Title

Differences in perceptions of health care between Asian Americans and non-Hispanic whites on the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Clinician and Group Adult Visit Survey 1.0.

Author

Ahmedov, Mohirjon

Publication Date

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Differences in perceptions of health care between Asian Americans and non-Hispanic whites on the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Clinician and Group Adult Visit Survey 1.0.

A dissertation submitted in partial satisfaction of

the requirements for the degree Doctor Philosophy

in Health Services

by

Mohirjon Ahmedov

2016

ABSTRACT OF THE DISSERTATION

Differences in perceptions of health care between Asian Americans and non-Hispanic whites on the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Clinician and Group Adult Visit Survey 1.0.

by

Mohirjon Ahmedov

Doctor of Philosophy in Health Services

University of California, Los Angeles, 2016

Professor Ronald D. Hays, Chair

Racial/ethnic disparities in patient experiences are widely reported. Asian Americans (Asians) consistently report worse care experiences in the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) surveys than non-Hispanic whites (Whites). However, little is known whether these race/ethnic differences in reports and ratings are due to differences in care experiences or differential response tendencies.

This dissertation consists of three studies. The first study compares reports and ratings of care between Asians and Whites using ordinary least squares analyses. The second study evaluates whether the hypothesized factor structure underlying the scoring of the CAHPS survey is confirmed in the survey dataset using categorical confirmatory factor analytic models. The third paper evaluates measurement invariance between Asians and Whites using a multiple group

confirmatory factor analysis. The dissertation uses the Clinician & Group CAHPS Adult Visit Surveys 1.0 data collected in 2011.

In the first study, Asians reported worse care experiences on *access to care*, *office staff courtesy and helpfulness*, *rated their doctor* lower and were less likely to *recommend their doctor to family and friends* than Whites. On *physician communication*, no significant difference was noted between Asians and Whites. The reported differences in care experiences between Asians and Whites are likely due to real racial/ethnic differences in care received rather than lack of measurement invariance. The study findings have several important policy implications and provide directions for future research. Quality improvement initiatives in primary care need to be tailored towards reducing racial/ethnic differences in care. Further research will be needed to understand what are the underlying reasons for differential care for Asians and Whites in ambulatory care.

The dissertation of Mohirjon Ahmedov is approved.

William E. Cunningham

Honghu Liu

Nadereh Pourat

Ronald D. Hays, Chair

University of California, Los Angeles

2016

TABLE OF CONTENTS

List of Tables

List of figures

List of appendices

List of acronyms

MEPS	Medical Expenditure Panel Survey
CAHPS®	Consumer Assessment of Healthcare Providers and Systems
AHRQ	Agency for Healthcare Research and Quality
CG-CAHPS	CAHPS Clinician and Group Survey
OLS	Ordinary least squares
CFA	Confirmatory factor analytic model
MG-CFA	Multiple group confirmatory factor analytic model
WLSMV	Mean- and Variance-adjusted Weighted Least Squares estimation method
ICC	Intraclass correlation coefficients
H-CAHPS	Hospital Consumer Assessment of Healthcare Providers and Systems
MLR	Maximum likelihood estimation with robust standard errors
ML	Maximum likelihood estimation
RMSEA	Root mean square error of approximation
CFI	Comparative fit index
TLI	Tucker-Lewis index

Acknowledgements

First I would like to thank my long-time mentor and dissertation chair, Dr. Ron Hays, for his unwavering support throughout this long journey. I am forever indebted to you. I am also very thankful for my committee members - Drs. William Cunningham, Honghu Liu and Nadereh Pourat – for their continued support and guidance.

Vita

1998	M.D., II Tashkent State Medical Institute, Uzbekistan
2000	Clinical fellow, II Tashkent State Medical Institute, Uzbekistan
2000 – 2003	Assistant Professor, Department of Internal Medicine, Department of Public Health, II Tashkent State Medical Institute, Uzbekistan
2005	M.P.H, University of California, Los Angeles, U.S.A.
2005 – 2008	Lecturer, Senior Lecturer, II Tashkent State Medical Institute, Uzbekistan
2010 – 2013	Head of unit, Tashkent Medical Diagnostics Center, Uzbekistan
2014 - 2016	Consultant, World Health Organization, Regional Office for Europe

Selected publications and reports

Ahmedov M, Green J, Azimov R, Avezova G, Inakov Sh, Mamatkulov B. Addressing the challenges of improving the quality of primary care for Chronic Heart Failure in Uzbekistan: a qualitative study. Health Policy and Planning, advance access 2012 Sep 16

Rechel B, **Ahmedov M**, Akkazieva B, Katsaga A, Khodjamurodov G, McKee M. Lessons from two decades of health reform in Central Asia, Health Policy and Planning, advance access 2011 May 24

Ahmedov M, Rechel B, Alimova V, Azimov R. Primary health care reform in Uzbekistan, International Journal of Health Planning and Management, 2007 Aug 29

Ahmedov M, Azimov R, Mutalova Z, Huseynov Sh, Tsoyi E, Rechel B. Uzbekistan: Health system review. Health Systems in Transition, World Health Organization, 2014

Ahmedov M, Azimov R, Alimova V, Rechel B. Uzbekistan: Health system review. Health Systems in Transition, 2007; 9(3): 1–210. World Health Organization

Chapter 1 Introduction to the dissertation

The racial/ethnic composition of the US has changed substantially in the last few decades. Asians Americans (Asians) grew faster than any other race/ethnic subgroup-- from 10.2 million in 2000 to 14.7 million in 2010 [1]. Along with the increase in percentage of minorities, there has been increased attention to racial/ethnic disparities in health care by the Institute of Medicine and other national health care agencies [2-4].

Patient perceptions of care are indicators of quality of care from the patient's perspective that are associated with other indicators of quality of care [5-12]. Patient evaluations of care are widely used by health plans, physician groups, hospitals, and other health care providers to inform patients about their health care options and improve quality of care [5, 8, 13-16].

A number of studies have reported disparities in patient experiences with care between Asians and other racial/ethnic subgroups. For example, Snyder and colleagues (2000) found that Asians reported the worst access to care among all racial/ethnic subgroups. The authors concluded that poor access to care by Asians might be explained in part by difficulties in communication and cultural differences. Phillips and colleagues analyzed the 1996 Medical Expenditure Panel Survey (MEPS) data and reported that Asians were more dissatisfied with their health care than other race/ethnic subgroups [17]. An analysis of the Medical Outcomes Study suggested that even though Asians had better or similar health as that of non-Hispanic Whites (Whites), they were less satisfied with the care received [18]. Asians were also the least satisfied subgroup in the 1998 National Research Corporation Healthcare Market Guide® survey [19]. Ngo-Metzger and colleagues (2004) found that Asians were more likely than Whites to

report that their regular doctors did not understand their background and values. Several other studies documented worse experiences with care for Asians than Whites [20-22].

Differences in experiences of care by race/ethnic groups have been examined in studies that have administered the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) surveys [23-25]. The CAHPS project has been funded by the Agency for Healthcare Research and Quality (AHRQ) since 1995 to standardize patient experience measures and has become since the most widely used measure of patient experiences in the US [26-28]. Race/ethnic differences on CAHPS surveys have been observed in several studies [29-36]. In particular, Morales and colleagues analyzed CAHPS 1.0 health plan survey data and found that Asians expressed worse experiences with physician communication and other aspects of care than Whites, but global ratings (health plan, health care, personal doctor or nurse, specialty care) were similar between these two groups [24]. Weech-Maldonado and colleagues also found that Asians fared worse than Whites in terms of CAHPS patient experience reports about care [25].

Disparities in perceptions of care for Asians and Whites could be due to true differences in care received, or possibly because of differences in expectations about care or response styles. Determining the true reasons for the worse reports about care for Asians has important policy implications. If differences in patient experiences are due to differences in care, then further research needs to be carried out to understand the reasons for differential care for racial/ethnic groups. Quality improvement initiatives need to be designed and implemented to address underlying reasons for differential care.

Differential response tendencies by survey language of administration and other characteristics may also account for differences in CAHPS survey results by different race/ethnic

subgroups [37]. Hence, research is needed to understand which CAHPS survey items are affected by differential racial/ethnic response tendencies. This research could also provide recommendations on how to account for any observed differential item functioning. For example, it might be necessary to adjust responses to survey items to make them comparable across subgroups.

The three papers in this dissertation build on earlier research findings and compare differences in reports and ratings of care on the CAHPS Clinician and Group (CG-CAHPS) Adult Visit Survey 1.0. between Asians and Whites and evaluate the extent to which differences are due to response tendencies The first paper compares reports and ratings of care between Asians and Whites using ordinary least squares regression models. Given prior studies, we hypothesize that Asians will have worse reports and ratings of care than Whites. Because Asians have been shown to have equal or better "objective" access to care, and to be less likely to change their physician due to dissatisfaction, we hypothesize that there may be some lack of measurement equivalence between the two racial/ethnic groups, The second and third papers address measurement equivalence between these two subgroups. In the second paper, the hypothesized three factor structure (access to care, physician communication and global ratings, and office staff courtesy and helpfulness) is evaluated in the overall sample. Categorical confirmatory factor analytic models are estimated. The third paper compares the factor structure in Asians and Whites using single and multiple-group confirmatory factor analyses.

References

1. U.S. Census Bureau. *United States Census 2010*. 2011 [cited 2016 August 20]; Available from: http://2010.census.gov/news/releases/operations/cb11-cn125.html

2. Nelson, A.R., B.D. Smedley, and A.Y. Stith, *Unequal treatment: confronting racial and ethnic disparities in health care,* in *National Academies Press*. 2002, Institute of Medicine: Washington DC.

3. AHRQ, *National Healthcare Disparities Report 2010*. 2010, Agency for Healthcare Research and Quality: Rockvillle, MD.

4. Richardson, W.C., et al., *Crossing the quality chasm: a new health system for the 21st century*. 2001, Institute of Medicine, National Academy Press Washington, DC.

5. Cleary, P. and B. McNeil, *Patient satisfaction as an indicator of quality care.* Inquiry : a journal of medical care organization, provision and financing, 1988. **25**(1): p. 25-36.

6. Vuori, H., *Patient satisfaction--an attribute or indicator of the quality of care?* QRB. Quality review bulletin, 1987. **13**(3): p. 106-108.

7. Donabedian, A., *The quality of care: How can it be assessed?* JAMA: The Journal of the American Medical Association, 1988. **260**(12): p. 1743-1748.

8. Rosenthal, G. and S. Shannon, *The Use of Patient Perceptions in the Evaluation of Health-Care Delivery Systems.* Medical care, 1997. **35**(11): p. NS58-NS68.

9. Sequist, T., et al., *Quality Monitoring of Physicians: Linking Patients' Experiences of Care to Clinical Quality and Outcomes.* Journal of General Internal Medicine, 2008. **23**(11): p. 1784-1790.

10. Boulding, W., et al., *Relationship Between Patient Satisfaction With Inpatient Care and Hospital Readmission Within 30 Days.* American Journal of Managed Care, 2011. **17**(1): p. 41-48.

11. Glickman, S.W., et al., *Patient Satisfaction and Its Relationship With Clinical Quality and Inpatient Mortality in Acute Myocardial Infarction.* Circulation: Cardiovascular Quality and Outcomes, 2010. **3**(2): p. 188-195.

12. Sequist, T.D., et al., *Measuring chronic care delivery: patient experiences and clinical performance.* International Journal for Quality in Health Care, 2012. **24**(3): p. 206-213.

13. Williams, B., *Patient satisfaction: A valid concept?* Social Science & Medicine, 1994. **38**(4): p. 509-516.

14. Cleary, P. and S. Edgman-Levitan, *Health care quality: Incorporating consumer perspectives.* JAMA: The Journal of the American Medical Association, 1997. **278**(19): p. 1608-1612.

15. Sitzia, J. and N. Wood, *Patient satisfaction: A review of issues and concepts.* Social Science & Medicine, 1997. **45**(12): p. 1829-1843.

16. Castle, N., J. Brown, and K. Hepner, Hays, R., *Review of the Literature on Survey Instruments Used to Collect Data on Hospital Patients' Perceptions of Care.* Health Services Research, 2005. **40**(6p2): p. 1996-2017.

17. Phillips, K., M. Mayer, and L. Aday, *Barriers to care among racial/ethnic groups under managed care.* Health Affairs, 2000. **19**(4): p. 65-75.

18. Meredith, L. and A. Siu, *Variation and Quality of Self-Report Health Data: Asians and Pacific Islanders Compared with Other Ethnic Groups.* Medical Care, 1995. **33**(11): p. 1120-1131.

19. Haviland, M., et al., *Do Health Care Ratings Differ by Race or Ethnicity?* Joint Commission Journal on Quality and Patient Safety, 2003. **29**(3): p. 134-145.

20. Murray-García, J., J. Selby, and J. Schmittdiel, Grumbach, K., Quesenberry, Ch., *Racial and Ethnic Differences in a Patient Survey: Patients' Values, Ratings, and Reports Regarding Physician Primary Care Performance in a Large Health Maintenance Organization.* Medical Care, 2000. **38**(3): p. 300-310.

21. Ngo-Metzger, Q., A. Legedza, and R. Phillips, *Asian Americans' Reports of Their Health Care Experiences.* Journal of General Internal Medicine, 2004. **19**(2): p. 111-119.

22. Rodriguez, H., T. von Glahn, and D. Grembowski, Rogers, W., Safran, D., *Physician Effects on Racial and Ethnic Disparities in Patients' Experiences of Primary Care.* Journal of General Internal Medicine, 2008. **23**(10): p. 1666-1672.

23. Morales, L., W. Cunningham, and J. Brown, Liu, H., Hays, R., *Are Latinos Less Satisfied with Communication by Health Care Providers?* Journal of General Internal Medicine, 1999. **14**(7): p. 409-417.

24. Morales, L., 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 Services Research, 2001. **36**(3): p. 595-617.

25. Weech-Maldonado, R., et al., *Race/Ethnicity, Language, and Patients' Assessments of Care in Medicaid Managed Care.* Health Services Research, 2003. **38**(3): p. 789-808.

26. CMS. *Centers for Medicare & Medicaid Services: Overview*. 2012 [cited 2012 August 13]; Available from: http://www.Medicare.gov/choices/Overview.asp.

27. Zaslavsky, A., et al., *How Consumer Assessments of Managed Care Vary Within and Among Markets.* Inquiry, 2000. **37**(2): p. 146-161.

28. Landon, B., et al., *COmparison of performance of traditional medicare vs medicare managed care.* JAMA: The Journal of the American Medical Association, 2004. **291**(14): p. 1744-1752.

29. Haviland, A., M. Elliott, and R. Weech-Maldonado, Hambarsoomian, K., Orr, N., & Hays, R. D. , *Racial/ethnic disparities in Medicare Part D experiences.* Medical Care, Supplement, in press.

30. Elliott, M., A. Haviland, and D. Kanouse, Hambarsoomian, K., & Hays, R., *Adjusting for Subgroup Differences in Extreme Response Tendency in Ratings of Health Care: Impact on Disparity Estimates.* Health Services Research, 2009. **44**(2p1): p. 542-561.

31. Fongwa, M., W. Cunningham, and R. Weech-Maldonado, Gutierrez, P., Hays, R., *Reports and ratings of care: Black and White Medicare enrollees.* Journal of Health Care for the Poor and Underserved, 2008. **19**(4): p. 1136-1147.

32. Weech-Maldonado, R., M. Fongwa, and P. Gutierrez, & Hays, R., *Language and Regional Differences in Evaluations of Medicare Managed Care by Hispanics.* Health Services Research, 2008. **43**(2): p. 552-568.

33. Weech-Maldonado, R., M. Elliott, and A. Oluwole, Schiller, K., & Hays, R., *Survey Response Style and Differential Use of CAHPS Rating Scales by Hispanics.* Medical Care, 2008. **46**(9): p. 963-968 10.1097/MLR.0b013e3181791924.

34. Fongwa, M., W. Cunningham, and R. Weech-Maldonado, Gutierrez, P., Hays, R. , *Comparison of Data Quality for Reports and Ratings of Ambulatory Care by African American and White Medicare Managed Care Enrollees.* Journal of Aging and Health, 2006. **18**(5): p. 707-721.

35. Hurtado, M., J. Angeles, and S. Blahut, & Hays, R., *Assessment of the Equivalence of the Spanish and English Versions of the CAHPS® Hospital Survey on the Quality of Inpatient Care.* Health Services Research, 2005. **40**(6p2): p. 2140-2161.

36. Weech-Maldonado, R., et al., *Health Plan Effects on Patient Assessments of Medicaid Managed Care Among Racial/Ethnic Minorities.* Journal of General Internal Medicine, 2004. **19**(2): p. 136-145.

37. Setodji, C., S. Reise, and L. Morales, Fongwa, M., Hays, R., *Differential Item Functioning by Survey Language Among Older Hispanics Enrolled in Medicare Managed Care: A New Method for Anchor Item Selection.* Medical care, 2011. **49**(5): p. 461-468 10.1097/MLR.0b013e318207edb5.

Chapter 2 Differences in perceptions of health care between Asian Americans and non-Hispanic whites: an analysis of the CAHPS Clinician and Group Adult Visit Survey 1.0.

Background

Patient evaluations of care are widely used by health plans, physician groups, hospitals, and other health care providers to inform patients about their health plan and provider options and to improve quality of care[1-4]. In 1995, the Agency for Healthcare Research and Quality (AHRQ) funded the development of standardized patient experience measures — the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) project. The CAHPS effort has generated surveys that are used to compare providers of care in both ambulatory and inpatient settings in the healthcare system. For example, the Centers for Medicare & Medicaid Services sponsors annual surveys of beneficiaries in fee-for-service and managed care[5-7].

CAHPS surveys ask patients to report and evaluate their health care experiences on aspects of care they are able to assess such as doctors' communication skills, office staff helpfulness and access to care [8]. The CAHPS survey development is guided by patient interests and needs in choosing health plans, physician groups, hospitals, nursing homes, and dialysis facilities [9]. The extent to which patient experience of care reports influence consumer choice is still an open question but CAHPS surveys have been shown to stimulate quality improvement efforts among health providers [10-15].

Racial/ethnic disparities in the US: Asian Americans and non-Hispanic whites

The racial/ethnic composition of the US has changed substantially over the last decade. Asian Americans grew faster than any other major race/ethnic group from 10.2 million in 2000 to 14.7 million in 2010[16]. A number of studies have found racial/ethnic disparities in patient experiences with care. Snyder and colleagues (2000) found that Asian Americans reported the worst access to care among all racial/ethnic groups. Access measures were self-reported and included blood cholesterol measurements, wait times, and reaching the doctor's office via phone. The authors concluded that poor access to care by Asian Americans, despite the higher education and income levels, might be explained in part by difficulties in communication and cultural differences[17].

Phillips and colleagues analyzed the 1996 Medical Expenditure Panel Survey (MEPS) data and reported that African Americans, Hispanics, Asian Americans and Pacific Islanders were less satisfied with the care received than non-Hispanic whites -- Asian Americans were the most dissatisfied group [18]. An analysis of the Medical Outcomes Study data suggested that even though Asian Americans had a better or similar health as non-Hispanic whites, they were less satisfied with the care received[19]. Asian Americans were also found to be less satisfied with care in the 1998 National Research Corporation Healthcare Market Guide® survey [20]. Ngo-Metzgar and colleagues (2004) found that Asian Americans were more likely than non-Hispanic whites to report that their regular doctors did not understand their background and values. They were also more likely than non-Hispanic whites to report that their providers did not listen to them enough, spend as much time, or involve them in decisions. Several other studies have confirmed worse experiences with care for Asian Americans than whites[21-23].

Differences in experiences of care by racial/ethnic groups have also been documented in studies with the CAHPS surveys [24-26]. Morales and colleagues analyzed CAHPS Health Plan 1.0 survey data collected from 54 commercial and 34 Medicaid health plans and found that most minorities reported experiences similar to non-Hispanic whites, except for Asian Americans, who expressed worse perceptions of care [25]. Weech-Maldonado and colleagues analyzed data collected using the CAHPS Health Plan 2.0 version of the survey and found that racial-ethnic minorities, including Asian Americans, fared worse than non-Hispanic whites [26]. There are several other published examples of racial-ethnic differences on different versions of CAHPS Health Plan surveys [27-34].

Differences in experiences of care for Asian Americans and non-Hispanic whites could be due to true differences in care received, or possibly because of differences in expectations about care or response styles [35]. Despite the worse perceptions of care noted above, Asian American Medicare enrollees were found to have equal or better odds than non-Hispanic whites of receiving good technical quality of care measured on Health Plan Employer Data and Information Set (HEDIS) measures such as breast cancer screening with mammograms and use of beta blockers after myocardial infarction [36].

Dissatisfaction with care should increase the likelihood of changing a doctor. But Asian Americans were shown to be less likely to change their physicians despite greater dissatisfaction with care [37]. Saha and Hickam (2003) also reported lower satisfaction with health insurance and life in general for Asian Americans than non-Asian Americans. They speculated that lower satisfaction rates in Asian Americans could be explained by differences in response tendencies.

The possibility of response differences between Hispanics who completed the CAHPS Medicare survey in English versus Spanish was explored by Setodji et al. (2011). They found evidence for differential item functioning (DIF) for 3 out of 9 items. DIF is found when response to questions is driven by the factors other than the underlying construct, such as responders' age, gender, native language, socio-economic status and race/ethnicity. In this study, English speakers were more likely to choose extreme responses for the question related to a doctor spending enough time with them, while Spanish speakers endorsed extreme response options for the question asking if the doctor respected what they said [38].

Rodriguez and Crane (2011) explored DIF across racial/ethnic groups using the Clinician & Group CAHPS (CG-CAHPS) survey data. The data was collected for a quality improvement initiative from eight southern California medical groups. The authors found negligible DIF by race/ethnicity and concluded that the earlier reported racial/ethnic differences in care experiences were likely due to true differences in care received rather than DIF[39].

Earlier efforts exploring differences in patient experiences with health care between Asian Americans and non-Hispanic whites primarily focused on CAHPS Health Plan survey data. We found no studies that compared Asian Americans and non-Hispanic whites in the CAHPS Clinician & Group Adult Visit 1.0 Survey data. Asian Americans also are frequently grouped together with Pacific Islanders in the studies exploring for racial/ethnic differences using CAHPS surveys. Whether the differences found in earlier research hold up in other CAHPS surveys and when Asian Americans are analyzed separately is largely unknown. This study examines differences in reports and ratings of care between Asian Americans and non-Hispanic whites in CAHPS Clinician & Group Adult Visit 1.0 survey collected in 2011. Asian Americans are analyzed separately from Pacific Islanders. Based on prior research, we

hypothesize that Asian-Americans will report worse experiences with care than Whites in a CG-CAHPS Adult Visit 1.0 survey.

Data description

Instrument

The CAHPS® Clinician & Group Survey 1.0 suite, released in 2008, includes Visit Survey (Adult and Child) surveys. This study analyzes the 2011 adult data with the CAHPS Clinician and Group Adult Visit Survey 1.0 from the CAHPS Database. The survey focuses on the respondents' most recent visit to a primary care physician (i.e. an internist, family practitioner, obstetrician/gynecologist) or a specialist (i.e. a surgeon) throughout 2011.

The survey includes 41 questions: 13 of them elicit reports about care: 6 for *physician communication*, 2 for *office staff courtesy and helpfulness*, and 5 for *access to care*. The physician communication composite questions ask if the physician listened carefully, spent enough time, and gave instructions that were easily understandable. The two items in the office staff helpfulness and courtesy composite ask respondents whether office staff were helpful, courteous and showed respect. Both these composites use a three-response option scale (*Yes, definitely; Yes, somewhat; No*) and refer to the most recent visit. Access questions elicit information about timeliness of care, timeliness of answers to phone questions, and appointment wait times using four-response options (*Always, Usually, Sometimes, Never*) and refer to a 12-month period

The survey also includes an item assessing global perceptions of the doctor using a 0-10 response scale, where 0 is the worst possible doctor and 10 is the best possible doctor. Another question asks respondents if they would recommend their doctor to family and friends on a three-

response option scale (*Yes, definitely; Yes, somewhat; No*). Table 1 shows all 15 questions used to elicit reports and ratings of care and their response options used in this study.

Age, education, gender, race/ethnicity, and overall health are also assessed in the survey. The reliability and validity of the CAHPS surveys have been evaluated extensively[40-46]. CAHPS surveys in this dataset are administered in English and Spanish.

Dataset

The CAHPS Database is a warehouse of data collected from various organization in the US and facilitates comparison of survey findings among participating organizations. Organizations that collect CAHPS survey data themselves or pay vendors to collect data following CAHPS specifications (sponsors) can voluntarily contribute their data to the warehouse. The data analyzed for this study is the data collected in 2011.

The Adult Visit Survey data include 769 providers and 266,327 respondents. Approximately 98% of the surveys were administered by mail vs. 2% by phone; only 0.5% of the surveys were administered in Spanish. Previous research indicates that the telephone and mail responses to different versions of CAHPS® survey yield similar results[47-49]. Table 2 provides descriptive demographic information about the sample. Since the focus of this study is on Asian Americans and non-Hispanic whites, a separate table (Table 3) provides key demographic information about these two groups.

About 48% of patients report experiences with either family medicine or internal medicine doctors, seven percent with obstetricians/gynecologists, about 5% with surgeons, and 39% report visit experiences with other specialty doctors.

Missing values for the composite and global rating items range from 1% to 6% and are of little concern for the purposes of this analysis. We analyzed the two survey items with the highest missing values (*getting an answer to a medical question after regular hours* and *doctor explained things clearly*) and found no major differences between groups that responded to the item and those who did not. Missing values, therefore, were dropped from the analysis.

Dependent variables

The three CAHPS composites (*access to care* (5 items, four-response option scale), *physician communication* (6 items, three-response option scale), and *office staff courtesy and helpfulness* (2 items, three-response option scale)) and global rating items (2 items, three-response option and 0-10 response option scales) are used as dependent variables in the regression analyses. Reports about care (composites) and global rating items are transformed linearly to a 0-100 possible range where 0 was the worst possible experience and 100 was the best possible experience. The following formula is used to linearly transform composite and global rating items:

*Transformed item = (observed item value - minimally possible value) * 100/ (maximally possible value – minimally possible value)*

The items in each composite are averaged together to obtain a composite score. Cronbach's alpha is estimated for each of the three composites to assess internal consistency.

Independent variables

The CAHPS surveys ask if the respondent is of Hispanic or Latino origin. In addition, the survey includes a question that asks for: 1) White, 2) Black or African American, 3) Asian, 4)

Native Hawaiian or Pacific Islander, 5) American Indian or Alaska Native, or 6) other. A race/ethnicity variable is created based on the above questions. Those respondents confirming Hispanic or Latino origin are coded as Hispanic. Only those who report no Hispanic origin or had a missing value are coded in accordance to the ethnicity they identify themselves with such as *white, African American, Asian American*. Native Hawaiians, Pacific Islanders, American Indians, Alaska Native and other races are coded into the *other* racial-ethnic group. Thus, the new variable for race/ethnicity includes five subgroups: *Hispanic, Asian American, African American, White* and *other*.

We control for gender, age, education, practice site, region and self-reported health. Gender is a dichotomous variable: *male* and *female*. Age has six categories in the survey: *18-24, 25-34, 45-54, 55-64, 65-74, and 75+*. Education also has six categories: a) *eighth grade or less; b) some high school but did not complete; c) high school graduate or GED; d) some college or 2-year degree; e) 4-year college graduate; and f) more than four-year college degree*. Self-rated health has five categories: *excellent, very good, good, fair* and *poor*.

Analysis plan

Ordinary least squares regression is used to explore the associations of non-Hispanic Asian American vs. non-Hispanic white on the 0-100 linearly transformed CAHPS multi-item scales (composites) and global ratings using the following model:

$$OLS\ (Ratings\ and\ Reports) = \beta_0 + \beta_1(Asian) + \beta_2(Gender) + \beta_3(Education) + \beta_4(Health\ status) + \beta_5(Age) + \beta_6\ (Practice\ site) + \beta_8(Interactions:\ Asian\ by\ other\ independent\ variables)$$

First, an ordinary least squares (OLS) model is run with the main effects in the model. We adjust for the practice site using the cluster option in STATA 14. Adjusted scores (means)

for Asian Americans and whites for composites and global ratings are estimated using recycled predictions in STATA 14. Recycled predictions are obtained from regression models and used to understand marginal effects of an independent variable on a dependent variable. Independent variables other than the one of primary interest are fixed [50]. In our analysis, we fix the covariates at means. When using this approach, the adjusted scores correspond to the sample mean on the dependent variable, but it also may lead to counterintuitive findings[51]. For instance, discordance between coefficients and recycled predictions can be observed, as variables with larger means may differently impact recycled predictions of variables with smaller means than vice versa. We found no counterintuitive results when comparing the adjusted scores to the coefficients from our models.

Secondly, we evaluate possible two-way interactions between race/ethnicity (*Asian American vs. non- Asian American*) and each independent variable in our model. Interaction terms are included into the model to explore whether care experienced reported by Asian Americans vary from non-Asian Americans in the in the sample by age, gender, levels of education and reported health status. Non-significant interaction terms ($p = 0.05$) are excluded from the final model.

The response rates presented in the dataset are those provided at survey sponsor levels. Non response rates at health plan or individual provider levels are not available. Response rates for the survey are reported by 470 physician groups out of 769. The lowest reported response rate is 6%, the highest 97% and the median – 35%. Because these are self-reported by sponsors and could be biased, appropriate caution is warranted. Nonresponse weighting has been shown to have limited impact on scores once CAHPS survey data are case-mix adjusted [52-54].

For the analyses of the 0-10 *rating of the doctor* and the *would recommend doctor to friend/family* items, we initially consider to estimate ordinal logistic regression models as a sensitivity analysis in addition to the OLS models. The Brant test is used to evaluate the proportionality assumption. The proportional odds assumption is not met for 16 out of 20 variables in the *0-10 rating of the doctor* model and 13 out of 20 variables in the *would recommend doctor to friend/family* model. Therefore, generalized models (using gologit2 command in STATA) are used.

Differential response tendencies by race/ethnicity and health plan for the CAHPS 0-10 rating question are reported in earlier studies[31, 55]. If not properly addressed, response tendencies can obscure the true differences in ratings between the racial/ethnic groups. Positive response tendencies can be corrected by standard case-mix adjustments (i.e. for age, education, and race/ethnicity) in a regression analysis. Extreme response tendencies, on the other hand, are not addressed by case-mix adjustments in the presence of skewed data. Several approaches are shown to address extreme response tendencies in CAHPS datasets[28, 31, 55]. Weech-Maldonado and colleagues (2008) recommend pooling responses at the lower *(0-6)* and top end *(9-10)* when examining racial/ethnic differences in CAHPS ratings of care. In this paper, in an generalized logistic regression analysis we use this categorization approach (*0–6, 7–8, 9-10*).

Results

Cronbach's alphas are high for all three composites: "*access*" - 0.81, "*communication*" - 0.88 and "*office helpfulness*" – 0.83. Product-moment correlations among the composites show that composites measured different dimensions. The correlations range from 0.26 to 0.32. Means and standard deviations for the five dependent variables on a 0-100 scale are presented in Table 4.

Descriptive information about non-transformed composite items is presented in Tables 5 and 6. The distribution of the *0-10 rating of the doctor* is compared for Asian Americans and non-Hispanic whites. Non-Hispanic whites are more likely than Asian Americans to endorse a score of 10 (10% difference) on the global rating of the doctor item (Table 5). The majority of responses are concentrated in the upper end of the scale for both groups.

On the *would recommend doctor to friend/family* item (*yes, definitely, yes, somewhat* and *no*), Asian Americans are more likely to endorse "yes, somewhat" choice compared to non-Hispanic whites (14% vs. 8%). Non-Hispanic whites, on the other hand, are more likely to endorse "yes, definitely" option compared to Asian Americans (89% vs. 82%) (Table 6).

Table 7 presents the recycled predictions (main effects and interaction terms models). Detailed information on the estimation of recycled predictions for the models with interaction terms is provided in Appendix 5. The results from the main effects only OLS models and OLS models that include significant interaction terms of Asian Americans with other independent variables are presented in the Appendices 3 and 4. Asian Americans report the worst access *(predicted score*: Asian Americans (72.10), non-Hispanic whites (79.10), African American (79.01), Hispanic (77.78) and Other (77.97)) and lowest (worse experiences) scores *(predicted score*: Asian Americans (92.74), non-Hispanic whites (94.85), African American (95.26), Hispanic (94.55) and Other (93.60)) on office staff courtesy and helpfulness measure of all five racial/ethnic groups. Asian Americans also report the lowest scores on rating their doctors and are also the least likely to recommend their doctors to family and friends of all five racial/ethnic groups. There are no significant differences between Asian Americans and non-Hispanic whites on physician communication. The "other" racial/ethnic group reports the worst physician communication.

We are not able to explore regional variations in reports and ratings among Asian Americans in our main model due to collinearity between the practice site and region. We run a secondary model where we replace the practice site with the region and found that Asian Americans from Northeast report better experience than Asian Americans from West. Asian Americans in the South rate care worse on most measures than Asian Americans in the West, Midwest and Northeast.

A number of interactions between Asian American race/ethnicity and gender, age, education, and health are significant. For instance, the findings of the analysis show that Asian Americans who rate their health as *excellent* report better experience than Asian Americans with other self-reported health states. On access measure, Asian Americans in the 45-54 age group report worse access to care compared to Asian Americans of other ages. Asian Americans with less than high school education also have the worst access among Asian Americans of various education levels. However, the interactions are not consistent in any direction.

The findings from the generalized ordinal logistic models for the 0-10 *rating of the doctor* and the *would recommend doctor to friend/family* items are in general consistent with the OLS model results and are presented in Appendix 6.

Discussion

Our findings confirm earlier studies that have shown that Asian Americans tend to report lower scores on CAHPS composite measures and rating items than non-Hispanic whites. In our analysis, Asian Americans report the lowest scores on *access to care, office staff courtesy and helpfulness, rating of their doctors* and were the least likely *to recommend their doctors to family and friends* of the all five racial/ethnic groups. Also in contrast to earlier research, there are no

significant differences between Asian Americans and non-Hispanic whites on *physician communication* in our dataset *[25, 26, 56]*.

Differences between our study and earlier research may explain differences in findings regarding *physician communication*. For instance, Morales and colleagues (2001) found that Asian Americans report worse access and communication than non-Hispanic whites. Differences in language and communication patterns between providers and Asian American patients were suggested as a possible reason. Morales and colleagues analyzed CAHPS Health Plan survey data, while this study uses CAHPS Clinician & Group survey data. The Morales study also grouped Asian Americans together with Pacific Islanders, while in our study Asian Americans are analyzed separately. The composition of Asian Americans and their language proficiency in the two studies could also be different.

Weech-Maldonado and colleagues (2003, 2004) divided Asian Americans into two groups based on whether English was the primary language or not. On four CAHPS domains (*getting needed care, timeliness of care, provider communication*, and *staff courtesy and helpfulness*), Asian Americans for whom English was the primary language did not differ significantly from non-Hispanic whites for whom English was the primary language. Asian American non-English speakers, on the other hand, had the lowest reports of care experiences among all racial/ethnic groups. In our dataset, it is not possible to differentiate Asian Americans into English and non-English speakers. Therefore, whether a smaller number of non-English speaking Asian Americans in our data might be responsible for similar reported experiences on *physician communication* for Asian Americans and non-Hispanic whites is not possible to determine. The two studies also use different versions of the CAHPS surveys (Weech-Maldonado study used CAHPS Health Plan survey).

Our study makes a number of important contributions to the literature. Earlier studies analyzed Asian Americans and Pacific Islanders together; a sufficiently large sample of Asian Americans in our dataset allowed us to analyze Asian Americans separately from Pacific Islanders. Our findings also show that regional variations in patient experiences among Asian Americans exist. Underlying reasons for regional variations among Asian Americans in CAHPS surveys are little studied and require further research. While previous studies used various CAHPS survey instruments, our study is the first to use CG-CAHPS Adult Visit Survey 1.0 data to explore racial/ethnic differences in perceptions of care between Asian Americans and non-Hispanic whites.

Racial-ethnic disparities can be driven by differential access or selection into plans or providers of differing quality. However, a number of studies report that "within provider" differences account for the significant share of disparities between Asian Americans and non-Hispanic whites [23, 34, 57]. In our study, we control for the "between providers" effects by including in our model provider identifications.

Our analysis presents some limitations. It is of note that Asian Americans are heterogeneous group themselves and previous research found that different Asian American subgroups may vary significantly in their reports and ratings of care [21, 37]. Regional differences among Asian Americans found in our secondary model analyses may be in part explained by differential regional concentrations of Asian American subgroups. English proficiency and acculturation may also have an effect on reports and rating of care. Unfortunately, the standard CAHPS survey race question does not distinguish between subtypes of Asian Americans (e.g., Chinese, Japanese, Vietnamese) and does not have items on English proficiency and acculturation.

Although the CAHPS is widely used throughout the US, participation in the CAHPS database is voluntary, thus the representativeness of the findings from this study is unknown. Low response rate presents a challenge in generalizing the findings from this study.

Differences in expectations from care can also lead to variations in how patients perceive their care experiences. Expectations from care are often shaped by previous experiences with care that could also be provided by doctors other than the patient's usual care provider. Poor quality care received can lead to lower expectations and thus inflated reports and ratings of care or vice versa. The CAHPS survey tool does not have items that capture previous patient experiences and in particular, with regard to care received from doctors other than a regular care provider.

Our analysis is the first study that explored for differences between Asian Americans and non-Hispanic whites in the CAHPS Clinician & Group Adult Visit 1.0 Survey data. The study findings showed that Asian Americans report and rate their experience with care lower than non-Hispanic whites. The differences in care experiences between Asian Americans and non-Hispanic whites are notable as small if any differences could be found between these two groups in utilization rates or quality of care received. Asian Americans are shown also to be comparable to Whites in rates of education and income. Differential response tendencies between Asian Americans and non-Hispanic whites may explain the observed differences in reports and ratings of care experiences. Setodji et al. (2011), for instance, explored the CAHPS Medicare survey data and found evidence for differential item functioning (DIF) for 3 out of 9 items between the surveys completed by Hispanics in English and Spanish [38]. Therefore, further research is needed to explore whether these differences found are due to cultural norms that lead to different

response tendencies. If differential response tendencies are not found, then future research is need to explore the reasons for the differential treatment.

Findings from this study and future research exploring differential response tendencies should inform and guide activities targeting to reduce racial/ethnic disparities in patient experiences with care. If differential response tendencies are found between Asian Americans and non-Hispanic whites, they may explain the differences in care experiences. Further efforts may be needed to understand the causes of differential response tendencies, so that CG-CAHPS Adult Visit Survey 1.0 items, data analysis and reporting practices can be improved to adjust for the differential response tendencies. Otherwise, targeted quality improvement initiatives will have to be developed to address differences in care provided to Asian Americans and non-Hispanic whites.

Tables and figures

Table 2.1 Reports about care and global rating items

Access to care 1. In the last 12 months, when you phoned this doctor's office to get an appointment for care you needed right away, how often did you get an appointment as soon as you thought you needed? 2. In the last 12 months, when you made an appointment for a check-up or routine care with this doctor, how often did you get an appointment as soon as you thought you needed? 3. In the last 12 months, when you phoned this doctor's office during regular office hours, how often did you get an answer to your medical question that same day? 4. In the last 12 months, when you phoned this doctor's office after regular office hours, how often did you get an answer to your medical question as soon as you needed? 5. Wait time includes time spent in the waiting room and exam room. In the last 12 months, how often did you see this doctor within 15 minutes of your appointment time?
Physician communication 1. During your most recent visit, did this doctor explain things in a way that was easy to understand? 2. During your most recent visit, did this doctor listen carefully to you? 3. During your most recent visit, did this doctor give you easy to understand instructions about taking care of these health problems or concerns? 4. During your most recent visit, did this doctor seem to know the important information about your medical history? 5. During your most recent visit, did this doctor show respect for what you had to say? 6. During your most recent visit, did this doctor spend enough time with you?

Table 2.1 Reports about care and global rating items (continued)

Office staff courtesy and helpfulness 1. During your most recent visit, were clerks and receptionists at this doctor's office as helpful as you thought they should be? 2. During your most recent visit, did clerks and receptionists at this doctor's office treat you with courtesy and respect?
Doctor rating • Using any number from 0 to 10, where 0 is the worst doctor possible and 10 is the best doctor possible, what number would you use to rate this doctor?
Recommending doctor to friends/family • Would you recommend this doctor's office to your family and friends?

Note: *Physician communication* and *office staff courtesy and helpfulness* composite items and the *recommending doctor to friends/family* question use a three response option scale (*Yes, definitely; Yes, somewhat; No*). *Access to care* composite items use a four response option scale (*Always, Usually, Sometimes, Never*). The *doctor rating* question uses a 0-10 response scale.

Table 2.2 Key sample demographic information, AHPS-CG Adult Visit Data from 2011

Gender		
	Percent	**Frequency**
Male	36.15	94,068
Female	63.85	160,236
Total	**100.00**	**260,236**
Missing values		6,091
Age		
18-24	2.83	7,364
25-34	7.96	20,733
35-44	9.29	24,207
45-54	16.93	44,118
55-64	24.51	63,872
65-74	21.44	55,872
75+	17.06	44,455
Total	**100.00**	**260,621**
Missing values		**5,706**
Education		
8th grade or less	2.47	6,329
Less than HS grad	4.62	11,845
HS grad (includes GED)	25.15	64,475
Some college	30.23	77,511
4-year grad	17.43	44,681
More than 4 years	20.10	51,542
Total	**100.00**	**256,383**
Missing values		**9,944**
Race/Ethnicity		
White	89.49	229,899
African-American	4.30	11,057
Asian	2.16	5,545
Native Hawaiian/Pacific Islander	0.13	338
American Indian/Native Alaskan	0.32	839
Other	2.17	5,575
Multi-racial	1.42	3,652
Total	**100.00**	**256,885**
Missing values		**9,442**

Table 2.2 Key sample demographic information, CAHPS-CG Adult Visit Data from 2011 (continued)

Hispanic/Latino origin or descent		
	Percent	**Frequency**
Yes	4.95	12,475
No	95.05	239,722
Total	**100.00**	**252,197**
Missing values		**14,130**
Health status		
Excellent	12.60	32,747
Very Good	34.41	89,428
Good	33.99	88,353
Fair	15.25	39,625
Poor	3.75	9,752
Total	**100.00**	**259,905**
Missing values		**6,422**
Practice regions		
Midwest	52.48	139,775
Northeast	22.93	61,056
South	10.08	26,843
West	14.51	38,653
Total	**100.00**	**266,327**
Missing values		**6,422**
Reported experiences with		
Family medicine or internal medicine doctors	49	45,727
Obstetricians/gynecologists	7	7,093
Surgeons	5	4,378
Other specialty doctors	39	36,714
Total	**100.00**	**93,912**
Missing values		**172,415**

Table 2.3 Gender, Age, Education, Health and Region for Asian Americans and non-Hispanic whites

	Asian	White
Gender		
	n = 5,119	*n = 216,041*
Male	34.64	36.45
Female	65.36	63.55
Age		
	n = 5,103	*n = 216,405*
18 to 24	5.13	2.71
25 to 34	17.56	7.69
35 to 44	17.42	8.76
45 to 54	16.58	16.65
55 to 64	18.87	24.81
65 to 74	13.70	21.93
75 and older	10.74	17.46
Education		
	n = 5,053	*n = 214,604*
8th grade or less	8.85	1.45
Some high school	4.61	3.59
High school graduate	12.29	24.94
Some college	19.87	30.63
4-year college graduate	24.70	18.33
> 4-year college	29.69	21.06
Health		
	n = 5,112	*n = 215,090*
Poor	3.72	3.43
Fair	12.77	14.23
Good	32.45	33.97
Very good	35.45	35.73
Excellent	15.61	12.64
Region		
	n = 5,170	*n = 217,458*
Midwest	48.57	55.62
Northeast	5.65	25.12
South	5.32	7.50
West	40.46	11.76

Note: Percentages do not always add up to 100.00 because of rounding. Chi square statistic was significant at p = 0.05 level for all variables in the table.

Table 2.4 Means and standard deviations for the five dependent variables

Composites	Total sample	Asian-Americans	non-Hispanic whites
Access			
Mean	78.69	68.76	79.89
Std. Dev.	23.99	26.74	23.09
N	263,737	5,098	215,719
Communication			
Mean	94.15	93.11	94.47
Std. Dev.	14.69	15.22	14.16
N	265,783	5,144	217,101
Office			
Mean	94.70	90.95	95.09
Std. Dev.	15.67	19.44	15.06
N	258,785	5,091	213,493
Rating a doctor			
Mean	91.25	88.60	91.41
Std. Dev.	14.16	15.47	13.74
N	263,216	5,089	215,535
Recommending to friends/family			
Mean	92.93	89.36	93.32
Std. Dev.	20.92	24.45	20.31
N	257,813	5,056	212,384

Note: All were scaled on a 0-100 possible range and the observed minimum and maximum were 0 and 100 for each variable.

Table 2.5 Distribution of the *rating a doctor* item (0-10 scale), Asians and Whites

	Asians		Whites	
	Frequency	Percent	Frequency	Percent
0: worst doctor possible	16	0.31	466	0.22
1	16	0.31	422	0.20
2	21	0.41	601	0.28
3	31	0.61	955	0.44
4	24	0.47	930	0.43
5	102	2.00	3,253	1.51
6	110	2.16	2,783	1.29
7	339	6.66	8,388	3.89
8	899	17.67	29,000	13.45
9	1,203	23.64	49,079	22.77
10: best doctor possible	2,328	45.75	119,658	55.52
Total	**5,089**	**100**	**215,535**	**100**

Table 2.6 Distribution of the *would recommend doctor to friend/family* item (1-3 scale), Asians and Whites

	Asians		Whites	
	Frequency	Percent	Frequency	Percent
No	181	3.58	5,233	2.46
Yes, somewhat	714	14.12	17,891	8.42
Yes, definitely	4,161	82.3	189,260	89.11
Total	**5,056**	**100**	**212,384**	**100**

Table 2.7 Recycled predictions from the main effects only model and the model with the interaction terms

Predicted scores from the main effects model										
	Access		**Communicate**		**Office**		**Rate**		**Recommend**	
	Margin	SE	Margin	SE	Margin	SE	Margin	SE	Margin	SE
Race/ethnicity										
White	79.10	0.03	94.31	0.02	94.85	0.02	91.30	0.02	93.11	0.02
Asian American	72.38	0.43	94.03	0.23	92.74	0.32	89.79	0.27	91.01	0.37
Hispanic	77.78	0.36	94.12	0.22	94.55	0.21	91.91	0.19	93.43	0.28
African American	79.01	0.45	94.55	0.25	95.26	0.19	91.95	0.30	93.58	0.40
Other	77.97	0.20	93.03	0.14	93.60	0.17	90.78	0.13	91.68	0.20
Predicted Scores from the model with the interaction terms										
Asian x Health (Poor)	72.38		94.03		92.04		89.79		91.01	
Asian x Fair	72.38		94.03		92.04		89.79		91.01	
Asian x Good	72.38		94.03		92.04		89.79		91.01	
Asian x Very Good	72.38		94.03		92.04		89.79		91.01	
Asian x Excellent	72.38		94.03		93.79		89.79		91.01	
Asian x Age (18-24)	72.76		94.03		92.74		89.79		91.01	
Asian x 25-34	72.76		94.03		92.74		89.79		91.01	
Asian x 35-44	72.76		94.03		92.74		89.79		91.01	
Asian x 45-54	71.25		94.03		92.74		89.79		91.01	
Asian x 55-64	72.76		94.03		92.74		89.79		91.01	
Asian x 65-74	72.76		94.03		92.74		89.79		91.01	
Asian x 75 & over	72.76		94.03		92.74		89.79		91.01	
Asian x Edu (8th Grade)	74.61		94.03		92.74		89.79		91.01	
Asian x Less Than HS	70.44		94.03		92.74		89.79		91.01	
Asian x HS	72.10		94.03		92.74		89.79		91.01	
Asian x Less Than Col	74.61		94.03		92.74		89.79		91.01	
Asian x Col	74.61		94.03		92.74		89.79		91.01	
Asian x Col Grad	74.61		94.03		92.74		89.79		91.01	

Note: HS – high school; Col – college; Col Grad – graduate education

Appendices

Appendix 2.1 Reports about care and global rating items: total number of responses and missing values

	% Not Answered	% Appropriately Skipped	% Total Missing	N
Access to care				
Getting care quickly when urgent	2.23	55.93	58.16	260,522
Getting care quickly when not urgent	3.22	26.75	29.97	258,018
Getting an answer to a medical question the same day when calling during regular hours	3.75	59.10	62.84	256,710
Getting an answer to a medical question quickly when calling after regular hours	5.77	92.63	98.40	251,794
Seeing a doctor within 15 minutes of appointment time	1.91	N/A	1.91	261,329
Physician communication				
Doctor explained things clearly	5.27	N/A	5.27	252,986
Doctor listened carefully	2.84	N/A	2.84	258,962
Doctor gave clear instructions	3.94	14.65	18.59	261,790
Doctor knew important information about medical history	1.20	N/A	1.20	263,164
Doctor showed respect	2.73	N/A	2.73	259,257
Doctor spent enough time	2.69	N/A	2.69	259,357
Office staff helpfulness				
Doctor office staff helpful	3.17	N/A	3.17	258,147
Courtesy and respect from doctor office staff	3.38	N/A	3.38	257,609
Global rating items				
Rating a doctor	1.18	N/A	1.18	263,216
Recommending a doctor to friends/family	3.30	N/A	3.30	257,813

Appendix 2.2 Estimation of recycled predictions for the models with interaction terms

To estimate recycled predictions for the model with significant interaction terms, a recycled prediction from the main effects only model was multiplied by the total number of dummy variables (1); regression coefficients of significant dummy variables were then subtracted from this value (2), the value was divided by the total number of dummy variables (3). We then added the regression coefficient of the interaction term for which recycled prediction was being estimated. All non-significant interaction terms were grouped together in estimation of the interaction terms models. The following provides an example of how recycled prediction was estimated for Asian Americans with less than high school education on Access composite:

(((Recycled prediction from the main effects model * number of dummy variables)-regression coefficients for dummy variables)/total number of dummy variables) + regression coefficients for Asian Americans with less than high school education on Access

Appendix 2.3 Ordinary least square regression: main effects only model

	Access (R^2-0.09)			Communicate (R^2-0.05)			Office (R^2-0.05)			Rate (R^2-0.05)			Recommend (R^2-0.04)		
	Coef.	Std. Err.	t	Coef.	Std. Err.	t	Coef.	Std. Err.	t	Coef.	Std. Err.	t	Coef.	Std. Err.	t
Race/ethnicity (White)															
Asian American	**-6.72**	**0.44**	**-15.32**	-0.28	0.24	-1.20	**-2.11**	**0.33**	**-6.43**	**-1.51**	**0.27**	**-5.56**	**-2.10**	**0.38**	**-5.55**
Hispanic	**-1.32**	**0.39**	**-3.40**	-0.19	0.23	-0.82	-0.29	0.22	-1.33	**0.60**	**0.20**	**3.06**	0.32	0.30	1.08
African American	-0.09	0.47	-0.19	0.24	0.26	0.92	**0.41**	**0.19**	**2.12**	**0.64**	**0.32**	**2.04**	0.47	0.41	1.15
Other	**-1.14**	**0.21**	**-5.33**	**-1.28**	**0.14**	**-8.89**	**-1.24**	**0.18**	**-6.84**	**-0.52**	**0.14**	**-3.76**	**-1.43**	**0.21**	**-6.79**
Health (Poor)															
Fair health	**0.64**	**0.30**	**2.12**	**2.23**	**0.26**	**8.62**	**0.91**	**0.22**	**4.14**	**1.42**	**0.24**	**6.03**	**2.52**	**0.31**	**8.01**
Good health	**2.43**	**0.29**	**8.25**	**4.21**	**0.27**	**15.72**	**1.64**	**0.23**	**7.08**	**2.70**	**0.24**	**11.12**	**4.28**	**0.34**	**12.61**
Very good health	**4.78**	**0.29**	**16.45**	**5.84**	**0.27**	**21.38**	**2.65**	**0.24**	**11.24**	**4.28**	**0.25**	**17.35**	**6.21**	**0.34**	**18.04**
Excellent health	**8.18**	**0.32**	**25.92**	**7.57**	**0.29**	**26.22**	**4.24**	**0.25**	**17.01**	**6.86**	**0.26**	**26.33**	**8.37**	**0.36**	**23.25**
Age (18-24)															
Age 25-34	**0.76**	**0.36**	**2.11**	**1.30**	**0.28**	**4.65**	**1.30**	**0.30**	**4.37**	**1.35**	**0.24**	**5.66**	**0.74**	**0.35**	**2.11**
Age 35-44	**2.67**	**0.36**	**7.47**	**3.38**	**0.28**	**12.28**	**2.98**	**0.28**	**10.48**	**3.31**	**0.24**	**13.94**	**3.20**	**0.34**	**9.48**
Age 45-54	**4.83**	**0.34**	**14.18**	**4.40**	**0.26**	**16.70**	**4.25**	**0.30**	**14.27**	**4.50**	**0.23**	**19.58**	**4.72**	**0.34**	**13.88**
Age 55-64	**6.63**	**0.37**	**17.70**	**5.81**	**0.27**	**21.25**	**5.97**	**0.29**	**20.46**	**6.01**	**0.24**	**25.19**	**6.32**	**0.35**	**18.07**
Age 65-74	**8.55**	**0.42**	**20.41**	**6.84**	**0.27**	**25.18**	**7.39**	**0.31**	**24.22**	**7.28**	**0.23**	**31.63**	**7.41**	**0.34**	**22.02**
Age 75 and over	**8.18**	**0.45**	**18.23**	**6.79**	**0.29**	**23.69**	**8.40**	**0.33**	**25.47**	**7.68**	**0.25**	**30.47**	**7.58**	**0.36**	**20.89**
Education (8th Grade)															
Edu Less Than HS	**0.92**	**0.38**	**2.40**	-0.27	0.31	-0.88	**0.73**	**0.30**	**2.44**	0.21	0.27	0.79	-0.41	0.40	-1.02
Edu HS	0.62	0.42	1.47	**-0.47**	**0.22**	**-2.11**	0.29	0.25	1.15	**-0.66**	**0.20**	**-3.27**	**-1.15**	**0.32**	**-3.58**
Edu Less Than Col	0.45	0.40	1.11	**-0.76**	**0.22**	**-3.40**	**-0.74**	**0.26**	**-2.86**	**-1.39**	**0.21**	**-6.69**	**-1.72**	**0.34**	**-5.02**
Edu Col	-0.06	0.41	-0.15	**-0.92**	**0.22**	**-4.17**	**-0.88**	**0.27**	**-3.30**	**-2.54**	**0.21**	**-11.94**	**-2.41**	**0.32**	**-7.60**
Edu Col Grad	-0.07	0.39	-0.18	**-1.06**	**0.22**	**-4.85**	**-1.37**	**0.28**	**-4.85**	**-2.69**	**0.21**	**-12.76**	-2.72	0.33	-8.21
Gender															
Male	**1.35**	**0.13**	**10.64**	**0.78**	**0.07**	**11.07**	**0.68**	**0.08**	**8.37**	**0.18**	**0.08**	**2.30**	**0.84**	**0.11**	**7.81**
Constant	**65.37**	**0.60**	**109.74**	**84.08**	**0.41**	**202.91**	**88.87**	**0.48**	**185.43**	**82.44**	**0.36**	**230.94**	**84.26**	**0.55**	**152.46**

Note: HS – high school; Col – college. Values with p-values at less or equal to 0.05 are in bold font.

Appendix 2.4 Ordinary least square regression: model with interaction terms of Asians and other case-mix variables

	Access (R^2-0.09)			Communicate (R^2-0.05)			Office (R^2-0.05)			Rate (R^2-0.05)			Recommend (R^2-0.04)		
	Coef.	Std. Err.	t	Coef.	Std. Err.	t	Coef.	Std. Err.	t	Coef.	Std. Err.	t	Coef.	Std. Err.	t
Race-ethnicity (White)															
Asians	**-5.85**	**0.53**	**-11.03**	-0.28	0.24	-1.20	**-2.44**	**0.37**	**-6.51**	**-1.51**	**0.27**	**-5.56**	**-2.10**	**0.38**	**-5.55**
Hispanic	**-1.33**	**0.39**	**-3.42**	-0.19	0.23	-0.82	-0.29	0.22	-1.32	**0.60**	**0.20**	**3.06**	0.32	0.30	1.08
African American	-0.10	0.47	-0.22	0.24	0.26	0.92	**0.41**	**0.19**	**2.11**	**0.64**	**0.32**	**2.04**	0.47	0.41	1.15
Other	**-1.14**	**0.21**	**-5.37**	**-1.28**	**0.14**	**-8.89**	**-1.25**	**0.18**	**-6.84**	**-0.52**	**0.14**	**-3.76**	**-1.43**	**0.21**	**-6.79**
Health (Poor)															
Fair health	**0.64**	**0.30**	**2.13**	**2.23**	**0.26**	**8.62**	**0.90**	**0.22**	**4.14**	**1.42**	**0.24**	**6.03**	**2.52**	**0.31**	**8.01**
Good health	**2.44**	**0.29**	**8.27**	**4.21**	**0.27**	**15.72**	**1.64**	**0.23**	**7.09**	**2.70**	**0.24**	**11.12**	**4.28**	**0.34**	**12.61**
Very good health	**4.79**	**0.29**	**16.48**	**5.84**	**0.27**	**21.38**	**2.65**	**0.23**	**11.26**	**4.28**	**0.25**	**17.35**	**6.21**	**0.34**	**18.04**
Excellent health	**8.19**	**0.32**	**25.94**	**7.57**	**0.29**	**26.22**	**4.19**	**0.25**	**17.04**	**6.86**	**0.26**	**26.33**	**8.37**	**0.36**	**23.25**
Age (18-24)															
Age 25-34	**0.74**	**0.36**	**2.05**	**1.30**	**0.28**	**4.65**	**1.30**	**0.30**	**4.37**	**1.35**	**0.24**	**5.66**	**0.74**	**0.35**	**2.11**
Age 35-44	**2.65**	**0.36**	**7.41**	**3.38**	**0.28**	**12.28**	**2.98**	**0.28**	**10.48**	**3.31**	**0.24**	**13.94**	**3.20**	**0.34**	**9.48**
Age 45-54	**4.86**	**0.34**	**14.20**	**4.40**	**0.26**	**16.70**	**4.25**	**0.30**	**14.27**	**4.50**	**0.23**	**19.58**	**4.72**	**0.34**	**13.88**
Age 55-64	**6.63**	**0.38**	**17.64**	**5.81**	**0.27**	**21.25**	**5.98**	**0.29**	**20.45**	**6.01**	**0.24**	**25.19**	**6.32**	**0.35**	**18.07**
Age 65-74	**8.54**	**0.42**	**20.36**	**6.84**	**0.27**	**25.18**	**7.39**	**0.31**	**24.22**	**7.28**	**0.23**	**31.63**	**7.41**	**0.34**	**22.02**
Age 75 and over	**8.17**	**0.45**	**18.17**	**6.79**	**0.29**	**23.69**	**8.40**	**0.33**	**25.46**	**7.68**	**0.25**	**30.47**	**7.58**	**0.36**	**20.89**
Education (8th Grade)															
Edu Less Than HS	**1.03**	**0.38**	**2.68**	-0.27	0.31	-0.88	**0.72**	**0.30**	**2.42**	0.21	0.27	0.79	-0.41	0.40	-1.02
Edu HS	0.67	0.42	1.60	**-0.47**	**0.22**	**-2.11**	0.28	0.25	1.12	**-0.66**	**0.20**	**-3.27**	**-1.15**	**0.32**	**-3.58**
Edu Less Than Col	0.47	0.40	1.17	**-0.76**	**0.22**	**-3.40**	**-0.75**	**0.26**	**-2.90**	**-1.39**	**0.21**	**-6.69**	**-1.72**	**0.34**	**-5.02**
Edu Col	-0.05	0.41	-0.11	**-0.92**	**0.22**	**-4.17**	**-0.89**	**0.27**	**-3.34**	**-2.54**	**0.21**	**-11.94**	**-2.41**	**0.32**	**-7.60**
Edu Col Grad	-0.05	0.39	-0.14	**-1.06**	**0.22**	**-4.85**	**-1.37**	**0.28**	**-4.88**	**-2.69**	**0.21**	**-12.76**	-2.72	0.33	-8.21
Gender															
Male	**1.35**	**0.13**	**10.64**	**0.78**	**0.07**	**11.07**	**0.68**	**0.08**	**8.36**	**0.18**	**0.08**	**2.30**	**0.84**	**0.11**	**7.81**
Constant	**65.33**	**0.60**	**109.66**	**84.08**	**0.41**	**202.91**	**88.89**	**0.48**	**186.18**	**82.44**	**0.36**	**230.94**	**84.26**	**0.55**	**152.46**

Note: HS – high school; Col – college; Col Grad – graduate education. Values with p-values at less or equal to 0.05 are in bold font.

Appendix 2.4 Ordinary least square regression: model with interaction terms of Asians and other case-mix variables (continued)

	Access (R^2-0.09)			**Communicate (R^2-0.05)**			**Office (R^2-0.05)**			**Rate (R^2-0.05)**			**Recommend (R^2-0.04)**		
	Coef.	Std. Err.	t	Coef.	Std. Err.	t	Coef.	Std. Err.	t	Coef.	Std. Err.	t	Coef.	Std. Err.	t
Asian x Health (Poor)															
Asian x Fair	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian x Good	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian x Very Good	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian x Excellent	-	-	-	-	-	-	**2.11**	**0.70**	**3.02**	-	-	-	-	-	-
Asian x Age (18-24)															
Asian x 25-34	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian x 35-44	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian x 45-54	**-2.27**	**1.02**	**-2.23**	-	-	-	-	-	-	-	-	-	-	-	-
Asian x 55-64	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian x 65-74	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian x 75 & over	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian x Edu (8th Grade)															
Asian x Less Than HS	**-4.17**	**1.95**	**-2.13**	-	-	-	-	-	-	-	-	-	-	-	-
Asian x HS	**-2.51**	**1.27**	**-1.97**	-	-	-	-	-	-	-	-	-	-	-	-
Asian x Less Than Col	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian x Col	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian x Col Grad	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Asian x Male	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Constant	**65.33**	**0.60**	**109.66**	**84.08**	**0.41**	**202.91**	**88.89**	**0.48**	**186.18**	**82.44**	**0.36**	**230.94**	**84.26**	**0.55**	**152.46**

Note: HS – high school; Col – college; Col Grad – graduate education. Values with p-values at less or equal to 0.05 are in bold font.

Appendix 2.5 Findings form generalized ordinal logistic regression (gologit2)

	Rate (1) (χ^2-0.00)			Rate (2) (χ^2-0.00)				Recommend (1) (χ^2-0.00)			Recommend (2) (χ^2-0.00)		
	Coef.	Std. Err.	z	Coef.	Std. Err.	z		Coef.	Std. Err.	z	Coef.	Std. Err.	z
Race-ethnicity (White)													
Asians	**-0.19**	**0.07**	**-2.80**	**-0.32**	**0.05**	**-6.99**		-0.15	0.09	-1.63	**-0.39**	**0.04**	**-8.94**
Hispanic	**-0.27**	**0.06**	**-4.52**	-0.04	0.04	-0.93		**-0.26**	**0.06**	**-4.42**	**-0.26**	**0.06**	**-4.42**
African American	0.07	0.04	1.62	0.07	0.04	1.62		0.01	0.05	0.29	0.01	0.05	0.29
Other	**-0.20**	**0.04**	**-5.38**	**-0.05**	**0.02**	**-2.23**		**-0.34**	**0.05**	**-6.98**	**-0.21**	**0.03**	**-7.15**
Health (Poor)													
Fair health	**0.31**	**0.04**	**7.60**	**0.06**	**0.03**	**2.18**		**0.58**	**0.05**	**10.69**	**0.18**	**0.03**	**5.45**
Good health	**0.70**	**0.04**	**15.79**	**0.22**	**0.03**	**7.10**		**1.02**	**0.06**	**18.46**	**0.41**	**0.04**	**10.71**
Very good health	**1.17**	**0.05**	**23.82**	**0.50**	**0.03**	**15.19**		**1.48**	**0.06**	**26.01**	**0.73**	**0.04**	**18.26**
Excellent health	**1.49**	**0.06**	**25.99**	**1.01**	**0.04**	**27.74**		**1.70**	**0.07**	**24.41**	**1.20**	**0.04**	**27.38**
Age (18-24)													
Age 25-34	**0.18**	**0.03**	**6.37**	**0.18**	**0.03**	**6.37**		0.04	0.05	0.92	**0.11**	**0.04**	**3.03**
Age 35-44	**0.43**	**0.03**	**14.46**	**0.43**	**0.03**	**14.46**		**0.37**	**0.04**	**9.55**	**0.37**	**0.04**	**9.55**
Age 45-54	**0.66**	**0.04**	**17.21**	**0.59**	**0.03**	**19.15**		**0.56**	**0.04**	**14.57**	**0.56**	**0.04**	**14.57**
Age 55-64	**0.96**	**0.04**	**22.79**	**0.81**	**0.03**	**24.40**		**0.93**	**0.05**	**18.63**	**0.78**	**0.04**	**19.44**
Age 65-74	**1.23**	**0.04**	**28.84**	**1.03**	**0.03**	**31.87**		**1.22**	**0.06**	**22.03**	**1.00**	**0.04**	**25.92**
Age 75 and over	**1.22**	**0.05**	**25.67**	**1.02**	**0.04**	**27.72**		**1.34**	**0.06**	**21.02**	**0.98**	**0.04**	**22.24**
Education (8th Grade)													
Edu Less Than HS	0.06	0.05	1.29	0.06	0.05	1.29		0.01	0.07	0.11	0.01	0.07	0.11
Edu HS	0.08	0.05	1.66	**-0.09**	**0.04**	**-2.29**		-0.10	0.06	-1.81	-0.10	0.06	-1.81
Edu Less Than Col	0.06	0.05	1.17	**-0.12**	**0.04**	**-2.82**		**-0.25**	**0.07**	**-3.64**	**-0.13**	**0.06**	**-2.19**
Edu Col	0.07	0.06	1.30	**-0.30**	**0.04**	**-7.21**		**-0.24**	**0.06**	**-4.16**	**-0.24**	**0.06**	**-4.16**
Edu Col Grad	0.03	0.06	0.47	**-0.29**	**0.04**	**-6.51**		**-0.38**	**0.07**	**-5.33**	**-0.27**	**0.06**	**-4.41**
Gender													
Male	**0.23**	**0.03**	**9.10**	0.00	0.02	-0.17		**0.17**	**0.03**	**5.12**	**0.10**	**0.02**	**4.90**
Constant	**1.29**	**0.07**	**17.52**	**0.35**	**0.06**	**5.74**		**1.95**	**0.08**	**23.25**	**1.00**	**0.09**	**11.56**

References

1. Rosenthal, G. and S. Shannon, *The Use of Patient Perceptions in the Evaluation of Health-Care Delivery Systems.* Medical care, 1997. **35**(11): p. NS58-NS68.

2. Cleary, P. and S. Edgman-Levitan, *Health care quality: Incorporating consumer perspectives.* JAMA: The Journal of the American Medical Association, 1997. **278**(19): p. 1608-1612.

3. Sitzia, J. and N. Wood, *Patient satisfaction: A review of issues and concepts.* Social Science & Medicine, 1997. **45**(12): p. 1829-1843.

4. Hays, R.D., *Patient Satisfaction. In Kattan M.W. (ed.), Encyclopedia of medical decision making*. Vol. 1. 2009, Thousand Oaks Sage Publications.

5. CMS. *Centers for Medicare & Medicaid Services: Overview*. 2012 [cited 2012 August 13]; Available from: http://www.Medicare.gov/choices/Overview.asp.

6. Zaslavsky, A., et al., *How Consumer Assessments of Managed Care Vary Within and Among Markets.* Inquiry, 2000. **37**(2): p. 146-161.

7. Landon, B., et al., *COmparison of performance of traditional medicare vs medicare managed care.* JAMA: The Journal of the American Medical Association, 2004. **291**(14): p. 1744-1752.

8. AHRQ. *Agency for Healthcare Research and Qualty: About CAHPS*. 2012 [cited 2012 August 13]; Available from: http://www.cahps.ahrq.gov/about.htm.

9. McGee, J., et al., *Making Survey Results Easy to Report to Consumers How Reporting Needs Guided Survey Design in CAHPS.* Medical Care, 1999. **37**(3): p. MS32-MS40.

10. Quigley, D.D., et al., *Use of CAHPS® patient experience survey data as part of a patient-centered medical home quality improvement initiative.* Journal of Healthcare Leadership, 2015. **7**: p. 41-54.

11. Faber, M., Bosch, M., Wollersheim, H., Leatherman, Sh., Grol, R., *Public Reporting in Health Care: How Do Consumers Use Quality-of-Care Information?: A Systematic Review.* Medical care, 2009. **47**(1): p. 1-8 10.1097/MLR.0b013e3181808bb5.

12. Farley, D., et al., *Effect of CAHPS Performance Information on Health Plan Choices by Iowa Medicaid Beneficiaries.* Medical Care Research and Review, 2002. **59**(3): p. 319-336.

13. Browne, K., D. Roseman, and D. Shaller, Edgman-Levitan, S., *Analysis & commentary: Measuring Patient Experience As A Strategy For Improving Primary Care.* Health Affairs, 2010. **29**(5): p. 921-925.

14. Schneider, E.C. and T. Lieberman, *Publicly disclosed information about the quality of health care: response of the US public.* Quality in Health Care, 2001. **10**(2): p. 96-103.

15. Fung, C., Lim, Y., Mattke, S., Damberg, Ch., Shekelle, P., *Systematic Review: The Evidence That Publishing Patient Care Performance Data Improves Quality of Care.* Annals of Internal Medicine, 2008. **148**(2): p. 111-123.

16. U.S. Census Bureau. *United States Census 2010*. 2011 [cited 2016 August 20]; Available from: http://2010.census.gov/news/releases/operations/cb11-cn125.html

17. Snyder, R.E., et al., *Access to Medical Care Reported by Asians and Pacific Islanders in a West Coast Physician Group Association.* Medical Care Research and Review, 2000. **57**(2): p. 196-215.

18. Phillips, K., M. Mayer, and L. Aday, *Barriers to care among racial/ethnic groups under managed care.* Health Affairs, 2000. **19**(4): p. 65-75.

19. Meredith, L. and A. Siu, *Variation and Quality of Self-Report Health Data: Asians and Pacific Islanders Compared with Other Ethnic Groups.* Medical Care, 1995. **33**(11): p. 1120-1131.

20. Haviland, M., et al., *Do Health Care Ratings Differ by Race or Ethnicity?* Joint Commission Journal on Quality and Patient Safety, 2003. **29**(3): p. 134-145.

21. Murray-García, J., J. Selby, and J. Schmittdiel, Grumbach, K., Quesenberry, Ch., *Racial and Ethnic Differences in a Patient Survey: Patients' Values, Ratings, and Reports Regarding Physician Primary Care Performance in a Large Health Maintenance Organization.* Medical Care, 2000. **38**(3): p. 300-310.

22. Ngo-Metzger, Q., A. Legedza, and R. Phillips, *Asian Americans' Reports of Their Health Care Experiences.* Journal of General Internal Medicine, 2004. **19**(2): p. 111-119.

23. Rodriguez, H., T. von Glahn, and D. Grembowski, Rogers, W., Safran, D., *Physician Effects on Racial and Ethnic Disparities in Patients' Experiences of Primary Care.* Journal of General Internal Medicine, 2008. **23**(10): p. 1666-1672.

24. Morales, L., W. Cunningham, and J. Brown, Liu, H., Hays, R., *Are Latinos Less Satisfied with Communication by Health Care Providers?* Journal of General Internal Medicine, 1999. **14**(7): p. 409-417.

25. Morales, L., 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 Services Research, 2001. **36**(3): p. 595-617.

26. Weech-Maldonado, R., et al., *Race/Ethnicity, Language, and Patients' Assessments of Care in Medicaid Managed Care.* Health Services Research, 2003. **38**(3): p. 789-808.

27. Haviland, A., M. Elliott, and R. Weech-Maldonado, Hambarsoomian, K., Orr, N., & Hays, R. D. , *Racial/ethnic disparities in Medicare Part D experiences.* Medical Care, Supplement, 2012. **59**.

28. Elliott, M., A. Haviland, and D. Kanouse, Hambarsoomian, K., & Hays, R., *Adjusting for Subgroup Differences in Extreme Response Tendency in Ratings of Health Care: Impact on Disparity Estimates.* Health Services Research, 2009. **44**(2p1): p. 542-561.

29. Fongwa, M., W. Cunningham, and R. Weech-Maldonado, Gutierrez, P., Hays, R., *Reports and ratings of care: Black and White Medicare enrollees.* Journal of Health Care for the Poor and Underserved, 2008. **19**(4): p. 1136-1147.

30. Weech-Maldonado, R., M. Fongwa, and P. Gutierrez, & Hays, R., *Language and Regional Differences in Evaluations of Medicare Managed Care by Hispanics.* Health Services Research, 2008. **43**(2): p. 552-568.

31. Weech-Maldonado, R., M. Elliott, and A. Oluwole, Schiller, K., & Hays, R., *Survey Response Style and Differential Use of CAHPS Rating Scales by Hispanics.* Medical Care, 2008. **46**(9): p. 963-968 10.1097/MLR.0b013e3181791924.

32. Fongwa, M., W. Cunningham, and R. Weech-Maldonado, Gutierrez, P., Hays, R. , *Comparison of Data Quality for Reports and Ratings of Ambulatory Care by African American and White Medicare Managed Care Enrollees.* Journal of Aging and Health, 2006. **18**(5): p. 707-721.

33. Hurtado, M., J. Angeles, and S. Blahut, & Hays, R., *Assessment of the Equivalence of the Spanish and English Versions of the CAHPS® Hospital Survey on the Quality of Inpatient Care.* Health Services Research, 2005. **40**(6p2): p. 2140-2161.

34. Weech-Maldonado, R., et al., *Health Plan Effects on Patient Assessments of Medicaid Managed Care Among Racial/Ethnic Minorities.* Journal of General Internal Medicine, 2004. **19**(2): p. 136-145.

35. Mayer L, et al., *Less Use of Extreme Response Options by Asians to Standardized Care Scenarios may Explain Some Racial/Ethnic Differences in CAHPS Scores.* Medical Care, In press.

36. Virnig, B.A., et al., *Racial Variation In Quality Of Care Among Medicare+Choice Enrollees.* Health Affairs, 2002. **21**(6): p. 224-230.

37. Saha, S. and D.H. Hickam, *Explaining Low Ratings of Patient Satisfaction Among Asian-Americans.* American Journal of Medical Quality, 2003. **18**(6): p. 256-264.

38. Setodji, C., S. Reise, and L. Morales, Fongwa, M., Hays, R., *Differential Item Functioning by Survey Language Among Older Hispanics Enrolled in Medicare Managed Care: A New Method for Anchor Item Selection.* Medical care, 2011. **49**(5): p. 461-468 10.1097/MLR.0b013e318207edb5.

39. Rodriguez, H. and P. Crane, *Examining Multiple Sources of Differential Item Functioning on the Clinician & Group CAHPS® Survey.* Health Services Research, 2011. **46**(6pt1): p. 1778-1802.

40. Hays, R., et al., *Psychometric Properties of the CAHPS(TM) 1.0 Survey Measures.* Medical care, 1999. **37**(3): p. MS22-MS31.

41. Hargraves, L., R. Hays, and P. Cleary, *Psychometric Properties of the Consumer Assessment of Health Plans Study (CAHPS®) 2.0 Adult Core Survey.* Health Services Research, 2003. **38**(6p1): p. 1509-1528.

42. Solomon, L., R. Hays, and A. Zaslavsky, Ding, L., Cleary, P. , *Psychometric Properties of a Group-Level Consumer Assessment of Health Plans Study (CAHPS) Instrument.* Medical Care, 2005. **43**(1): p. 53-60.

43. Martino, S., et al., *Psychometric Properties of an Instrument to Assess Medicare Beneficiaries' Prescription Drug Plan Experiences.* Health Care Financing Review, 2009. **30**(3): p. 41.

44. Scholle, S., O. Vuong, and L. Ding, Gallagher, P., Brown, J. A., Hays, R. D., & Cleary, P.D. , *Development of and field-test results for the CAHPS PCMH survey.* Medical Care, Supplement, 2012.

45. Dyer, N., J. Sorra, and S. Smith, Cleary, P., & Hays, R. , *Psychometric properties of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) clinician and group adult visit survey.* Medical Care, Supplement, 2012.

46. Chong, K., P. Damiano, and R. & Hays, *Psychometric properties of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) 4.0 adult health plan survey.* Primary Health Care: Open Access 2012.

47. Fowler, F., P. Gallagher, and S. Nederend, *Comparing Telephone and Mail Responses to the CAHPS(TM) Survey Instrument.* Medical Care, 1999. **37**(3): p. MS41-MS49.

48. Hepner, K., J. Brown, and R. Hays, *Comparison of Mail and Telephone in Assessing Patient Experiences in Receiving Care from Medical Group Practices.* Evaluation & the Health Professions, 2005. **28**(4): p. 377-389.

49. de Vries, H., M. Elliott, and K. Hepner, Keller, S., Hays, R., *Equivalence of Mail and Telephone Responses to the CAHPS® Hospital Survey.* Health Services Research, 2005. **40**(6p2): p. 2120-2139.

50. Basu, A. and P.J. Rathouz, *Estimating marginal and incremental effects on health outcomes using flexible link and variance function models.* Biostatistics, 2005. **6**(1): p. 93-109.

51. Hays, R.D., et al., *Associations of cancer and other chronic medical conditions with SF-6D preference-based scores in Medicare beneficiaries.* Quality of Life Research, 2014. **23**(2): p. 385-391.

52. Elliott, M.N., et al., *Patterns of Unit and Item Nonresponse in the CAHPS® Hospital Survey.* Health Services Research, 2005. **40**(6p2): p. 2096-2119.

53. Elliott, M.N., et al., *Effects of Survey Mode, Patient Mix, and Nonresponse on CAHPS® Hospital Survey Scores.* Health Services Research, 2009. **44**(2p1): p. 501-518.

54. Klein, D.J., et al., *Understanding Nonresponse to the 2007 Medicare CAHPS Survey.* The Gerontologist, 2011. **51**(6): p. 843-855.

55. Damiano, P., et al., *Differential use of the CAHPS 0–10 global rating scale by medicaid and commercial populations.* Health Services and Outcomes Research Methodology, 2004. **5**(3): p. 193-205.

56. Goldstein, E., M. Elliott, and W. Lehrman, Hambarsoomian, K., Giordano, L., *Racial/Ethnic Differences in Patients' Perceptions of Inpatient Care Using the HCAHPS Survey.* Medical Care Research and Review, 2010. **67**(1): p. 74-92.

57. Bao, Y., S.A. Fox, and J.J. Escarce, *Socioeconomic and Racial/Ethnic Differences in the Discussion of Cancer Screening:"Between " versus "Within " Physician Differences.* Health services research, 2007. **42**(3p1): p. 950-970.

Chapter 3 Evaluation of a factor structure of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Clinician & Group Adult Visit 1.0 Survey

Introduction

The Consumer Assessment of Healthcare Providers and Systems (CAHPS®) project was initiated in 1995 by the Agency for Healthcare Research and Quality (AHRQ) to provide consumers with comparative information about health plan performance [1]. The CAHPS survey captures care quality domains that can be used by consumers in making decisions about ambulatory care (health plans, physician groups, individual physicians) and inpatient care (hospitals, nursing homes, surgical care). Health care providers can also use the survey findings to assess the patient-centeredness of the care provided and inform quality improvement initiatives [2]. CAHPS survey reporting composites consist of multiple items and are used to reduce reporting data points when presenting survey findings to consumers [3-7].

CG-CAHPS factor structure

Factor analyses have been reported for previous CAHPS surveys [8-12]. Some of this prior work suggests two overarching dimensions for CAHPS health plan report items: one factor relates to customer service and the other to the care delivered at providers' office including communication, office staff courtesy and respect, and access to care [8, 13].

Less information is available about the underlying structure of the CAHPS Clinician & Group (CG-CAHPS) surveys. Rodriguez and Crane (2011) found that the model with two factors (*physician communication* and *access to care*) had a better fit than a single factor model [14]. Dyer and colleagues performed confirmatory factor analysis of the CAHPS® Clinician & Group

Adult Visit 1.0 Survey at the individual patient level and at the practice level (a multi-level model) [15]. The individual and multi-level model analyses provided support for the three composites: *access to care*, *physician communication* and *office staff courtesy and helpfulness*.

This paper aims to study whether the assumed three factor structure (*access to care*, *physician communication* and *office staff courtesy and helpfulness* with associated items) of the CG-CAHPS survey can be confirmed in the CG-CAHPS Adult Visit 1.0 Survey comparative database for data collected in 2011 [16]. The CG-CAHPS Adult Visit 1.0 Survey is a version of the CG-CAHPS survey that focuses on experiences during a single visit, while other versions of the survey ask respondents to report about care experiences over the last 12 months.

Evaluation of the factor structure in this dataset is important for two reasons. First, while a number of studies have evaluated factor structure of CAHPS surveys, only Dyer et al. [15] have evaluated the CAHPS Clinician & Group Adult Visit 1.0 survey. They found support for the hypothesized three factor structure [15]. However, the authors used a confirmatory factor analytic (CFA) model that assumes continuous measures. The study also used only a subset of the full data for the analysis. Only respondents that answered all CG-CAHPS core questions were included in the analysis – 21 318 out of 103 442 respondents. Confirming a factor structure in a more recent dataset using a different analytical method (categorical CFA) will provide additional information about the reporting framework for CG-CAHPS Adult Visit 1.0 Survey. Second, the data analyzed in this paper were used in an earlier study that found significant differences in care experiences between Asian Americans and non-Hispanic whites [17]. The validity of this comparison hinges on a number of assumptions, an important one of which is that the instrument has the same underlying factor structure in these two racial-ethnic groups.

Confirming an underlying structure of the survey instrument is a prerequisite for comparing factor structure and evaluating measurement invariance across these two groups.

Instrument

We analyzed CG-CAHPS Adult Visit 1.0 Survey data collected in 2011 from the CAHPS Database. CG-CAHPS Adult Visit 1.0 Survey asks patients about their most recent visit to a doctor (a primary care doctor or a specialist).

CG-CAHPS Adult Visit 1.0 Survey used for this study has 41 questions. Thirteen questions are used to create three reporting composites – *access to care* (6 questions), *physician communication* (5 questions) and *office staff courtesy and helpfulness* (2 questions). There are two global rating questions and questions that ask for age, gender, education, race/ethnicity and overall health.

Substantive CAHPS survey questions are only asked of those for which they apply. Therefore, some of the missing responses are missing appropriately. Inappropriately missing values ranged from 1% to 6% and were of little concern for the purposes of this analysis (Appendix 1).

The *access to care* composite items ask about timeliness of care, timeliness of responses to medical questions over the phone and waiting times for appointments during the last 12 months using four-category response options (*Always, Usually, Sometimes, Never*). The *physician communication* composite questions refer to the most recent visit and ask whether the doctor listened carefully, gave clear instructions, showed respect, spent enough time with and seemed knowledgeable about respondent's medical history using a three-category response scale (*Yes, definitely; Yes, somewhat; No*). The *office staff courtesy and helpfulness* composite questions refer to the most recent visit and ask whether clerks and receptionists were helpful and treated

the respondent with courtesy and respect using the same three-category response scale used for physician communication.

The global rating items ask respondents *(1)* to rate their doctor on a 0-10 rating scale where 0 is the worst possible doctor and 10 is the best doctor and *(2)* whether they would recommend their doctor to family and friends using the three-category response scale (*Yes, definitely; Yes, somewhat; No*).

Table 2 shows the 15 questions used to elicit reports and ratings of care and their response options.

Methods

We conducted an individual level categorical confirmatory factor analysis using Mplus Version 6.12 [18]. CAHPS surveys have categorical response options. Therefore, we used the robust weighted least squares estimation procedure, available in Mplus as Mean- and Variance-adjusted Weighted Least Squares estimation method (WLSMV).

We followed earlier recommendations and evaluated the model fit using two types of goodness of fit tests: a parsimony correction index (RMSEA) and a comparative fit index (CFI, NNFI) [19, 20]. We did not rely on the chi-square test as with large samples such as the one used in this analysis, it becomes not informative, since even small differences between the covariance matrices become statistically significant so [21].

Parsimony correction indices reward parsimony and are insensitive to sample size. We used RMSEA to evaluate the model fit, where 0.05 or less would be considered a close fit. Comparative fit indices compare a suggested model to a more restricted model. We used CFI and

NNFI to test the model fit, where 0.95 or greater would be considered an acceptable fit. We acknowledged a good model fit if at least RMSEA and CFI met the fit criteria [22-24].

In our confirmatory factor model, there were three factors - *(1) access to care, (2) physician communication and global ratings* and *(3) office staff helpfulness and courtesy (*Figure 1*)*. The hypothesized factor structure was based on preliminary findings from a principal components analysis we ran prior to this analysis (findings not presented). The items used to create factors/composites were loaded on respective latent variables (Table 2). Each item loaded only on one factor. Latent variables were allowed to correlate. We fixed the variance of each factor at 1.0. A hypothesized factor structure is presented in Table 2. We used a factor loadings of 0.40 or greater as a cut off for the appropriateness of the item to the proposed factor [25]. We examined for the estimated correlations among the factors. Although correlations between different factors/composites are expected, very high correlations (0.80 or above) imply that the factors may not be unique.

We estimated Cronbach's alpha coefficient of internal consistency and reliability for each factor/composite. We considered a reliability of 0.70 or above to be acceptable reliability for group comparisons [26, 27]. In order to evaluate whether the data have a truly nested structure, we estimated intraclass correlation coefficients (ICC) for each factor. The ICC values requiring the use of multilevel models vary from 0.1 to 0.05 [28-30]. We used 0.1 as the guiding criterion.

Results

Key sample characteristics and an overview of missingness in the data are provided in Table 1 and Appendix 1. Respondents were predominantly white and educated. Means and standard deviations for the CAHPS items are provided in Table 3. Reports about care and global rating items are negatively skewed, with skewness ranging from -0.93 to -4.52.

All factor loadings were statistically significant (P>0.01) and above the 0.40 cut point (Table 4). The smallest loading was found for the item asking about how often a respondent was seen within 15 minutes of an appointment (0.57).

All model fit indices supported model fit. The CFI and NNFI were 0.99 – both above the 0.95 criterion. The RMSEA was 0.03, which is below the 0.05 cut point. Table 5 provides more details on these indices.

Cronbach's alpha coefficient for all factors were all acceptable. The coefficient for the *physician communication and global ratings* factor was 0.91. The *access* and *office staff* factors had coefficients of 0.81 and 0.83, respectively. Estimated correlations among factors are presented in Table 6. None of the correlations were above the 0.80 threshold, and ranged from 0.46 to 0.60. The ICCs for the three factors ranged from 0.03 to 0.1 (*access to care* – 0.1, *physician communication and global ratings – 0.04, office staff helpfulness and courtesy – 0.03*) supporting the use of an individual level analysis.

Discussion

The CAHPS composites summarize information from the CAHPS surveys to consumers and providers in an easy to understand format that can be used to compare care provided to different groups. However, the validity of reporting composites and group comparisons rests on whether an assumed factor structure holds in a specific dataset.

We conducted a confirmatory factor analysis to assess the assumed three factor structure in our CG-CAHPS Adult Visit 1.0 Survey dataset. The confirmatory factor analysis showed that the hypothesized three factor structure fits well in our dataset. All the items loaded significantly on to the respective hypothesized factors. The lowest loadings were observed for the item on wait

times at the doctor's office. This item was shown to load weakly on access to care related factors in other CAHPS surveys as well, which in part can be explained with findings from other studies that suggest overlapping association for this item with physician communication and office staff courtesy composites [8, 31, 32]. All three factors showed an acceptable internal consistency reliability.

There are few studies that examined the underlying factor structure of CG-CAHPS Adult Visit 1.0 Survey data. Dyer and colleagues assessed the three factor structure model fit using an earlier version of the CG-CAHPS Adult Visit 1.0 Survey data. The study findings supported the underlying three factor structure [15]. However, Dyer and colleagues used a factor structure suggested by CAHPS developers in their analysis, which was reflective of the three survey reporting composites – *access to care, physician communication* and *office staff helpfulness and courtesy*. In our study, we conducted preliminary principal component analysis to explore for a factor structure, which supported combining *physician communication* composite and *global rating* items together as one factor. The two other factors (*access to care* and *office staff courtesy and helpfulness*) in our factor structure were the same as the reporting composites suggested by the survey developers.

The grouping of *physician communication* composite items and *global rating* items together can be justified on theoretical grounds. Physician communication plays an important role in patient perceptions of care [33]. Communication between a physician and a patient was shown to be significantly associated with how a patient rated a doctor and his/her willingness to recommend a doctor to other people [8, 34].

This study makes an important contribution to the literature on the underlying factor structures of CG-CAHPS surveys. The study findings suggest CG-CAHPS Adult Visit 1.0 Survey data have an underlying three factor structure. The composition of three underlying factors is, however, different from the composition recommended by the survey developers as it groups together *physician communication* and *global rating* items. This may have important implications on how the CG-CAHPS Adult Visit 1.0 Survey findings are reported to consumers.

Although the three factor structure tested in our study showed higher factor loadings and better model fit compared to the structure in the study by Dyer and colleagues, future studies may be needed that evaluate the factor structures in other CG-CAHPS Adult Visit 1.0 Survey datasets to confirm which factor structure represent the true underlying factor structure.

Our study also had a number of limitation. Although the sample analyzed represented respondents with various racial ethnic and demographic backgrounds, the respondents were predominantly white and educated. Therefore, the generalizability of the three factor structure supported in this analysis to non-white racial and specific socio-economic groups may be limited and requires further research. The CAHPS® Clinician & Group Adult Visit 1.0 Survey explores patient experiences over a single visit, which may contain more random error. The same factor structure should also be tested in the CG-CAHPS Adult Survey datasets that report on patient experiences over a period of time such as the CG-CAHPS Adult 12-Month Surveys. Our study tested the factor structure at an individual level guided by the ICCs values. Whether this factor structure holds in a multi-level model that includes practice as a level also may have to be explored in future research.

Tables and figures

Figure 3.1 Three-factor model

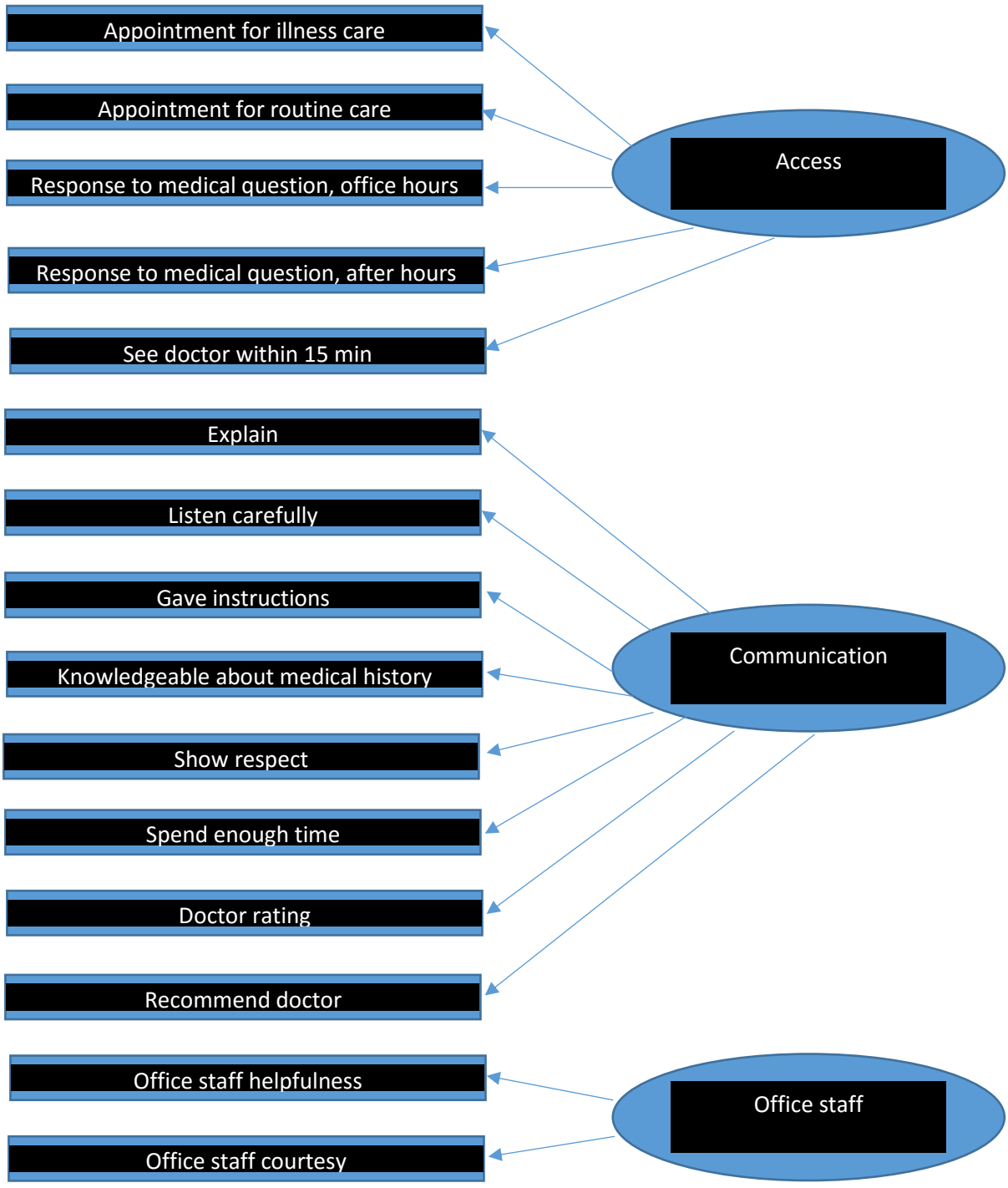

Table 3.1 Key sample demographic information, CAHPS-CG Adult Visit Data from 2011

Gender		
	Percent	**Frequency**
Male	36.15	94,068
Female	63.85	160,236
Total	**100.00**	**260,236**
Missing values		6,091
Age		
18-24	2.83	7,364
25-34	7.96	20,733
35-44	9.29	24,207
45-54	16.93	44,118
55-64	24.51	63,872
65-74	21.44	55,872
75+	17.06	44,455
Total	**100.00**	**260,621**
Missing values		**5,706**
Education		
8th grade or less	2.47	6,329
Less than HS grad	4.62	11,845
HS grad (includes GED)	25.15	64,475
Some college	30.23	77,511
4 year grad	17.43	44,681
More than 4 years	20.10	51,542
Total	**100.00**	**256,383**
Missing values		**9,944**
Race/Ethnicity		
White	89.49	229,899
African-American	4.30	11,057
Asian	2.16	5,545
Native Hawaiian/Pacific Islander	0.13	338
American Indian/Native Alaskan	0.32	839
Other	2.17	5,575
Multi-racial	1.42	3,652
Total	**100.00**	**256,885**
Missing values		**9,442**

Note: Percentages do not always add up to 100.00 because of rounding

Table 3.1 Key sample demographic information, CAHPS-CG Adult Visit Data from 2011 (continued)

Hispanic/Latino origin or descent		
	Percent	**Frequency**
Yes	4.95	12,475
No	95.05	239,722
Total	**100.00**	**252,197**
Missing values		**14,130**
Health status		
Excellent	12.60	32,747
Very Good	34.41	89,428
Good	33.99	88,353
Fair	15.25	39,625
Poor	3.75	9,752
Total	**100.00**	**259,905**
Missing values		**6,422**
Practice regions		
Midwest	52.48	139,775
Northeast	22.93	61,056
South	10.08	26,843
West	14.51	38,653
Total	**100.00**	**266,327**
Missing values		**6,422**

Note: Percentages do not always add up to 100.00 because of rounding

Table 3.2 Reports and ratings of care items

Physician communication and global ratings 7. During your most recent visit, did this doctor explain things in a way that was easy to understand? 8. During your most recent visit, did this doctor listen carefully to you? 9. During your most recent visit, did this doctor give you easy to understand instructions about taking care of these health problems or concerns? 10. During your most recent visit, did this doctor seem to know the important information about your medical history? 11. During your most recent visit, did this doctor show respect for what you had to say? 12. During your most recent visit, did this doctor spend enough time with you? 13. Using any number from 0 to 10, where 0 is the worst doctor possible and 10 is the best doctor possible, what number would you use to rate this doctor? 14. Would you recommend this doctor's office to your family and friends?
Access to care 6. In the last 12 months, when you phoned this doctor's office to get an appointment for care you needed right away, how often did you get an appointment as soon as you thought you needed? 7. In the last 12 months, when you made an appointment for a check-up or routine care with this doctor, how often did you get an appointment as soon as you thought you needed? 8. In the last 12 months, when you phoned this doctor's office during regular office hours, how often did you get an answer to your medical question that same day? 9. In the last 12 months, when you phoned this doctor's office after regular office hours, how often did you get an answer to your medical question as soon as you needed? 10. Wait time includes time spent in the waiting room and exam room. In the last 12 months, how often did you see this doctor within 15 minutes of your appointment time?
Office staff courtesy and helpfulness 3. During your most recent visit, were clerks and receptionists at this doctor's office as helpful as you thought they should be? 4. During your most recent visit, did clerks and receptionists at this doctor's office treat you with courtesy and respect?

Note: First six items in *physician communication and global ratings*, both items in *office staff courtesy and helpfulness* and the *recommending doctor to friends/family* question use a three response option scale (*Yes, definitely; Yes, somewhat; No*). *Access to care* composite items use a four response option scale (*Never, Sometimes, Usually, Always*). The *doctor rating* question uses a 0-10 response scale.

Table 3.3 Means and standard deviations for the factor/composite items

Factor/Composite items	Factor loadings	Std. Dev.
Physician communication and global ratings		
1 Explained things in a way that was easy to understand	2.90	0.33
2 Listened carefully	2.91	0.32
3 Gave easy to understand instructions	2.88	0.38
4 Was knowledgeable about medical history	2.81	0.46
5 Showed respect	2.93	0.30
6 Spent enough time	2.89	0.37
7 Rating a doctor	3.25	0.99
8 Recommending a doctor to friends/family	2.86	0.42
Access		
1 Timely appointment for urgent care	3.52	0.76
2 Timely appointment for check-up or routine care	3.62	0.65
3 Received answer to medical questions the same day, regular hours	3.44	0.81
4 Received answer to medical questions timely, after regular hours	3.39	0.92
5 Was seen within 15 minutes of appointment time	3.15	0.96
Office staff		
1 Was helpful	2.87	0.38
2 Treated with courtesy and respect	2.92	0.29

Note: First six items in *physician communication and global ratings*, both items in *office staff courtesy and helpfulness* and the *recommending doctor to friends/family* question use a three response option scale (*Yes, definitely; Yes, somewhat; No*). All these items were reverse coded to (*1- No; 2 - Yes, somewhat; 3- Yes, definitely*). *Access to care* composite items use a four response option scale (*Never, Sometimes, Usually, Always*). The *doctor rating* question uses a 0-10 response scale and was recoded (*0-7=1; 8-2; 9-3; 10-4*).

Table 3.4 Factor loadings and standard errors for the factor/composite items

Factor/Composite items	Factor loadings	Std. Err.
Physician communication and global ratings		
1 Explained things in a way that was easy to understand	0.90	0.001
2 Listened carefully	0.96	0.001
3 Gave easy to understand instructions	0.89	0.001
4 Was knowledgeable about medical history	0.78	0.002
5 Showed respect	0.96	0.001
6 Spent enough time	0.88	0.001
7 Rating a doctor	0.84	0.001
8 Recommending a doctor to friends/family	0.92	0.001
Access		
1 Timely appointment for urgent care	0.88	0.002
2 Timely appointment for check-up or routine care	0.84	0.002
3 Received answer to medical questions the same day, regular hours	0.78	0.003
4 Received answer to medical questions timely, after regular hours	0.83	0.006
5 Was seen within 15 minutes of appointment time	0.57	0.002
Office staff		
1 Was helpful	0.98	0.003
2 Treated with courtesy and respect	0.94	0.003

Table 3.5 Model fit indices

Fit statistics	CFA
Chi square	23123.5
df	87
CFI	0.99
TFI	0.99
RMSEA	0.03

Table 3.6 Estimated factor correlations

Factors	Physician communication and global ratings	Access	Office staff
Physician communication and global ratings	1		
Access	0.60	1	
Office staff	0.46	0.51	1

Appendix

Appendix 3.1 Reports about care and global rating items: total number of responses and missing values

	% Not Answered	% Appropriately Skipped	% Total Missing	N
Access to care				
Getting care quickly when urgent	2.23	55.93	58.16	260,522
Getting care quickly when not urgent	3.22	26.75	29.97	258,018
Getting an answer to a medical question the same day when calling during regular hours	3.75	59.10	62.84	256,710
Getting an answer to a medical question quickly when calling after regular hours	5.77	92.63	98.40	251,794
Seeing a doctor within 15 minutes of appointment time	1.91	N/A	1.91	261,329
Physician communication and global ratings				
Doctor explained things clearly	5.27	N/A	5.27	252,986
Doctor listened carefully	2.84	N/A	2.84	258,962
Doctor gave clear instructions	3.94	14.65	18.59	261,790
Doctor knew important information about medical history	1.20	N/A	1.20	263,164
Doctor showed respect	2.73	N/A	2.73	259,257
Doctor spent enough time	2.69	N/A	2.69	259,357
Rating a doctor	1.18	N/A	1.18	263,216
Recommending a doctor to friends/family	3.30	N/A	3.30	257,813
Office staff helpfulness				
Doctor office staff helpful	3.17	N/A	3.17	258,147
Courtesy and respect from doctor office staff	3.38	N/A	3.38	257,609

References

1. AHRQ. *Agency for Healthcare Research and Qualty: About CAHPS*. 2012 [cited 2012 August 13]; Available from: http://www.cahps.ahrq.gov/about.htm.

2. Price, R.A., et al., *Examining the role of patient experience surveys in measuring health care quality*. Medical Care Research and Review, 2014. **71**(5): p. 522-554.

3. Damiano, P.C., et al., *CAHPS® in practice: The Iowa demonstration.* The Journal of ambulatory care management, 2002. **25**(2): p. 32-42.

4. Farley, D., et al., *Effect of CAHPS Performance Information on Health Plan Choices by Iowa Medicaid Beneficiaries.* Medical Care Research and Review, 2002. **59**(3): p. 319-336.

5. Farley, D.O., et al., *Effects of CAHPS health plan performance information on plan choices by New Jersey Medicaid beneficiaries.* Health Services Research, 2002. **37**(4): p. 985-1007.

6. Derose, K.P., et al., *Developing a Spanish-language Consumer Report for CAHPS® Health Plan Surveys.* The Joint Commission Journal on Quality and Patient Safety, 2007. **33**(11): p. 681-688.

7. Quigley, D.D., et al., *Use of CAHPS® patient experience survey data as part of a patient-centered medical home quality improvement initiative.* Journal of Healthcare Leadership, 2015. **7**: p. 41-54.

8. Hargraves, L., R. Hays, and P. Cleary, *Psychometric Properties of the Consumer Assessment of Health Plans Study (CAHPS®) 2.0 Adult Core Survey.* Health Services Research, 2003. **38**(6p1): p. 1509-1528.

9. Marshall, G., L. Morales, and M. Elliott, Spritzer, K., Hays, R., *Confirmatory factor analysis of the Consumer Assessment of Health Plans Study (CAHPS) 1.0 core survey.* Psychological Assessment, 2001. **13**(2): p. 216-229.

10. O'Malley, A.J., et al., *Exploratory Factor Analyses of the CAHPS® Hospital Pilot Survey Responses across and within Medical, Surgical, and Obstetric Services.* Health Services Research, 2005. **40**(6p2): p. 2078-2095.

11. Reise, S., J. Morizot, and R. Hays, *The role of the bifactor model in resolving dimensionality issues in health outcomes measures.* Quality of Life Research, 2007. **16**(0): p. 19-31.

12. Bender, R. and S. Garfinkel, *Differences in the Structure of CAHPS Measures among the Medicare Fee-for-Service, Medicare Managed Care, and Privately Insured Populations.* Health Services Research, 2001. **46**(3): p. 489–508.

13. Setodji, C., S. Reise, and L. Morales, Fongwa, M., Hays, R., *Differential Item Functioning by Survey Language Among Older Hispanics Enrolled in Medicare*

Managed Care: A New Method for Anchor Item Selection. Medical care, 2011. **49**(5): p. 461-468 10.1097/MLR.0b013e318207edb5.

14. Rodriguez, H. and P. Crane, *Examining Multiple Sources of Differential Item Functioning on the Clinician & Group CAHPS® Survey.* Health Services Research, 2011. **46**(6pt1): p. 1778-1802.

15. Dyer, N., J. Sorra, and S. Smith, Cleary, P., & Hays, R. , *Psychometric properties of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) clinician and group adult visit survey*. Medical Care, Supplement, 2012.

16. *Agency for Healthcare Research and Quality: CAHPS Database* [cited 2016; Available from: http://www.ahrq.gov/cahps/cahps-database/index.html.

17. Ahmedov, M., *Chapter 2: Differences in perceptions of health care between Asian Americans and non-Hispanic whites: an analysis of the CAHPS Clinician and Group Adult Visit Survey 1.0.*, in *Health policy and management*. 2016, University of California, Los Angeles.

18. Muthén, L. and B. Muthén, *Mplus.* Statistical analysis with latent variables. Version, 2007. **3**.

19. Brown, T.A., *Confirmatory factor analysis for applied research*. 2015: Guilford Publications.

20. Kline, R.B., *Principles and practice of structural equation modeling*. 2015: Guilford publications.

21. Chan, Y., et al., *Confirmatory factor analysis of the Child Abuse Potential Inventory: Results based on a sample of Chinese mothers in Hong Kong.* Child abuse & neglect, 2006. **30**(9): p. 1005-1016.

22. Browne, M.W., et al., *Alternative ways of assessing model fit.* Sage focus editions, 1993. **154**: p. 136-136.

23. Hu, L.t. and P.M. Bentler, *Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives.* Structural equation modeling: a multidisciplinary journal, 1999. **6**(1): p. 1-55.

24. Reeve, B.B., et al., *Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS).* Medical care, 2007. **45**(5): p. S22-S31.

25. Peterson, R.A., *A meta-analysis of variance accounted for and factor loadings in exploratory factor analysis.* Marketing Letters, 2000. **11**(3): p. 261-275.

26. Cronbach, L., *Coefficient alpha and the internal structure of tests.* Psychometrika, 1951. **16**(3): p. 297-334.

27. Nunnally, J., *C.(1978). Psychometric theory*. 1978, New York: McGraw-Hill.

28. Hox, J.J., M. Moerbeek, and R. van de Schoot, *Multilevel analysis: Techniques and applications*. 2010: Routledge.

29. Muthen, B., *Latent variable modeling of longitudinal and multilevel data.* Sociological methodology, 1997. **27**(1): p. 453-480.

30. Julian, M.W., *The Consequences of Ignoring Multilevel Data Structures in Nonhierarchical Covariance Modeling.* Structural Equation Modeling: A Multidisciplinary Journal, 2001. **8**(3): p. 325-352.

31. Anderson, R.T., F.T. Camacho, and R. Balkrishnan, *Willing to wait?: the influence of patient wait time on satisfaction with primary care.* BMC health services research, 2007. **7**(1): p. 1.

32. Hays, R.D., et al., *Classical test theory and item response theory analyses of multi-item scales assessing parents' perceptions of their children's dental care.* Medical care, 2006: p. S60-S68.

33. Ong, L.M.L., et al., *Doctor-patient communication: A review of the literature.* Social Science & Medicine, 1995. **40**(7): p. 903-918.

34. Solomon, L., R. Hays, and A. Zaslavsky, Ding, L., Cleary, P. , *Psychometric Properties of a Group-Level Consumer Assessment of Health Plans Study (CAHPS) Instrument.* Medical Care, 2005. **43**(1): p. 53-60.

Chapter 4 Measurement invariance between non-Hispanic whites and Asian Americans on the Consumer Assessment of Healthcare Providers and Systems Clinician and Group Adult Visit Survey

Introduction

The Consumer Assessment of Healthcare Providers and Systems (CAHPS®) surveys are the most widely used measures of patient experiences with health care providers and plans in the US. [1]. The CAHPS surveys are increasingly used as a part of mandated quality reports and a basis for performance based payment schemes [2-4]. Differences in patient experiences between various racial/ethnic groups, including non-Hispanic Whites and Asian Americans (hereafter Whites and Asians) in the CAHPS surveys are widely reported. In CAHPS surveys, Asians consistently tend to report worse patient experiences than Whites, with the effect size of differences reaching up to 2 (CAHPS Medicare Plan D surveys) [5-10]. While the proportion of Asians in the US population is rapidly increasing, underlying reasons for the reported differences is little studied [9]. The reported differences in patient experiences between Whites and Asians could be attributed to actual differences in care experiences or to measurement issues related to survey instruments (e.g. group differences in survey response tendencies).

Measurement issues related to the use of self-report instruments are highlighted as an important area of concern in studying and addressing health disparities [11]. Evaluation of measurement invariance is frequently used to explore for survey measurement issues. The measurement invariance can be defined as the same probability of obtaining a score by different group members if they have the same levels of the latent variable being measured [12-14]. A limited number of studies explored measurement invariance across different racial/ethnic groups

using CAHPS surveys. Studies evaluating measurement invariance across Asians and other non-white populations show contrasting findings – while some studies present evidence of differences in response styles, the others support differences in care experiences.

Zhu and colleagues examined measurement invariance across seven racial/ethnic groups on the Hospital Consumer Assessment of Healthcare Providers and Systems (H-CAHPS) data [15]. They found lack of measurement (scalar) invariance for *physician communication* composite measures between Whites and Asians and Whites and other racial/ethnic groups (e.g. non-Hispanic African Americans and Native Hawaiians). Lack of measurement invariance can arise from factors such as differences in response styles and differences in interpretation of survey items and imply that any differences in reports are due to survey measurement issues and not differences in care experiences [16]. Scalar invariance is a form of measurement invariance where factor loadings and indicator thresholds are equal between the groups of interest.

Carle and colleagues looked at measurement invariance in a sample of Medicaid managed care enrollees on the CAHPS Cultural Competence Survey [14]. They investigated invariance by English versus Spanish language and across three racial/ethnic groups (White, Black and Hispanic) [17, 18]. Their findings supported the scalar invariance, thus suggesting comparability of the CAHPS Cultural Competence Survey scores across English and Spanish versions of the survey and three racial/ethnic groups. Hays and colleagues assessed measurement invariance between Asians and Whites on the CAHPS Medicare Survey data and also found scalar invariance and thus support for the comparability of reported mean scores between these two racial/ethnic groups [19].

A very limited number of studies explored for differences in care experiences between Whites and Asians using the Consumer Assessment of Healthcare Providers and Systems Clinician and Group (CG-CAHPS) Adult Visit Survey data [10]. The unpublished analysis showed that Asians reported worse care experiences compared to Whites in this dataset [10]. These findings, however, are not conclusive without evidence of measurement invariance between these two racial/ethnic groups. We found no published study that have explored for measurement invariance between Asians and non-Hispanic whites using national CG-CAHPS Adult Visit Survey data. A follow up analysis evaluating measurement invariance between these two groups in the national CG-CAHPS Adult Visit Survey data, therefore, is well-warranted. This study builds on earlier analyses and examines measurement equivalence between Asians and Whites on the CG-CAHPS Adult Visit Survey.

Methods

Data

The 2011 CAHPS Clinician and Group Adult Visit Survey 1.0 data in the CAHPS Database is used in the analyses [20]. The CAHPS Database is a warehouse of data collected from various organization in the US through voluntary submission. The database serves to facilitate comparison of survey findings among participating organizations.

The CAHPS Clinician and Group Adult Visit Survey 1.0 focuses on the respondents' most recent visit to a primary care physician (i.e. an internist, family practitioner, obstetrician/gynecologist) or a specialist (i.e. a surgeon) throughout 2011. The survey consists of 41 questions. Thirteen items ask for reports about care: 6 for *physician communication*, 2 for *office staff courtesy and helpfulness*, and 5 for *access to care*. The questions about physician communication ask if the physician listened carefully, spent enough time, and gave instructions

that were easily understandable. The questions about the office staff helpfulness and courtesy ask whether office staff were helpful, courteous and showed respect. Both these question sets use a three-response option scale (*Yes, definitely; Yes, somewhat; No*) and refer to the most recent visit. Questions about access to care ask about timeliness of care, timeliness of answers to phone questions, and appointment wait times have four-response options (*Always, Usually, Sometimes, Never*) and refer to a 12-month recall period.

The survey also includes two global rating items. One item assesses global perceptions of the doctor using a 0-10 response scale (0 is the worst possible doctor and 10 is the best possible doctor). The second item asks respondents if they would recommend their doctor to family and friends on a three-response option scale (*Yes, definitely; Yes, somewhat; No*). Age, education, gender, race/ethnicity, and overall health are also assessed in the survey. All 15 survey items used to elicit reports and ratings of care are used in the analysis (Table 1).

The dataset used in this analysis includes 769 providers and 266,327 respondents. The vast majority of the surveys (98%) were administered by mail and the rest (2%) by phone. Only 0.5% of the surveys were administered in Spanish. Previous research indicates that the telephone and mail responses to different versions of CAHPS® survey yield similar results [21-23]. Table 2 provides descriptive demographic information about the sample of Asian Americans and non-Hispanic whites used in this study. Missing values for the composite and global rating items ranged from 1% to 6%.

Two questions in the CAHPS surveys ask about race/ethnicity. The first question asks if the respondent is of Hispanic or Latino origin. The next question asks whether the respondent is: 1) White, 2) Black or African American, 3) Asian, 4) Native Hawaiian or Pacific Islander, 5)

American Indian or Alaska Native, or 6) other. Similar to the US Decennial Census, those respondents confirming Hispanic or Latino origin are coded as Hispanic. Only those who reported no Hispanic origin or had a missing value were coded in accordance to the ethnicity they identified themselves with such as *white, African American, Asian American [1]*. Native Hawaiians, Pacific Islanders, American Indians, Alaska Native and other races were coded into the *other* racial-ethnic group.

Analysis Plan

In another study, we verified the hypothesized three-factor structure ((*(1) access to care, (2) physician communication and global ratings* and *(3) office staff courtesy and helpfulness*) of the CG-CAHPS Adult Visit Survey items using exploratory factor analysis (Table 3) [24]. In this study, we used a multiple group confirmatory factor analysis (MG-CFA) to assess measurement invariance between Asians and Whites. Since response variables in our dataset were categorical, we used Mean- and Variance-adjusted Weighted Least Squares estimation method (WLSMV) with the polychoric correlation matrix. WLSMV uses variances-covariances and kurtosis as weights to adjust for multivariate normality violations. When used with categorical data, WLSMV provides robust standard errors (SE) [25]. All analyses were conducted with Mplus 7 [26]. We used Mplus theta parameterization and full information maximum likelihood estimation. Theta parameterization is a model specification method that takes into account variance differences across groups and thus improves the model analysis [27].

We tested for configural, metric and scalar invariance, since CAHPS reports rely on group mean scores for comparisons across racial/ethnic groups. Configural invariance confirms that the number of factors and the pattern of indicator-factor loadings are identical across the groups. Metric and scalar invariance confirm equality of factor loadings, and factors loadings and

intercepts respectively [28]. MG CFA analysis requires three major steps [28]. In the first step, we performed a single group confirmatory factor analysis (CFA) separately in the White and Asian subgroups to confirm that the measurement model has an acceptable fit in the two groups. Model fit indices and parameter estimates for the two groups were evaluated.

In the second step, we conducted a test of configural invariance to explore whether the pattern of free factor loadings and thresholds were similar across groups. Lack of configural invariance implies that different latent variables in the two groups may have been measured.

In the third step, we first constrained factor loadings to equality in the two groups to conduct a metric invariance test. Evidence of metric invariance implies that latent variables are related to survey items the same way across groups. That is, one-unit change in item score translates to the same unit change in estimated latent variable score in the groups. Once the metric invariance test results show an acceptable model fit, a new constraint – equality of thresholds across groups – was added to test for scalar invariance. Evidence of scalar invariance implies that individuals with the same score on a latent variable are likely to have the same score on survey items corresponding to the latent variable. Thus, score means across groups can be compared validly.

We performed two versions of a metric invariance test. In the first version, we followed an approach suggested by Brown, where we fixed factor means to 0 in both groups and factor variances to 1 in one group and placed no constraints in the other [28]. This model produced a difference of 12 constraints (difference in degrees of freedom) between the configural and metric models, while we placed a constraint on 15 parameters (15 factor loadings). In order to address this issue, we fixed factor means to 0 and factor variance to 1 in both groups and obtained a difference in degrees of freedom corresponding to the number of constrained parameters.

We also performed two versions of scalar invariance testing. In the first approach (proposed by Brown), we fixed factor means to 0 and factor variances to 1 in one group and set both parameters free in the second group. The test results showed a difference in degrees of freedom between the two models to be 33, while we placed constraints on 36 parameters. In the second approach, we fixed factor means to 0 and factor variances to 1 in both groups, which yielded a difference in degrees of freedom of 36.

In general, in nested models, chi square values should be higher for the more constrained models compared to the less constrained models. This was not observed in model outputs when we used the WLSMV estimation in our analysis. In our analysis, chi square values in the metric models were consistently lower than those in the configural models (Table 4). Chi square values from nested models are considered not to be directly comparable when the WLSMV estimation is used. This places limitation on evaluation of a model fit, as many practical fit indices are derived from chi square values. To confirm the model findings, we also estimated models with an MLR estimation. The MLR based models also produced lower chi square values in the metric models compared to those in the configural models. Therefore, in the final step, we run the models with the ML estimation. The chi square values in more constrained models were higher than those in less constrained models when using the ML estimation, thus permitting comparison of model fit indices across models.

We evaluated the model fit using a combination of three approximate model fit indices. The single group CFA and configural invariance models were evaluated by the root mean square error of approximation (RMSEA), the comparative fit index (CFI), and the Tucker-Lewis index (TLI). For the nested models, a CFI difference test is recommended [29]. However, CFI difference tests are recommended based on findings from ML models on continuous data,

therefore applicability of the recommendations to other estimation methods and particularly categorical ordered data need to be interpreted cautiously [30]. Furthermore, CFI values in nested models using the WLSMV and MLR estimations may not be directly comparable [30].

The following criteria were used to inform our decisions about the model fit – RMSEA < 0.08, CFI > 0.90, TLI > 0.90 [29, 31-34]. Many researchers recommend the model fit evaluation to be informed by combination of a set of fit indices and not any single index [28, 29, 35]. In our study, the model fit was found to be acceptable when two out of three fit indices had met the above outlined criteria.

Results

Single group CFA model outputs supported an acceptable model fit for both Asians and Whites. The model for Asians produced RMSEA (the upper and lower limits of 90% confidence interval) = 0.037 (0.040; 0.034), CFI = 0.991 and TLI = 0.989 and the model for Whites – RMSEA (the upper and lower limits of 90% confidence interval) = 0.031 (0.031; 0.030), CFI = 0.994 and TFI = 0.992 (Table 4). Standardized factor loadings for the two single group models are presented in Table 5. All loadings are statistically significant at $p<0.001$ and greater than 0.50. While the loadings for Asians are uniformly higher than for Whites for *access to care* and *office staff courtesy and helpfulness* composites, the loadings for *physician communication and global ratings* are not uniformly higher for either one of the groups. The magnitude of difference in the standardized loadings between the two groups is the greatest for the item on timely response to a medical question after office hours (0.075; $p<0.01$).

In an MG CFA analysis using WLSMV estimation, the fit indices for the configural invariance model are in the acceptable range – RMSEA (the upper and lower limit of 90%

confidence interval) = 0.031 (0.031; 0.030), CFI = 0.993 and TLI = 0.992. No modification indices beyond the specified value of 4.0 are observed, suggesting no model misspecification.

We conducted two versions of metric and scalar invariance testing using WLSMV estimation. Both versions of the tests showed an acceptable model fit (Table 4). The metric invariance model with factor means fixed to 0 and factor variances fixed to 1 in both groups produced a slightly better fit compared with the model with factor means fixed to 0 in both groups and factor variances fixed to 1 in one group and no constraints placed in the other group (RMSEA = 0.024 vs. 0.027, CFI = 0.996 vs. 0.994, TLI = 0.995 vs. 0.994). Neither of the scalar invariance models produced consistently better fit indices (RMSEA = 0.024 vs. 0.025, CFI = 0.995 vs. 0.994, TLI = 0.995 vs. 0.995).

Table 4 presents model parameters for the metric and scalar invariance testing using MLR and ML estimation methods. Both models supported metric invariance when using approximate fit indices. The standardized parameter estimates for the scalar invariance testing model using the WLSMV estimation are presented in Table 6.

Discussion

In this study, we investigated whether worse care experiences reported by Asians compared with Whites in CAHPS surveys could be explained by measurement variance. We conducted MG CFA to test for measurement invariance in a CG-CAHPS Adult Visit Survey 1.0 dataset between the two racial/ethnic groups. Our study provides support for measurement invariance between Asians and Whites in the CG-CAHPS Adult Visit Survey measures of *access to care, physician communication and global ratings*, and *office staff courtesy and helpfulness*. In our analysis, the criteria for both metric and scalar invariance were met, suggesting that the mean

differences reported in the CG-CAHPS Adult Visit Survey data between these two groups were likely to be due to differences in care experiences.

There are a limited number of studies that have explored measurement invariance among different racial/ethnic groups using CAHPS surveys [15, 17, 18, 36]. Very few have specifically compared Asians and Whites [15, 19]. To our knowledge, our study is the first that tested for measurement invariance of latent constructs between Asians and Whites in a large CG-CAHPS Adult Visit Survey dataset. Furthermore, we conducted MG CFA using three estimation approaches: theta parametrization and WLSMV estimation (used for categorical variables), MLR and ML estimation. Our analysis also found that chi square values may not be directly comparable across models in MG CFA analyses when MLR estimation is used with categorical data and therefore, may limit the use of fit indices across models. Further research is needed to explore whether this is supported in other categorical datasets.

Our study had a number of limitations. The response rate for the survey dataset was not available. CAHPS adult surveys in samples other than Medicare tend to yield response rates below 40%. If non-responders differ in how they interpret and respond to survey questions, the study results could be biased and generalizability of the findings may be limited too. Furthermore, Asians are a diverse racial/ethnic group comprising of various subgroups (e.g., Chinese, Japanese, Vietnamese). The subgroups may have different patient care experiences and may differ in how they interpret and respond to CAHPS survey items. In this study, we do not have information about subgroups of Asians and grouped all Asians together. Asians, who have limited English proficiency, may have worse care experiences due to difficulties in communication. They are also less likely to respond to survey questions, thus possibly biasing the study findings. Asians who marked Hispanic origin (n=135) are not included in our sample

of Asians. Future studies should explore for reported differences and measurement invariance among various Asian subgroups, including Hispanic Asians.

In our analysis, the two groups had different sample sizes, which might lead to unequal contribution to model parameters that were sensitive to sample sizes. On the other hand, using equal sample sizes would have reduced the power to detect differences in groups.

Various model fit indices accounting for sample size and model complexity are used in measurement invariance testing [13, 29, 35]. Given the large sample size and various estimation methods, we determined measurement invariance applying primarily comparative fit indices. Interpretation of more conventional fit indices used for nested models such as the chi square difference test and change in CFI values are limited due to lack of information on use with categorical data.

For instance, the test of chi square difference is frequently used to evaluate for measurement invariance. A non-significant chi square difference test confirms that differences in model parameters across compared groups are not significant, hence measurement invariance exists. However, the chi square based tests are sensitive to sample sizes, so that even very small differences between the groups result in significant test results in large samples [13, 29, 34]. Chi square values in nested models did not conform to expected pattern when using WLSMV and MLR estimation methods, thus making limiting the use of nested model fit indices.

In conclusion, the validity of racial/ethnic comparisons of reports of patient experiences is critical to informing quality improvement initiatives and policy decisions. Without the evidence of measurement invariance, the validity of comparisons should not be assumed. The findings of this study show that our demonstration of measurement invariance supports the use of the CG-

CAPHS Adult Visit Survey measures in comparisons of patient perceptions of care across Asians and Whites in the US. The differences in reports and ratings of care reported between Asians and Whites in the CG-CAPHS Adult Visit Survey data are likely due to differences in care experiences. Future studies are required to explore the underlying reasons for the racial-ethnic differences in physician communication, access to care and office staff support. These studies should aim to inform care providers and payers about underlying reasons for differences and help to tailor quality improvement initiatives to address racial-ethnic disparities in care.

Tables and figures

Figure 4.1 Three-factor model

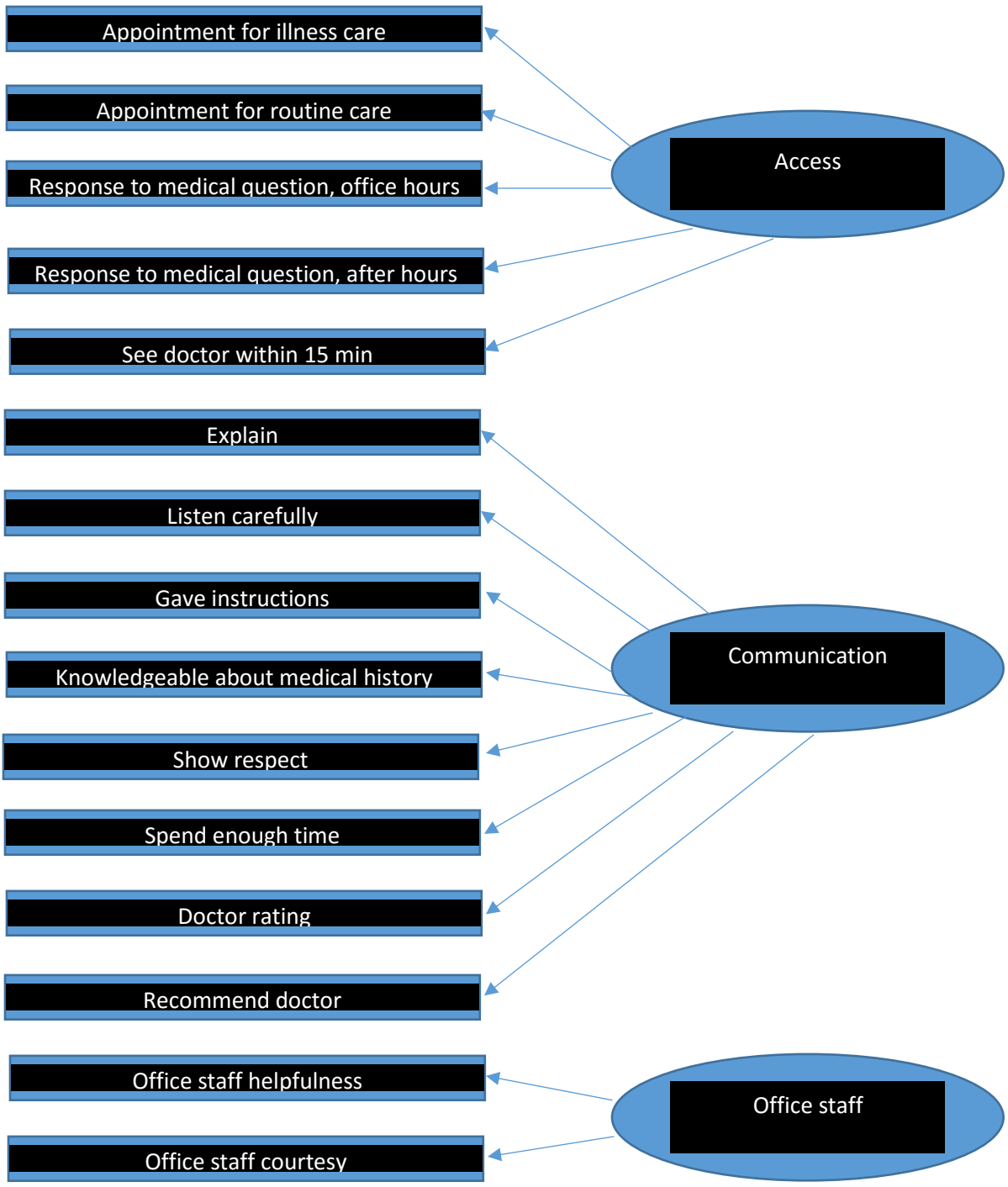

Table 4.1 Reports about care and global rating items

Access to care

11. In the last 12 months, when you phoned this doctor's office to get an appointment for care you needed right away, how often did you get an appointment as soon as you thought you needed?

12. In the last 12 months, when you made an appointment for a check-up or routine care with this doctor, how often did you get an appointment as soon as you thought you needed?

13. In the last 12 months, when you phoned this doctor's office during regular office hours, how often did you get an answer to your medical question that same day?

14. In the last 12 months, when you phoned this doctor's office after regular office hours, how often did you get an answer to your medical question as soon as you needed?

15. Wait time includes time spent in the waiting room and exam room. In the last 12 months, how often did you see this doctor within 15 minutes of your appointment time?

Physician communication

15. During your most recent visit, did this doctor explain things in a way that was easy to understand?

16. During your most recent visit, did this doctor listen carefully to you?

17. During your most recent visit, did this doctor give you easy to understand instructions about taking care of these health problems or concerns?

18. During your most recent visit, did this doctor seem to know the important information about your medical history?

19. During your most recent visit, did this doctor show respect for what you had to say?

20. During your most recent visit, did this doctor spend enough time with you?

Table 4.1 Reports about care and global rating items (continued)

Office staff courtesy and helpfulness 5. During your most recent visit, were clerks and receptionists at this doctor's office as helpful as you thought they should be? 6. During your most recent visit, did clerks and receptionists at this doctor's office treat you with courtesy and respect?
Doctor rating • Using any number from 0 to 10, where 0 is the worst doctor possible and 10 is the best doctor possible, what number would you use to rate this doctor?
Recommending doctor to friends/family • Would you recommend this doctor's office to your family and friends?

Note: *Physician communication* and *office staff courtesy and helpfulness* composite items and the *recommending doctor to friends/family* question use a three response option scale (*Yes, definitely; Yes, somewhat; No*). *Access to care* composite items use a four response option scale (*Always, Usually, Sometimes, Never*). The *doctor rating* question uses a 0-10 response scale.

Table 4.2 Gender, Age, Education, Health and Region for Asian Americans and non-Hispanic whites

	Asian	White
Gender		
	n = 5,119	*n = 216,041*
Male	35%	36%
Female	65%	64%
Age		
	n = 5,103	*n = 216,405*
18 to 24	6%	3%
25 to 34	18%	8%
35 to 44	17%	9%
45 to 54	17%	17%
55 to 64	19%	25%
65 to 74	14%	22%
75 and older	11%	17%
Education		
	n = 5,053	*n = 214,604*
8th grade or less	9%	1%
Some high school	5%	4%
High school graduate	12%	25%
Some college	20%	31%
4-year college graduate	25%	18%
> 4-year college	30%	21%
Health		
	n = 5,112	*n = 215,090*
Poor	4%	3%
Fair	12%	14%
Good	32%	34%
Very good	35%	36%
Excellent	16%	13%
Region		
	n = 5,170	*n = 217,458*
Midwest	49%	56%
Northeast	6%	25%
South	5%	7%
West	40%	12%

Note: Percentages are rounded off to the nearest one.

Table 4.4 Model fit indices for single group CFA and MG-CFA analyses models, WLSMV, MLR and ML estimation methods

	Free Parameters	Chi-Square Test of Model Fit	Degrees of Freedom	P-Value	Root Mean Square Error Of Approximation	Comparative Fit Index	Tucker-Lewis Index
Single group CFA							
Asian Americans	54.00	702.08	87.00	0.00	0.037	0.991	0.989
Non-Hispanic whites	54.00	17993.25	87.00	0.00	0.031	0.994	0.992
Multiple Group CFA, WLSMV*							
Configural	108.00	18326.90	174.00	0.00	0.031	0.993	0.992
Metric	96.00	15416.08	186.00	0.00	0.027	0.994	0.994
Scalar	63.00	15885.50	219.00	0.00	0.025	0.994	0.995
Multiple Group CFA, WLSMV**							
Configural	108.00	18326.90	174.00	0.00	0.031	0.993	0.992
Metric	93.00	11889.05	189.00	0.00	0.024	0.996	0.995
Scalar	57.00	14248.89	225.00	0.00	0.024	0.995	0.995
Multiple Group CFA, MLR*							
Configural	96.00	33147.88	174.00	0.00	0.041	0.945	0.933
Metric	84.00	32773.24	186.00	0.00	0.040	0.945	0.938
Scalar	72.00	33976.08	198.00	0.00	0.039	0.943	0.940
Multiple Group CFA, MLR**							
Configural	96.00	33147.88	174.00	0.00	0.041	0.945	0.933
Metric	81.00	32627.64	189.00	0.00	0.039	0.945	0.939
Scalar	66.00	34865.43	204.00	0.00	0.039	0.942	0.940

Note: * factor means fixed to 0 in both groups and factor variances fixed to 1 in one group and freed in the second group; ** factor means fixed to 0 and factor variances fixed to 1 in both groups

Table 4.4 Model fit indices for single group CFA and MG-CFA analyses models, WLSMV, MLR and ML estimation methods (continued)

	Free Parameters	Chi-Square Test of Model Fit	Degrees of Freedom	P-Value	Root Mean Square Error Of Approximation	Comparative Fit Index	Tucker-Lewis Index
Multiple Group CFA, ML*							
Configural	96.00	61099.59	174.00	0.00	0.056	0.950	0.939
Metric	84.00	61359.40	186.00	0.00	0.054	0.950	0.943
Scalar	72.00	61858.19	198.00	0.00	0.053	0.949	0.946
Multiple Group CFA, ML**							
Configural	96.00	61099.59	174.00	0.00	0.056	0.950	0.939
Metric	81.00	61963.07	189.00	0.00	0.054	0.949	0.943
Scalar	66.00	63750.85	204.00	0.00	0.053	0.948	0.946

Note: * factor means fixed to 0 in both groups and factor variances fixed to 1 in one group and freed in the second group; ** factor means fixed to 0 and factor variances fixed to 1 in both groups

Table 4.5 Standardized parameter estimates by race/ethnicity, single group CFA

Factor/items	Race/ethnicity	Estimate	Standard Errors	P-value
Access				
Appointment for illness care				
	White	0.868	0.003	0.001
	Asian	0.882	0.014	0.001
Appointment for routine care				
	White	0.822	0.003	0.001
	Asian	0.839	0.013	0.001
Response to medical question, office hours				
	White	0.767	0.003	0.001
	Asian	0.780	0.017	0.001
Response to medical question, after hours				
	White	0.816	0.008	0.001
	Asian	0.891	0.026	0.001
See doctor within 15 minutes				
	White	0.554	0.003	0.001
	Asian	0.569	0.015	0.001
Communication				
Explain in easy to understand ways				
	White	0.902	0.001	0.001
	Asian	0.877	0.009	0.001
Listen carefully				
	White	0.963	0.001	0.001
	Asian	0.943	0.007	0.001
Gave instruction in easy to understand ways				
	White	0.893	0.002	0.001
	Asian	0.884	0.01	0.001
Knowledgeable about medical history				
	White	0.773	0.002	0.001
	Asian	0.781	0.011	0.001
Show respect				
	White	0.959	0.001	0.001
	Asian	0.954	0.007	0.001
Spend enough time				
	White	0.882	0.002	0.001
	Asian	0.865	0.009	0.001
Doctor rating				
	White	0.846	0.001	0.001
	Asian	0.816	0.007	0.001
Recommend doctor				
	White	0.916	0.001	0.001
	Asian	0.877	0.008	0.001

Table 4.5 Standardized parameter estimates by race/ethnicity, single group CFA (continued)

Factor/items	Race/ethnicity	Estimate	Standard Errors	P-value
Office staff helpfulness				
Office staff helpfulness				
	White	0.985	0.003	0.001
	Asian	0.951	0.014	0.001
Office staff courtesy				
	White	0.942	0.003	0.001
	Asian	0.940	0.013	0.001

Table 4.6 Standardized parameter estimates, MG CFA Scalar invariance testing*

Factor/items	Estimate	Standard Errors	P-value
Access			
Appointment for illness care	0.869	0.003	0.000
Appointment for routine care	0.822	0.003	0.000
Response to medical question, office hours	0.768	0.003	0.000
Response to medical question, after hours	0.822	0.008	0.000
See doctor within 15 minutes	0.555	0.003	0.000
Communication			
Explain in easy to understand ways	0.901	0.001	0.000
Listen carefully	0.964	0.001	0.000
Gave instruction in easy to understand ways	0.893	0.002	0.000
Knowledgeable about medical history	0.773	0.002	0.000
Show respect	0.958	0.001	0.000
Spend enough time	0.881	0.002	0.000
Doctor rating	0.845	0.001	0.000
Recommend doctor	0.915	0.001	0.000
Office staff helpfulness			
Office staff helpfulness	0.983	0.003	0.000
Office staff courtesy	0.943	0.003	0.000

* WLSMV estimation, factor means fixed to 0 and factor variances fixed to 1 in both groups

References

1. U.S. Census Bureau. *United States Census 2010*. 2011 [cited 2016 August 20]; Available from: http://2010.census.gov/news/releases/operations/cb11-cn125.html

2. Medicine, C.o.R.H.I.P.M.I.o., *Rewarding Provider Performance: Aligning Incentives in Medicare*. 2006: National Academies Press.

3. Rodriguez, H.P., et al., *The Effect of Performance-Based Financial Incentives on Improving Patient Care Experiences: A Statewide Evaluation.* Journal of General Internal Medicine, 2009. **24**(12): p. 1281-1288.

4. *Centers for Medicare and Medicaid Services*. [cited 2016; Available from: https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/AnalysisAndPayment.html.

5. Morales, L., 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 Services Research, 2001. **36**(3): p. 595-617.

6. Goldstein, E., M. Elliott, and W. Lehrman, Hambarsoomian, K., Giordano, L., *Racial/Ethnic Differences in Patients' Perceptions of Inpatient Care Using the HCAHPS Survey.* Medical Care Research and Review, 2010. **67**(1): p. 74-92.

7. Haviland, A., M. Elliott, and R. Weech-Maldonado, Hambarsoomian, K., Orr, N., & Hays, R. D. , *Racial/ethnic disparities in Medicare Part D experiences.* Medical Care, Supplement, 2012. **59**.

8. Weech-Maldonado, R., et al., *Health Plan Effects on Patient Assessments of Medicaid Managed Care Among Racial/Ethnic Minorities.* Journal of General Internal Medicine, 2004. **19**(2): p. 136-145.

9. *The Rise of Asian Americans*, in *Washington, DC: Pew Social & Demographic Trends*. 2012.

10. Ahmedov, M., *Chapter 2: Differences in perceptions of health care between Asian Americans and non-Hispanic whites: an analysis of the CAHPS Clinician and Group Adult Visit Survey 1.0.*, in *Health policy and management*. 2016, University of California, Los Angeles.

11. Stewart, A.L., xe, and A.M. poles-Springer, *Advancing Health Disparities Research: Can We Afford to Ignore Measurement Issues?* Medical Care, 2003. **41**(11): p. 1207-1220.

12. Millsap, R.E., *Statistical approaches to measurement invariance*. 2012: Routledge.

13. Wu, A.D., Z. Li, and B.D. Zumbo, *Decoding the meaning of factorial invariance and updating the practice of multi-group confirmatory factor analysis: A demonstration with TIMSS data.* Practical Assessment, Research and Evaluation, 2007. **12**(3): p. 1-26.

14. Meredith, W. and J.A. Teresi, *An essay on measurement and factorial invariance.* Medical care, 2006. **44**(11): p. S69-S77.

15. Zhu, J., et al., *Racial/ethnic disparities in patient experience with communication in hospitals: real differences or measurement errors?* Medical care, 2015. **53**(5): p. 446-454.

16. McHorney, C.A. and J.A. Fleishman, *Assessing and understanding measurement equivalence in health outcome measures: issues for further quantitative and qualitative inquiry.* Medical care, 2006. **44**(11): p. S205-S210.

17. Carle, A.C., et al., *Evaluating Measurement Equivalence across Race and Ethnicity on the CAHPS® Cultural Competence Survey.* Medical care, 2012. **50**(9 Suppl 2): p. S32.

18. Carle, A.C. and R. Weech-Maldonado, *Does the Consumer Assessment of Healthcare Providers and Systems Cultural Competence Survey provide equivalent measurement across English and Spanish versions?* Medical care, 2012. **50**(9 Suppl 2): p. S37.

19. Hays, R.D., et al., *Measurement equivalence of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Medicare survey items between Whites and Asians.* Quality of Life Research, 2016: p. 1-8.

20. *Agency for Healthcare Research and Quality CAHPS® Survey Products Web Site*. [cited 2016; Available from: http://www.ahrq.gov/cahps/index.html.

21. Fowler, F., P. Gallagher, and S. Nederend, *Comparing Telephone and Mail Responses to the CAHPS(TM) Survey Instrument.* Medical Care, 1999. **37**(3): p. MS41-MS49.

22. Hepner, K., J. Brown, and R. Hays, *Comparison of Mail and Telephone in Assessing Patient Experiences in Receiving Care from Medical Group Practices.* Evaluation & the Health Professions, 2005. **28**(4): p. 377-389.

23. de Vries, H., M. Elliott, and K. Hepner, Keller, S., Hays, R., *Equivalence of Mail and Telephone Responses to the CAHPS® Hospital Survey.* Health Services Research, 2005. **40**(6p2): p. 2120-2139.

24. Ahmedov, M., *Chapter 3: Evaluation of a factor structure of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Clinician & Group Adult Visit 1.0 Survey*, in *Health policy and management*. 2016, University of California, Los Angeles.

25. Millsap, R.E. and J. Yun-Tein, *Assessing factorial invariance in ordered-categorical measures.* Multivariate Behavioral Research, 2004. **39**(3): p. 479-515.

26. Muthén, L.K. and B.O. Muthén, *MPLUS (7).* Computer software). Los Angeles: Muthén & Muthén. http://statmodel. com, 2012.

27. Muthén, B. and T. Asparouhov, *Latent variable analysis with categorical outcomes: Multiple-group and growth modeling in Mplus.* Mplus web notes, 2002. **4**(5): p. 1-22.

28. Brown, T.A., *Confirmatory factor analysis for applied research*. 2015: Guilford Publications.

29. Cheung, G.W. and R.B. Rensvold, *Evaluating goodness-of-fit indexes for testing measurement invariance.* Structural equation modeling, 2002. **9**(2): p. 233-255.

30. Sass, D., *Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework.* Journal of Psychoeducational Assessment, 2011: p. 0734282911406661.

31. Steiger, J.H., *Understanding the limitations of global fit assessment in structural equation modeling.* Personality and Individual differences, 2007. **42**(5): p. 893-898.

32. Chen, F., et al., *An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models.* Sociological methods & research, 2008. **36**(4): p. 462-494.

33. Hu, L.t. and P.M. Bentler, *Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives.* Structural equation modeling: a multidisciplinary journal, 1999. **6**(1): p. 1-55.

34. Hooper, D., J. Coughlan, and M. Mullen, *Structural equation modelling: Guidelines for determining model fit.* Articles, 2008: p. 2.

35. Van de Schoot, R., P. Lugtig, and J. Hox, *A checklist for testing measurement invariance.* European Journal of Developmental Psychology, 2012. **9**(4): p. 486-492.

36. Hurtado, M.P., et al., *Assessment of the Equivalence of the Spanish and English Versions of the CAHPS® Hospital Survey on the Quality of Inpatient Care.* Health Services Research, 2005. **40**(6p2): p. 2140-2161.

Chapter 5 Conclusion

Patient experiences with care are increasingly important indicators of quality care and have become an indelible part of measurement metrics used to evaluate health plan and provider performance and guide pay-for-performance programs. Similar to other quality of care domains, racial/ethnic disparities are evident in patient experiences with care. In particular, Asian Americans (Asians) are consistently shown to report worse care experiences compared with non-Hispanic Whites (Whites). In light of the increasing number of Asians in the US population, disparities in reports about care for Asians have gained prominence. It is uncertain to what extent worse perceptions of care by Asians are due to "real" differences in care or differential response tendencies.

In this study, the experiences of care of Asians were compared to Whites on the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Clinician & Group (CG-CAHPS) Adult Visit Survey data collected in 2011.

The first study in this dissertation compared mean scores on CAHPS composites (scales) and global ratings of care between Asians and Whites. The results of this comparison were similar to those found in many prior studies, with Asians reporting worse experiences than Whites. Given the similarity to prior work, it was then important to evaluate whether the measurement structure underlying the scoring of the CAHPS survey was confirmed in these data. The second study evaluated the hypothesized three-factor structure in the overall sample. Finally, the third study evaluated whether the factor structure was equivalent for Asians and Whites, indicating measurement equivalence for the two subgroups.

Study findings

In the first study, ordinary least squares regression models and generalized ordinary logistic regression models as a sensitivity analysis for global ratings were estimated. Asians reported worse care experiences on *access to care, office staff courtesy and helpfulness, rated their doctor* lower and were less likely to *recommend their doctor to family and friends* than Whites. There was no significant difference between Asians and Whites on reports of *physician communication*.

The second study confirmed the hypothesized three factor structure (*access to care, physician communication and global ratings* and *office staff helpfulness and courtesy*) on the CG-CAHPS Adult Visit Survey 1.0 survey using individual-level categorical confirmatory factor analyses.

In the third study, single group and multiple group confirmatory factor analyses were carried out to assess measurement invariance between Asians and Whites. The single group individual CFA analyses showed acceptable fit indices, supporting the hypothesized three factor structure in both subgroups. The multiple group CFA (MG-CFA) to evaluate the factor structure was estimated using Mean- and Variance-adjusted Weighted Least Squares (WLSMV) estimation and the polychoric correlation matrix. The study focused on scalar invariance (equality of factor loadings and thresholds) between Asians and Whites. The findings of the study supported scalar invariance between the two racial/ethnic groups.

Overall, this dissertation shows racial/ethnic differences between Asians and Whites on the CG-CAHPS Adult Visit Survey 1.0 survey. Asians reported worse experiences on global ratings of care as well as reports about *access to care* and *office staff courtesy and helpfulness*, but reports about *physician communication* were similar. Findings from the measurement invariance

evaluation showed that these reported racial/ethnic differences were likely due to real differences in patient experiences rather than lack of measurement equivalence.

Contributions, limitations and implications

The three studies make an important contribution to the field of racial/ethnic disparities in patient experiences in several ways. This dissertation is unique in comparing patient experiences between Asians and Whites on the CAHPS Clinician & Group Adult Visit 1.0 Survey. While prior studies have often combined Asians with Pacific Islanders, this dissertation had a sufficiently large sample of Asians in the dataset to study them as a separate group.

While the factor structure in the CAHPS Clinician & Group Adult Visit 1.0 Survey were evaluated in two other studies, they both had important limitations [1, 2]. Dyer and colleagues analyzed only non-missing data (those with appropriately missing data were dropped), which reduced the sample size significantly (21,318 vs. 103,442). Analyzing only non-missing data may introduce bias as respondents with non-missing values could be different from those with missing data. Rodriguez and colleagues analyzed data collected from eight southern California medical groups. In contrast to the two studies, this dissertation analyzed the full sample and the data come from survey sponsors across the nation. The third paper is also the first study to evaluate measurement invariance between Asians and Whites on the CG-CAHPS Adult Visit Survey 1.0 using MG-CFA estimation methods for categorical data.

Despite these contributions, it is important to acknowledge the limitations of the studies. First, Asians are a heterogeneous group consisting of various subgroups with different expectations and response styles. The CAHPS surveys do not collect data on Asian subgroups, therefore the dissertation was not able to carry out analyses at the subgroup level. Also, the survey was carried out in two languages (English and Spanish) effectively leaving out Asians with limited English

proficiency. Earlier studies show that Asians with limited proficiency may have different response tendencies than Asians who are proficient in English [1].

The dissertation findings have several important policy and methodological implications for the field and provide directions for future research. The study findings suggest that racial/ethnic difference in patient experiences in ambulatory care are real – Asians are likely to have worse care experiences in ambulatory care than Whites. Future research is needed to understand the reasons why Asians are treated by care providers differently. Findings from this research should be used to inform quality improvement initiatives aimed at reducing racial/ethnic disparities in care. Future research will also be needed to collect and analyze data on Asian American subgroups, so that quality improvement initiatives can be targeted at those subgroups that are most at disadvantage. Our findings also show that future research is needed in order to improve research methods used in evaluation of measurement invariance in large datasets with categorical survey items. For instance, in this dissertation, chi square based model fit indices were less informative given the large size. Also, chi square values in nested models in this dissertation did not conform to typically expected patterns when using WLSMV and MLR estimation methods. These two factors limited the use of nested model fit indices in the evaluations.

References

1. Rodriguez, H. and P. Crane, *Examining Multiple Sources of Differential Item Functioning on the Clinician & Group CAHPS® Survey.* Health Services Research, 2011. **46**(6pt1): p. 1778-1802.

2. Dyer, N., J. Sorra, and S. Smith, Cleary, P., & Hays, R. , *Psychometric properties of the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) clinician and group adult visit survey.* Medical Care, Supplement, 2012.