantially reduce bias, leading to an overall reduction in mean squared error. For this reason, variance of the survey results is not discussed further. We review case-mix, extreme response tendency, and mode of survey administration using evidence from the English General Practice Patient Survey,¹¹ CAHPS Medicare Fee-for-Service Survey,¹⁷ and CAHPS Hospital Survey.¹⁰

An unadjusted comparison of patient experiences across health care providers (eg, health plans, physician groups, hospitals, physicians) is problematic if characteristics of the patients (case-mix) treated by the health care entities being compared vary and are related to how patients respond to patient experience of care survey questions. For example, younger and more educated individuals provide less favorable evaluations of health care.^{9,18,19} As a result, providers that treat younger and more educated individuals will tend to have unadjusted patient experience scores that are less positive due to the characteristics of the patients they treat.

From the *RAND Corporation, Santa Monica; and †Division of General Internal Medicine & Health Services Research, University of California, Los Angeles, Los Angeles, CA.

This research was supported by the Agency for Healthcare Research and Quality: U18 HS025920.

The authors declare no conflicts of interest.

Correspondence to: Marc N. Elliott, PhD, RAND Corporation, 1776 Main Street, Santa Monica, CA 90401. E-mail: elliott@rand.org.

Copyright © 2021 Wolters Kluwer Health, Inc. All rights reserved.
ISSN: 0025-7079/21/000-000

ADJUSTMENT FOR POSITIVE RESPONSE TENDENCIES

Certain groups of patients have the tendency to respond more favorably to surveys, a concept known as positive response tendency. Since providers serve a varying mix of patients, positive response tendency can bias comparisons of providers. Case-mix adjustment (CMA) aims to remove the effects of patient characteristics that vary across providers so fairer comparisons can be made.¹⁹ The principle behind the adjustment is to assess the performance that would have been observed if all providers treated the same set of patients by removing the effects of patient characteristics not under provider control, but which may influence response and vary across providers. This helps to level the playing field and minimize incentives to avoid patients more likely to provide lower scores irrespective of care.²⁰ CMA should not include factors that are aspects of care (eg, utilization) or are products of the quality of care (eg, hospital-born infections); controlling for factors that have a causal link to quality of care will bias the results and remove real differences in care between providers.¹⁰ Case-mix adjusters should be associated with patient experience and vary across providers.¹⁸

A common CMA method is linear regression predicting outcomes from provider intercepts and case-mix variables. The coefficients on the case-mix variables represent the within-provider associations between those variables and the patient experience outcomes, and the predicted value from the CMA model represents the average outcome among patients with similar characteristics. The difference between the observed and predicted values represents the patient-level deviation from the average outcome of patients with similar characteristics. Averaging the patient-level deviations within a provider yields information on the performance of a provider relative to what would have been expected based on their case-mix.

The linear regression approach to CMA requires several technical assumptions, 2 of which are highlighted here. First, the CMA model is assumed to be correctly specified,⁽¹⁹⁾ that is, the model used to predict the average outcomes across patient characteristics must capture all the important relationships between the case-mix adjusters and the outcome. Second, case-mix variables are assumed to be measured without error. Measurement error in the case-mix adjusters would bias the adjustment. For example, if age is incorrectly measured such that the patients of a certain provider are coded as younger than their actual age, the provider will receive a different case-mix adjusted score than if age were correctly measured. This occurs because the provider's score gets adjusted upward due to the tendency of younger individuals to respond less favorably. Assessing the impact of CMA is important both for understanding the importance of CMA and for data-driven CMA model building.

We review evidence from the English General Practice Patient Survey¹¹ to highlight these concepts. This study sought to “quantify the impact of [CMA] on practice-level scores in a national survey of patient experience, to identify why and when it may be useful to adjust for case mix, and to discuss unresolved policy issues regarding the use of [CMA].” In that study, CMA had little impact on an access item (getting through on the phone), with >1% of the

variance of practice scores accounted for by CMA and <1% of practices gaining 10+ percentile points in rank compared to unadjusted practice-level mean scores. However, CMA affected nurse communication substantially, accounting for 7% of the variance and changing the percentile rank of 14% of practices by 10+ points.

ADJUSTMENT FOR EXTREME RESPONSE TENDENCIES

CMA addresses the tendency of some groups to respond more favorably to a survey (positive response tendency). Another systematic response tendency is differing proclivity for the endpoints (most negative or positive option) of the response scale. For reporting results using top box scoring where the proportion of individuals endorsing the top of the response scale is reported, CMA and extreme response tendency (ERT) adjustment will have similar effects since the outcome itself is an extreme response. However, when survey results are reported more granularly, such as reporting top, middle, and bottom box score or the mean score, CMA assumes constant differences across the whole response scale.

Past research has shown counterintuitive patterns of patient experiences among subgroups, even after CMA.¹⁷ CMA can only account for mean differences caused by ERT and will not adjust for distributional effects beyond the mean. For example, Black patients receive worse care than non-Hispanic White patients according to clinical guidelines,²¹ but often report more positive evaluations of care.^{22–25} These results could be the result of lower expectations of care. They may also be due to greater ERT among Black than White patients; ERT is larger on average among Black, less-educated, lower-income, and older respondents. ERT can bias the measurement of the true disparity in patient experiences even after typical CMAs.¹⁷

One method for adjusting for ERT is to use a 2-stage heuristic model.¹⁷ The first stage uses educational attainment to adjust for ERT within categories of the response scale (eg, 0–6, 7–8, 9–10) and then uses educational attainment to adjust for positive response tendency. This removes within-response category differences associated with educational attainment. Then, positive response tendency is addressed as a between-response category effect of educational attainment, which removes the tendency of different educational attainment groups to respond more favorably, absent differences in ERT.

In the second stage, education-adjusted ratings are adjusted for positive response tendency using other case-mix variables in a standard CMA model. While standard approaches found that Medicaid eligibility in CAHPS Medicare Fee-for-Service data was associated with more positive reports of care and having additional private insurance was associated with worse reports of care, adjustment for ERT reversed these results into the intuitive direction.¹⁷

ADJUSTMENTS FOR VARIABLES FIXED AT THE PROVIDER-LEVEL: MODE OF SURVEY ADMINISTRATION

Because CMA methods rely on within-provider variation, they are best suited for accounting for patient-level factors. However, there may be a need to adjust for provider-level factors

in some settings. For example, the HCAHPS survey allows for hospitals to choose among 4 survey modes: mail, telephone, mail with telephone follow-up, or active interactive voice response and then apply it to all surveyed patients in a hospital.²⁶ The choice of mode affects the composition of respondents (ie, those who respond to a mail survey are different from those who respond to a telephone survey). CMA can account for compositional differences of respondents across hospitals due to survey mode,¹⁰ but it cannot adjust for the effect of survey mode on responses among those who respond.

Without proper adjustment, providers could choose the survey mode that results in the most positive patient experience scores. Survey mode adjustment requires an experiment in which individuals are randomly assigned a survey mode within providers or similar units, so that the mode differences can be assessed. The survey mode effect estimated from the experiment can then be used to adjust for survey mode in future administrations. Both CMA and survey mode adjustment should be used to account for the effects of survey mode, such that case-mix adjusters are used as covariates in the model estimating mode effects, using the same functional form in which they will be used in later applications.¹⁰

Even if case-mix and survey mode had similar effects at the patient level, survey mode may have a greater impact on provider-level scores because survey mode is often confounded with provider. One study showed a statistically significant ($P < 0.05$) mode effect on the mean responses for 6 out of 8 HCAHPS measures and more positive evaluations were observed for phone and interactive voice response modes than mail for all 8 HCAHPS measures evaluated. Failure to account for these mode effects would result in substantial errors in ranking the hospitals. Without mode adjustment, “a hospital surveyed by telephone that was truly at the 25th percentile on Pain Management ... would appear to rank at the 66th percentile.”¹⁰ Allowing hospitals the choice of 4 different survey administration modes in HCAHPS allows flexibility but requires mode adjustment to ensure hospital-levels scores are equitable and comparable.

DISCUSSION

It is best practice to adjust for any known source of bias when analyzing survey data. In patient experience surveys publicly reported or linked to provider payment, it is imperative to ensure that the results are equitable. If results from patient experience provided systematic advantage to certain entities, then there is the possibility of gaming the approach by excluding patients seen as promoting lower scores, which could exacerbate problems of comparability and access to care.²⁰

Although not required, a common assumption for the CMA described in this article is that the within-provider relationship between the case-mix variables and the outcome is the same for all provider entities. This assumption can be tested by including an interaction between the provider entity and the case-mix variables.¹⁹ If a significant interaction is detected, we recommend interpreting the heterogeneity substantively (ie, reflecting disparities in care).

The CMA method reviewed here is the most common approach for detecting and accounting for how people respond to

patient experience surveys. Addressing some aspects of response bias is possible when differential item functioning is observed by estimating subgroup specific item parameters,²⁷ but this alternative is limited to multi-item scales or composites calibrated using item response theory methods.^{28,29}

Failure to adjust for patient case-mix, extreme response tendency, and survey mode in patient experience surveys may lead to erroneous comparisons of providers.

ACKNOWLEDGMENTS

The authors would like to thank Biayna Darabidian for help with preparation of the manuscript.

REFERENCES

1. Browne K, Roseman D, Shaller D, et al. Analysis & commentary. Measuring patient experience as a strategy for improving primary care. *Health Aff (Millwood)*. 2010;29:921–925.
2. Davies E, Shaller D, Edgman-Levitan S, et al. Evaluating the use of a modified CAHPS survey to support improvements in patient-centred care: lessons from a quality improvement collaborative. *Health Expect*. 2008;11:160–176.
3. Friedberg MW, Steelfisher GK, Karp M, et al. Physician groups' use of data from patient experience surveys. *J Gen Intern Med*. 2011;26:498–504.
4. Goldstein E, Cleary PD, Langwell KM, et al. Medicare Managed Care CAHPS(®): a tool for performance improvement. *Health Care Financ Rev*. 2001;22:101–107.
5. Quigley DD, Mendel PJ, Predmore ZS, et al. Use of CAHPS(®) patient experience survey data as part of a patient-centered medical home quality improvement initiative. *J Health Leadersh*. 2015;7:41–54.
6. Quigley DD, Palmaru A, Lerner C, et al. A review of best practices for monitoring and improving inpatient pediatric patient experiences. *Hosp Pediatr*. 2020;10:277–285.
7. Anhang Price R, Elliott MN, Zaslavsky AM, et al. Examining the role of patient experience surveys in measuring health care quality. *Med Care Res Rev*. 2014;71:522–554.
8. Cleary PD. Evolving concepts of patient-centered care and the assessment of patient care experiences: optimism and opposition. *J Health Polit Policy Law*. 2016;41:675–696.
9. O'Malley AJ, Zaslavsky AM, Elliott MN, et al. Case-mix adjustment of the CAHPS Hospital Survey. *Health Serv Res*. 2005;40:2162–2181.
10. Elliott MN, Zaslavsky AM, Goldstein E, et al. Effects of survey mode, patient mix, and nonresponse on CAHPS hospital survey scores. *Health Serv Res*. 2009;44:501–518.
11. Paddison C, Elliott M, Parker R, et al. Should measures of patient experience in primary care be adjusted for case mix? Evidence from the English General Practice Patient Survey. *BMJ Qual Saf*. 2012;21:634–640.
12. Chen AY, Elliott MN, Spritzer KL, et al. Differences in CAHPS reports and ratings of health care provided to adults and children. *Med Care*. 2012;50(suppl):S35–S39.
13. Johnson ML, Rodriguez HP, Solorio MR. Case-mix adjustment and the comparison of community health center performance on patient experience measures. *Health Serv Res*. 2010;45:670–690.
14. Holbrook AL, Green MC, Krosnick JA. Telephone versus face-to-face interviewing of national probability samples with long questionnaires: comparisons of respondent satisficing and social desirability response bias. *Public Opin Q*. 2003;67:79–125.
15. Heeringa SG, West BT, Berglund PA. *Applied Survey Data Analysis*. Boca Raton, FL: Chapman and Hall/CRC; 2010.
16. Elliott M, Haviland A. Use of a web-based convenience sample to supplement a probability sample. *Surv Methodol*. 2007;33:211–215.
17. Elliott MN, Haviland AM, Kanouse DE, et al. Adjusting for subgroup differences in extreme response tendency in ratings of health care: impact on disparity estimates. *Health Serv Res*. 2009;44:542–561.
18. Zaslavsky AM, Zaboriski LB, Ding L, et al. Adjusting performance measures to ensure equitable plan comparisons. *Health Care Financ Rev*. 2001;22:109–126.
19. Elliott MN, Swartz R, Adams J, et al. Case-mix adjustment of the National CAHPS benchmarking data 1.0: a violation of model assumptions? *Health Serv Res*. 2001;36:555–573.

20. Institute of Medicine, National Academies of Sciences Engineering, and Medicine. *Accounting for Social Risk Factors in Medicare Payment: Identifying Social Risk Factors*. Washington, DC: The National Academies Press; 2016.
21. Agency for Healthcare Research and Quality. *2017 National Healthcare Quality and Disparities Report*. Rockville, MD: Agency for Healthcare Research and Quality; 2018.
22. Bashshur RL, Metzner CA, Worden C. Consumer satisfaction with group practice, the CHA case. *Am J Public Health Nations Health*. 1967;57:1991–1999.
23. Morales LS, Elliott MN, Weech-Maldonado R, et al. Differences in CAHPS adult survey reports and ratings by race and ethnicity: an analysis of the National CAHPS benchmarking data 1.0. *Health Serv Res*. 2001;36:595–617.
24. Weech-Maldonado R, Elliott MN, Morales LS, et al. Health plan effects on patient assessments of Medicaid managed care among racial/ethnic minorities. *J Gen Intern Med*. 2004;19:136–145.
25. Dayton E, Zhan C, Sangl J, et al. Racial and ethnic differences in patient assessments of interactions with providers: disparities or measurement biases? *Am J Med Qual*. 2006;21:109–114.
26. Giordano LA, Elliott MN, Goldstein E, et al. Development, implementation, and public reporting of the HCAHPS survey. *Med Care Res Rev*. 2010;67:27–37.
27. Paz SH, Spritzer KL, Morales LS, et al. Evaluation of the Patient-Reported Outcomes Information System (PROMIS(R)) Spanish-language physical functioning items. *Qual Life Res*. 2013;22:1819–1830.
28. Setodji CM, Reise SP, Morales LS, et al. Differential item functioning by survey language among older Hispanics enrolled in Medicare managed care: a new method for anchor item selection. *Med Care*. 2011;49:461–468.
29. Yang FM, Heslin KC, Mehta KM, et al. A comparison of item response theory-based methods for examining differential item functioning in object naming test by language of assessment among older Latinos. *Psychol Test Assess Model*. 2011;53:440–460.