

Patient-Reported Experience of Care & Assessment of Provider Performance

Sukyung Chung, PhD

Palo Alto Medical Foundation Research Institute

Objectives

- Review definition and measurement of patient satisfaction
- Understand how patient satisfaction surveys are used in assessing provider performance: issues and challenges – experience of Palo Alto Medical Foundation

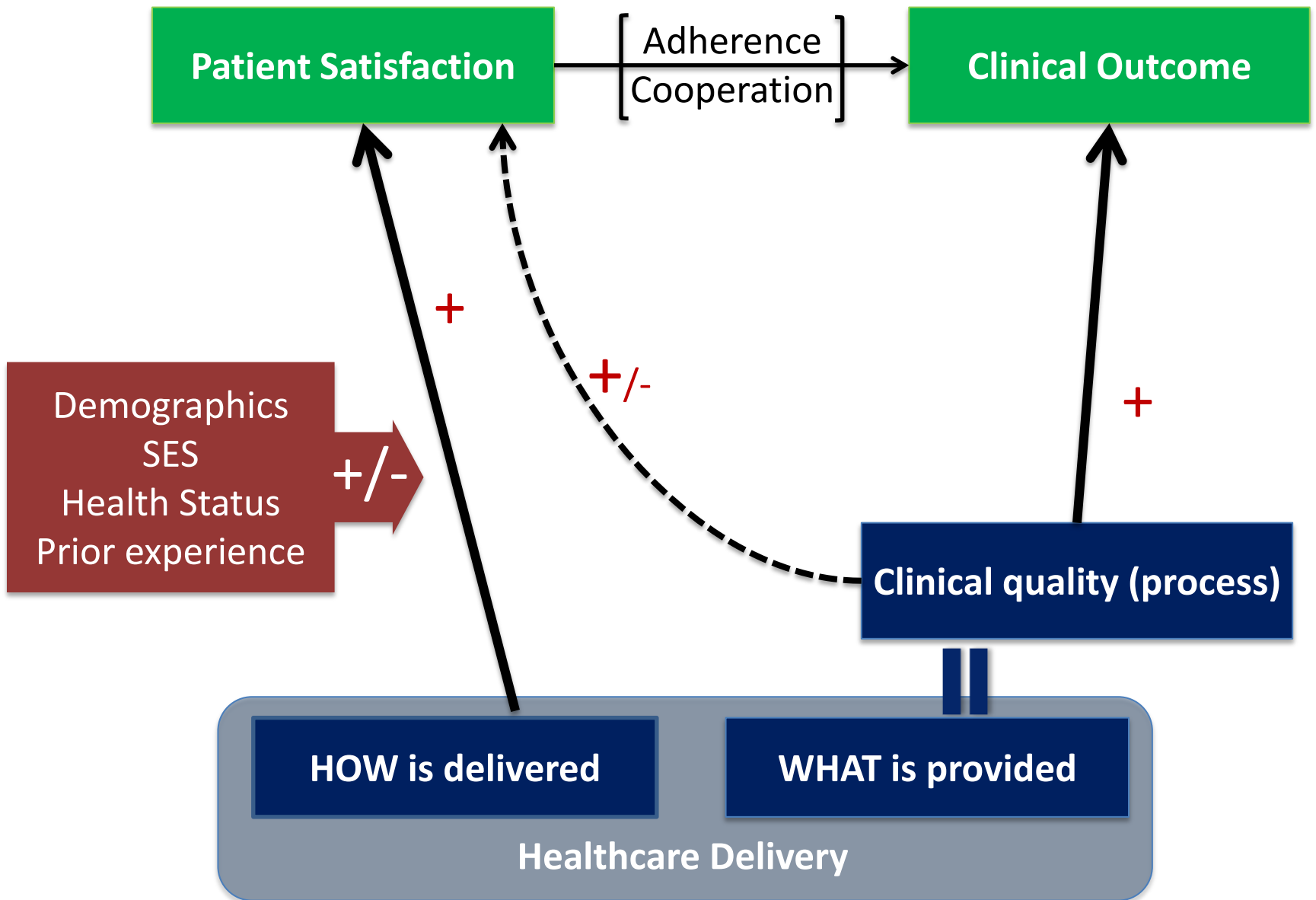
Why do we care about patient satisfaction?

- A measure of care quality from patient perspectives
 - (Analogous to patient-reported experience of care)
 - An indicator of “patient-centeredness” of care
 - Viewed as health outcome in its own right
- Leads to better clinical quality
 - Satisfied patients are more likely to be cooperative and to adhere treatment regimens
 - Positive association with inpatient mortality, readmission, length of hospital stay, guideline adherence, self-management of chronic conditions

Doyle et al. 2013; Boulding et al. 2011, Glickman et al. 2010, Isaac 2010, Jha et al. 2008, Manary et al. 2013, Gupta et al 2013, Wong et al 2008, Zantan et al. 2012, Heisler et al. 2012, Kaplan et al. 1989, Heather et al. 2013, Chang et al. 2006, Gray et al. 2014

Why do we care about patient satisfaction? (cont.)

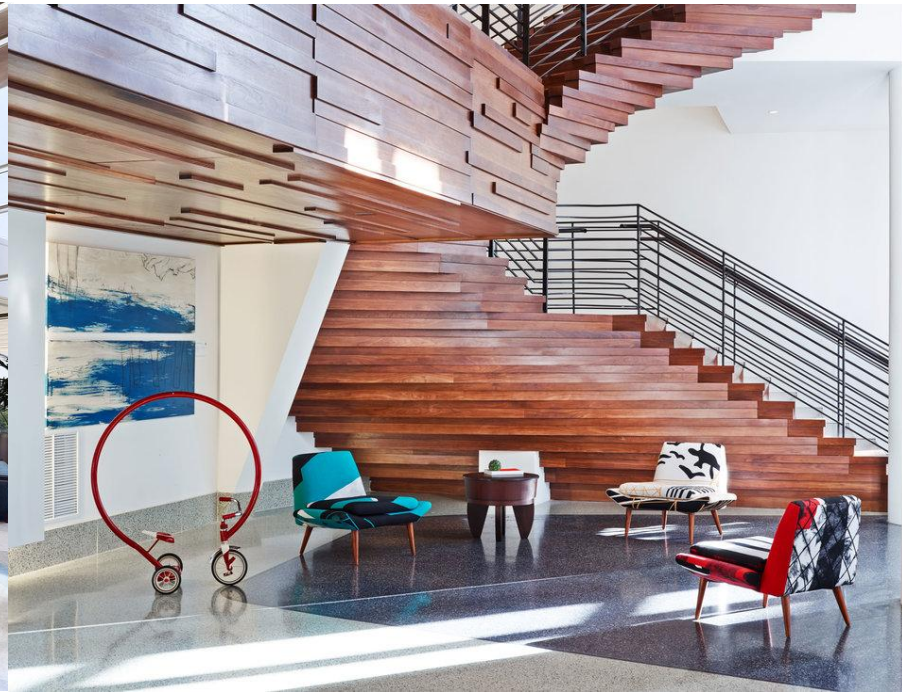
- High satisfaction does not always mean better quality
 - Focusing on patient satisfaction may divert attention from clinically appropriate care
 - Sometimes inappropriate overuse or underuse of necessary care is preferred by patients
- Better satisfaction means lower liability exposure



Hospital or Hotel?



Jeffrey Totaro for RTKL Associates Inc



Eric Laignel for Lauren Rottet of Rottet Studio

Hospital or Hotel?



한국경제매거진



연합뉴스

What do patient satisfaction *surveys* typically measure?

- Access
 - Seeing preferred providers, convenience of office hours and scheduling, timely care
- Care process
 - Wait time, information about delays, prompt attention, coordination/referral
- Provider communication
 - Ease of understanding explanations, shared decision making, affective/respectful, follow-up/self-management instructions
- Staff courtesy
 - Friendly, respectful, helpful office staff, respect privacy
- Structure, etc.
 - Basic amenity, cleanliness
- Overall satisfaction of the visit or stay

Where are patient satisfaction surveys used?

- Internal quality improvement
 - To understand/monitor patients' experience of care
 - To assess health care delivery interventions
- Comparing across organizations
 - Provider performance indicator in public reporting, P4P
 - US examples: Hospital Compare, Hospital Value Based Purchasing, Medicare Shared Savings Program, Physician Group Value-Based Payment Modifier Program
- International comparison
 - World Health Survey of Health System Responsiveness
 - European Task Force on Patient Evaluations of General Practice

Methodological considerations in *developing survey measures*

- Directness of measurement
 - “Are you satisfied with communication with the provider?” vs.
 - “Did the provider explain what to do if problems or symptoms continued, got worse, or came back?”
- Specificity (particular visit or general impressions)
 - General assessment
 - 3/6/12-month reference period
 - Visit-specific
- Dimensionality/scope (#aspects of care ask about)

Methodological considerations in *survey methods*

- Method, moment, place
 - Comment box, in-person, phone, mail, online
 - Hospital/clinic, home
 - Timing (how long after satisfaction was measured)
- Sampling
 - Random sampling of [*unit of analysis*]
- Number of surveys, considering
 - Response rate
 - Representativeness of survey respondents
 - Costs

Methodological considerations in *analysis and reporting*

- Representativeness
 - Sample: limited sample size per provider
 - Responses: from unusually (dis)satisfied patients
 - Comments are useful but tend to be more biased
- Ceiling effects
 - Rank plunges by few lower than top ratings
- My patients are sicker [grumpier]
 - Responses are subject to patients' demographic and cultural backgrounds and clinical conditions

Development of HCAHPS (CMS/AHRQ)

The initial input for measures: CAHPS Health Plan Survey as a prototype, additional "Call for Measures" (June 2003), web chat questions and comments, stakeholders' meeting, vendors' meeting, responses to an electronic mailbox, literature review, cognitive testing

OMB clearance of the initial draft: draft instrument submitted this to the U.S. Office of Management and Budget (OMB). In the clearing process, seek comments on the draft instrument and input about implementation options. Six times Iteratively.

Pilot testing: in 132 hospitals, ~20,000 completed surveys

Focus groups: 7 focus groups in various hospital sites

Additional field testing: 375 volunteer hospitals over a 6-month period

Pre-implementation testing: to identify ways to minimize the potential burden and disruption posed by this survey. Researchers investigated various approaches to integrating the survey items into existing questionnaires as well as alternative protocols for administering the survey.

Endorsement by NQF: Submitted a 25-item instrument to the National Quality Forum's (NQF) review and consensus-building process for endorsement. An NQF committee made some recommendations. After public review and comment on the 27-item version, NQF endorsed the CAHPS Hospital Survey as a measure in May 2005.



Use of Patient Satisfaction Scores in Assessing Provider Performance: Experience from Palo Alto Medical Foundation

Where does PAMF use patient satisfaction survey data?

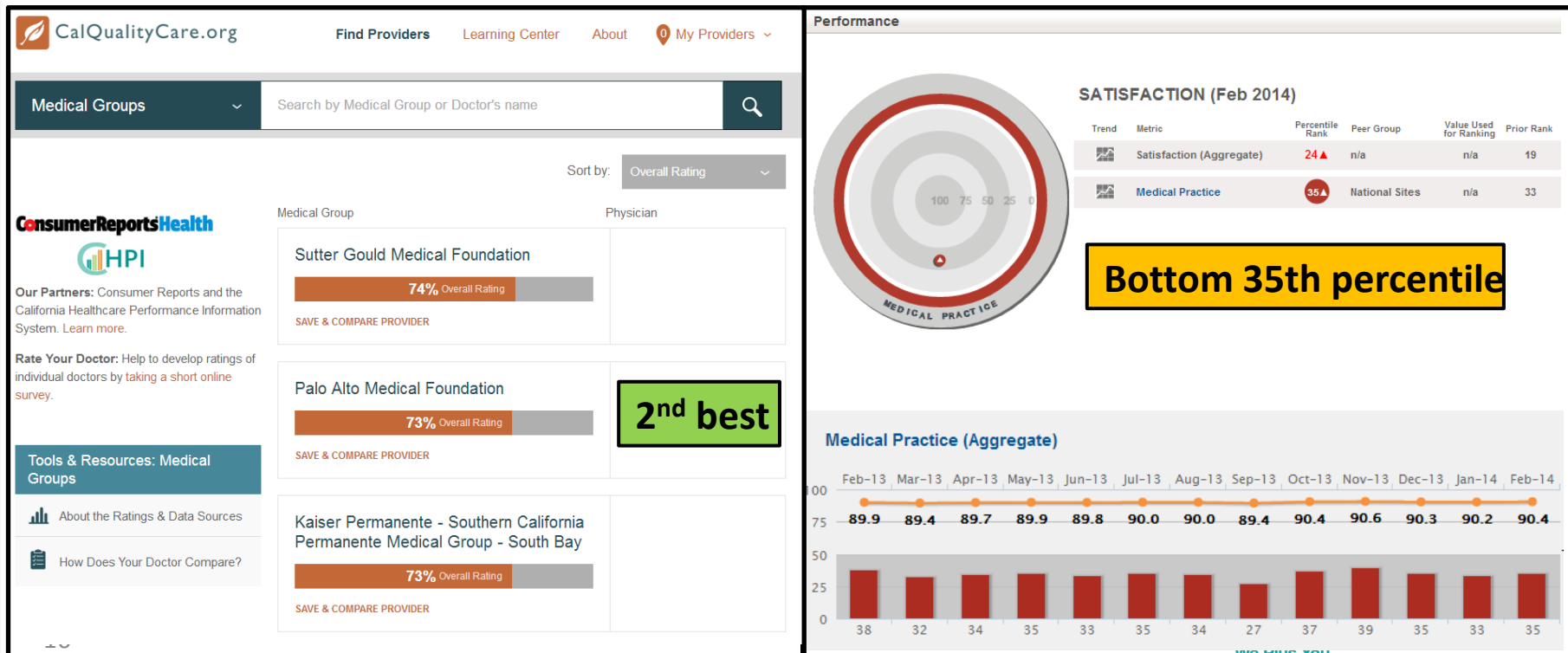
- Participate in national and regional initiatives
 - Public reporting, P4P, VBP programs
- Provider/group performance assessment at PAMF
 - Rate/rank individual providers
 - Distribution of financial incentives to department and region
- Monitoring and assessing QI effort
 - *Lean* implementation

Motivation

- PAMF strives for high quality, patient-centered care, but national ranking of patient satisfaction score is low

California ranking

National ranking



Research questions

- To identify the contribution of race (Asians in particular)
 - As compared to other demographic factors
 - Within a clinic/provider
 - On provider ratings and ranking
- To explore potential pathways of racial differences
 - Actual care received
 - Preferences, expectations
 - Response effect

Demographics and patient satisfaction literature

- Difference by demographic/cultural background such as age, education, sex, marital status, race/ethnicity
 - Race/ethnicity is one of the strongest predictor among these factors, particularly for Asian race
 - Asians rate their health care experience lower than do non-Hispanic whites (NHWs) in the same settings
- Cultural norms in rating similar experiences
 - Reticence to select extremes in survey response
 - Even with much lower ratings in satisfaction, Asian patients were not more likely to change providers

Literature: patient-provider sorting

- Lower ratings by patients of particular demographic (e.g., Asians) can disproportionately affect providers in a group or region
 - Recent Asian immigrants often live in racial enclaves, so some providers may have more proportion of Asian patients than others in nearby locations.
 - Patients may prefer a physician of the same race due to concerns about language, empathic treatment, and personal preference.

Survey data

- Patient satisfaction surveys
 - Press-Ganey survey collected during 2011-2014
 - CG-CAHPS collected 2013-2014
- Surveys were mailed to patients of randomly selected visits
 - 30 returned surveys/provider /6months (or 60 per year)
- Total N=197,479 surveys available as of March 2014

Press-Ganey survey questions

Access (8 items)	courtesy of staff, ease of getting clinic on the phone, ease of obtaining referrals, helpfulness on the phone, promptness in returning calls/messages
Moving Through the Visit (2 items)	wait time, informed about any delays
Nurse/Assistant (2 items)	friendliness, concern shown for patient problems
Care Provider (10 items)	friendliness, concern for questions/worries, efforts to include the patient in decisions, clear/understandable language/instruction, time spent with patients, confidence in the provider, overall recommendation of the provider
Personal Issues (4 items)	staff protect safety, sensitivity to patient needs/privacy, cleanliness
Billing (4 items)	accuracy/clarity of billing statement , handling of billing questions, courtesy of billing personnel
Overall Assessment (2 items)	how well the staff worked together, overall recommendation of the practice

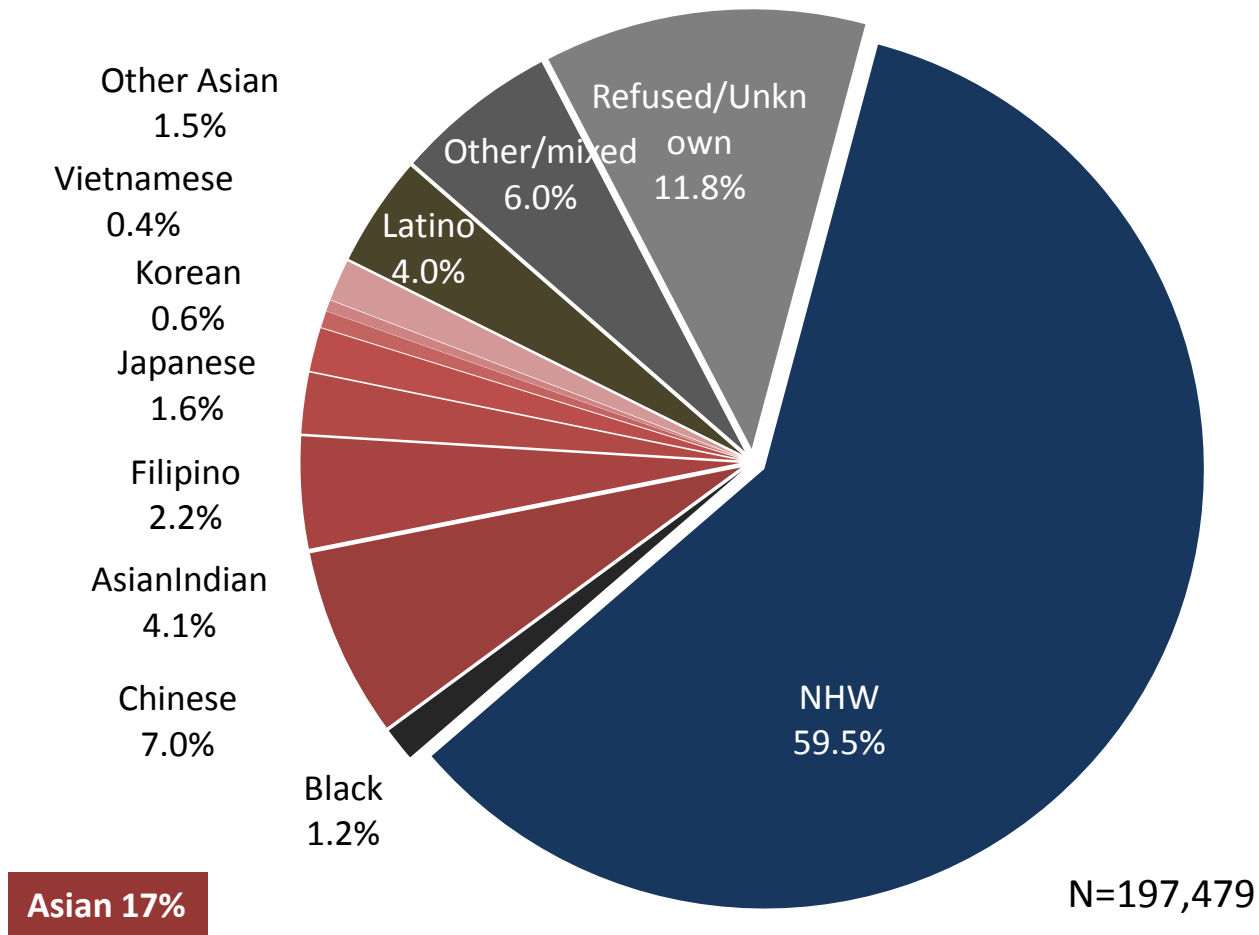
Selected CG-CAHPS questions

- "Would you recommend this provider's office to your family and friends?"
 1. Yes, definitely; 2. Yes, somewhat; 3. No
- Using any number from 0 to 10, where 0 is the worst provider possible and 10 is the best provider possible, what number would you use to rate this doctor?
 - 0 Worst provider possible ... 10 Best provider possible
- Compared to similar questions in Press-Ganey survey in 5-category Likert scale:
 1. Very Poor; 2. Poor; 3. Fair; 4. Good; 5. Very Good

Results

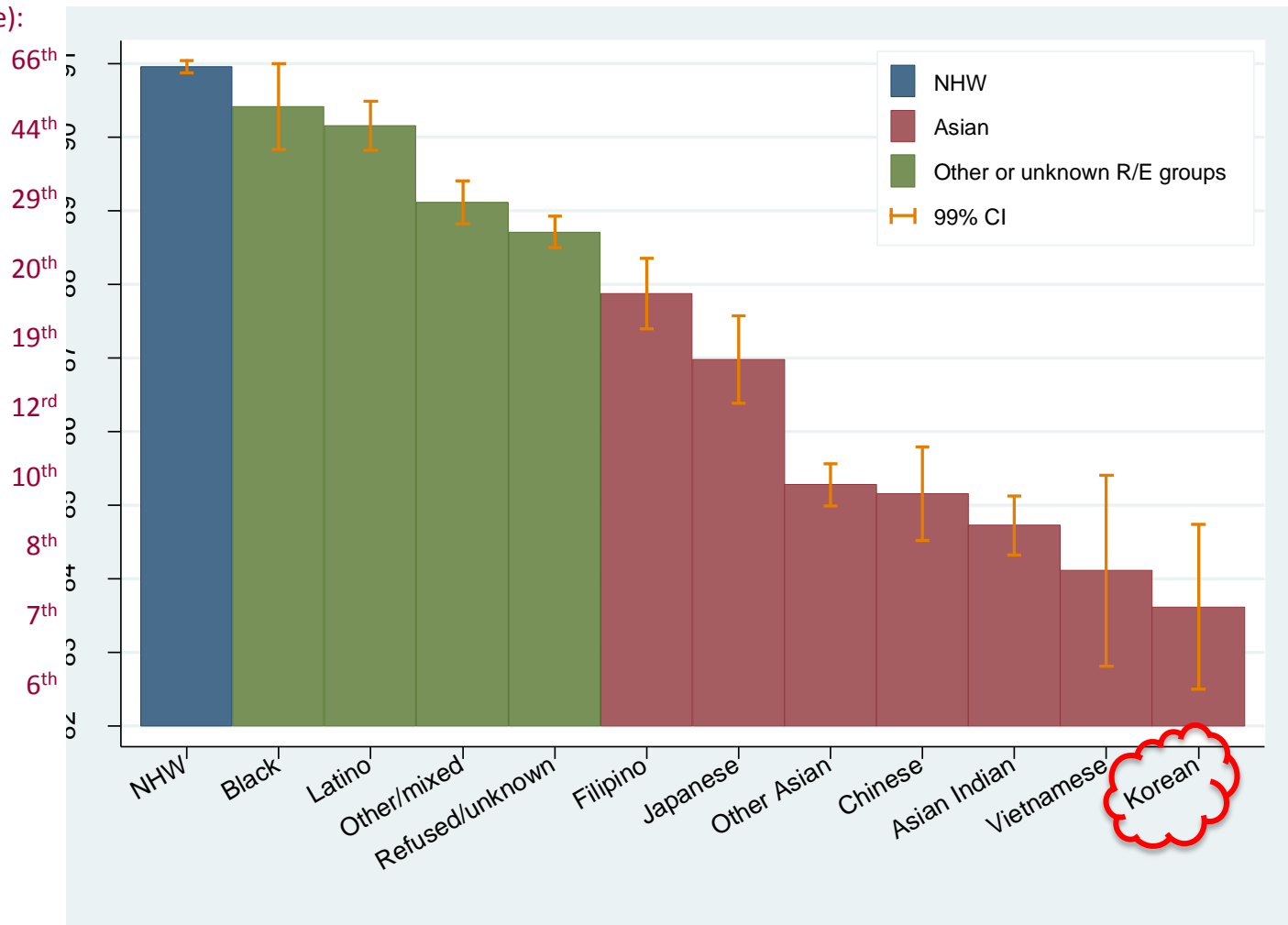
Variation across providers: role of patient race/ethnicity

Respondents racial/ethnic composition

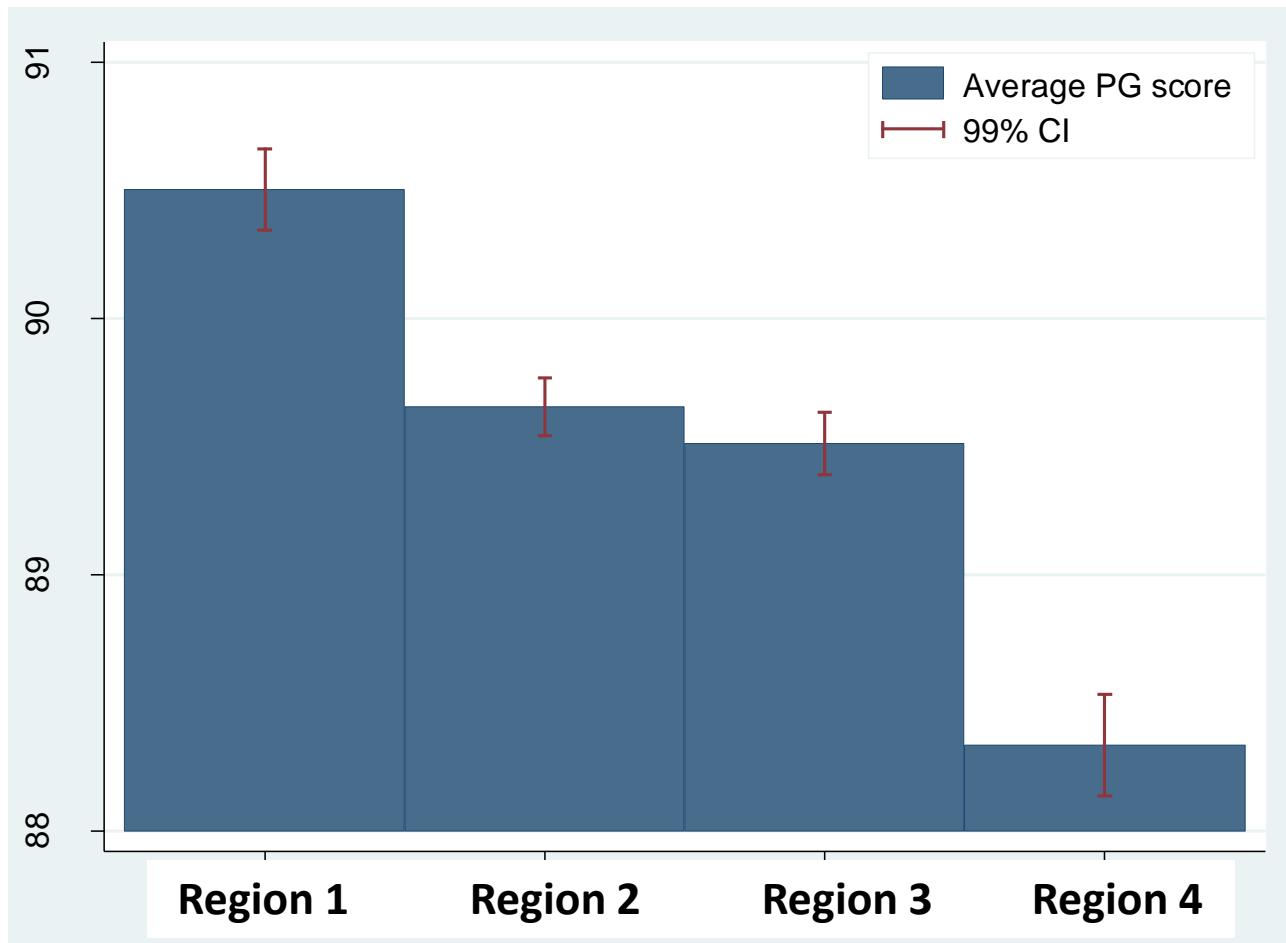


Average Press-Ganey Score by R/E

Corresponding national ranking (in percentile):



Regional difference in Press-Ganey Score



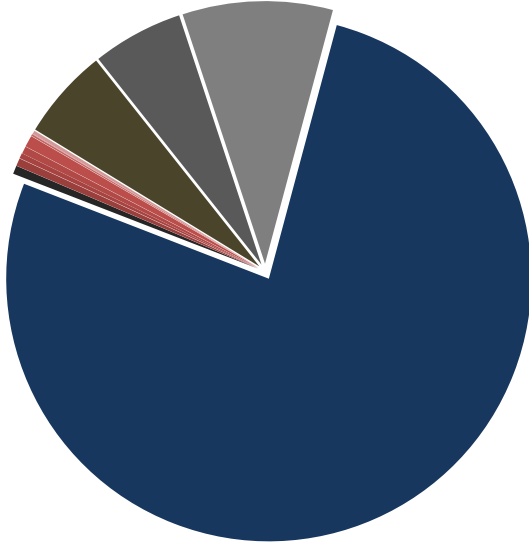
Palo Alto Medical Foundation Medical Center Locations



R/E Composition (% Asian) of Each Region

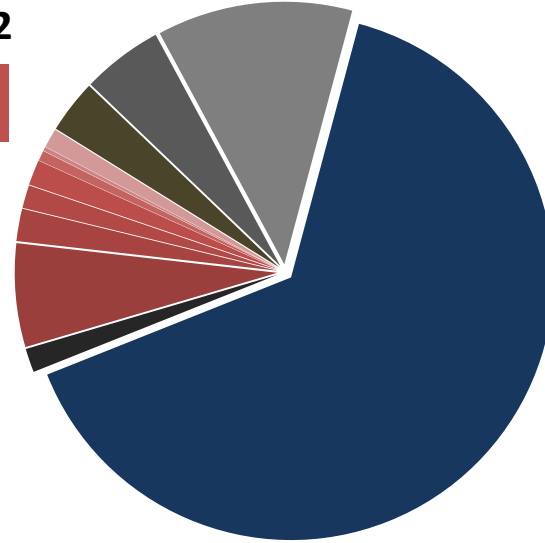
Region 1

2%



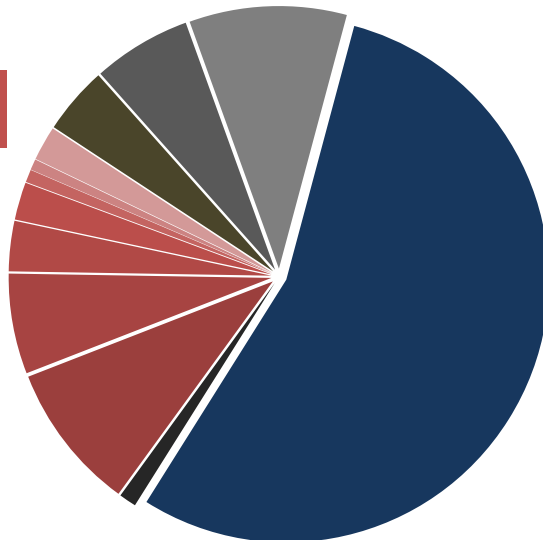
Region 2

13%



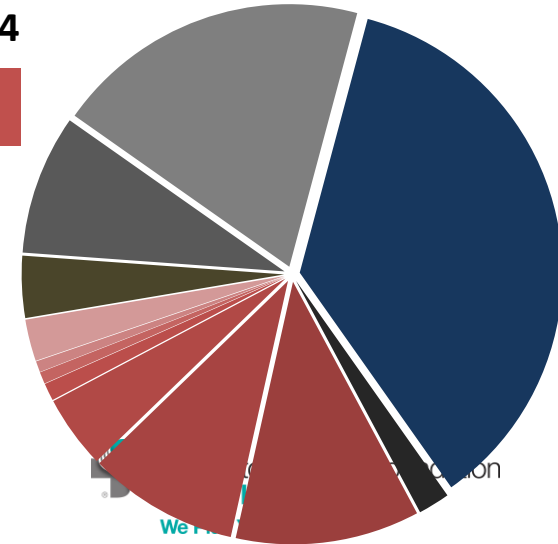
Region 3

24%



Region 4

30%



Does the difference persist after adjusting patient age, sex, and race/ethnicity?

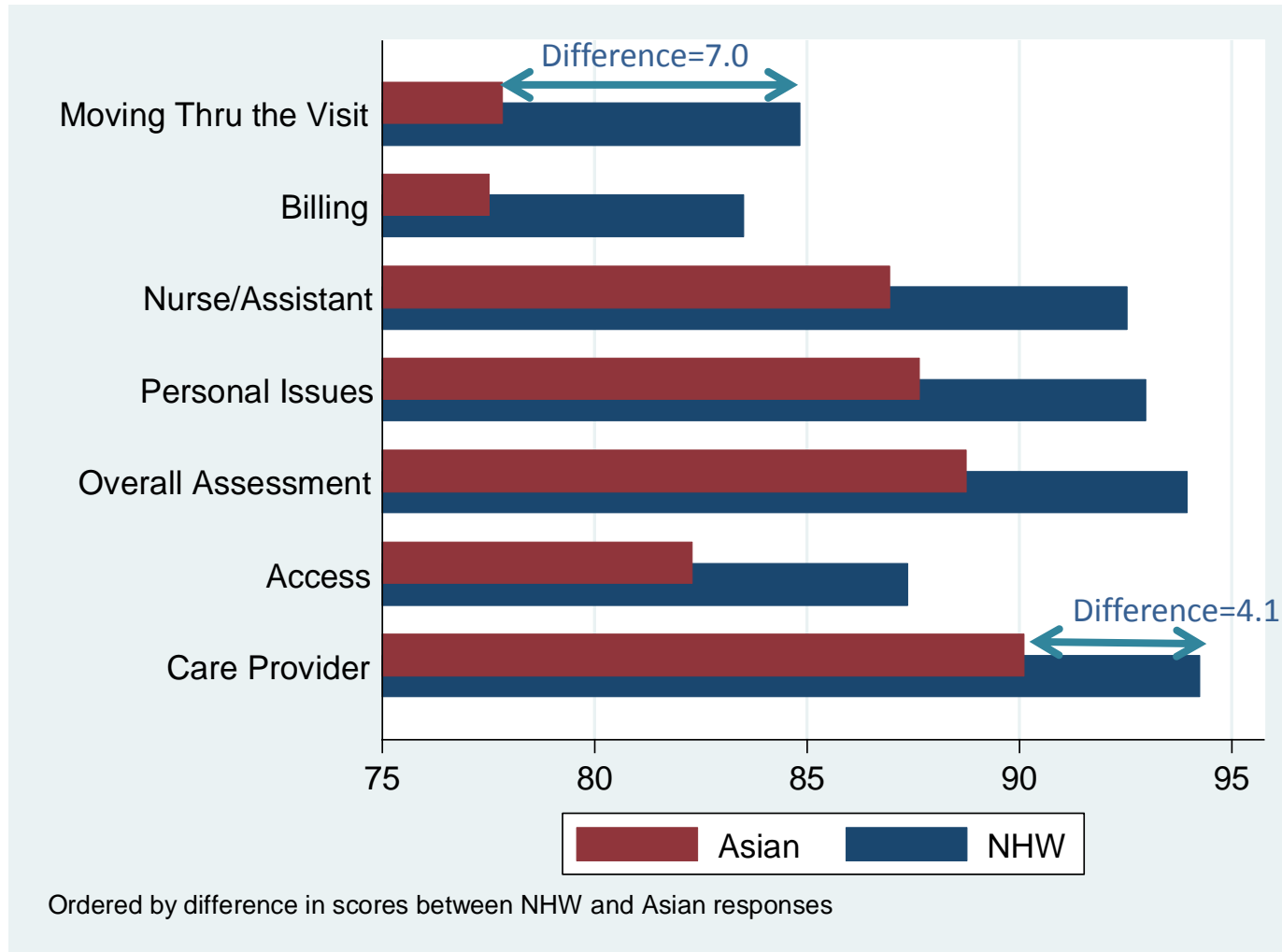
	Without any adjustment	Adjusting patient age/sex and provider specialty	Further adjusting patient race/ethnicity
Region 2 (largest)	89.66	88.21	89.78
<i>Compared to Region 2</i>			
Region 1	0.85	0.53	0.03
Region 3	-0.14	-0.01	0.45
Region 4	-1.32	-0.81	-0.02
R-squared	0.003	0.02	0.04

N=191,924; Statistically significant difference (P<0.01) is in **bold**

Result

Actual vs. Perceived Experience vs. Expectation

Difference across survey domains



Does R/E matter after adjusting age/sex?

Key variables	Access	Moving through the visit	Nurse/assistant	Care provider	Personal issues	Billing	Overall assessment
Asian (ref: NHW)	-4.3	-6.0	-5.0	-3.8	-4.9	-5.1	-4.5
Age (ref: 18-34)							
0-17				1.8	0.7		1.1
45-64	1.8	3.8	1.4	2.3	1.4	1.9	2.3
65+	2.7	4.5	2.4	2.4	1.8	4.2	3.1
Female							
Constant	85.5	81.7	91.1	92.3	91.7	81.0	91.8

All the results presented with coefficient are statistically significant ($p < 0.01$).

Within-provider fixed effects model was used for estimation.

Number of providers=1,097

Number of responses ranges from 37,332 (Billing) to 150,953 (Access)

Racial difference in wait time measures: satisfaction, patient-reported, actual

	Wait time satisfaction: %very good	Patient-reported wait time (min)		EHR-recorded, <u>wait room</u> time (min)		
		Wait room	Exam room	Total	Prior to scheduled t	Post scheduled t
NHW	54.5	7.9	5.9	19.3	7.8	11.6
Asian	38.4	9.0	6.9	17.6	6.8	11.0
Chinese	37.1	8.6	6.7	16.8	6.4	10.6
Asian Indian	36.2	9.0	7.1	17.5	4.8	13.0
Filipino	44.2	9.9	7.9	17.9	9.1	8.7
Japanese	42.8	8.4	6.0	20.1	9.2	11.1
Korean	37.2	8.8	6.9	17.9	7.6	10.5
Vietnamese	37.3	10.3	8.0	17.5	7.0	10.9
Other Asian	37.9	9.4	7.0	17.2	6.9	10.2

Asian and each Asian subgroups was significantly different from NHW ($P < .01$), **after adjusting for patient age, sex, scheduled visit length, time of the day, own provider, and provider fixed effects.** except for those *italicized in gray*. N ranges from 134,586 (satisfaction) to 73,670 (EHR total wait time, EHR post appointment wait time)

Asians rate satisfaction lower, when everything is equal*

Satisfaction of wait time (%very good)

Chinese	-16.1	-15.3	-12.7
Asian Indian	-17.1	-16.8	-12.4
Filipino	-10.4	-10.2	-4.7
Japanese	-10.4	-11.8	-11.3
Korean	-16.3	-14.1	-10.9
Vietnamese	-15.2	-18.0	-11.3
Other Asian	-15.7	-13.2	-10.0
EHR prior-appointment wait time		-0.1	<i>-0.01</i>
EHR post-appointment wait time		-0.1	<i>-0.05</i>
Patient-reported wait room time			<i>-1.4</i>
Patient-reported exam room time			<i>-1.2</i>

* Difference from NHW, after adjusting for patient age, sex, scheduled visit length, time of the day, provider fixed effects, AND EHR recorded and/or patient-reported wait time.

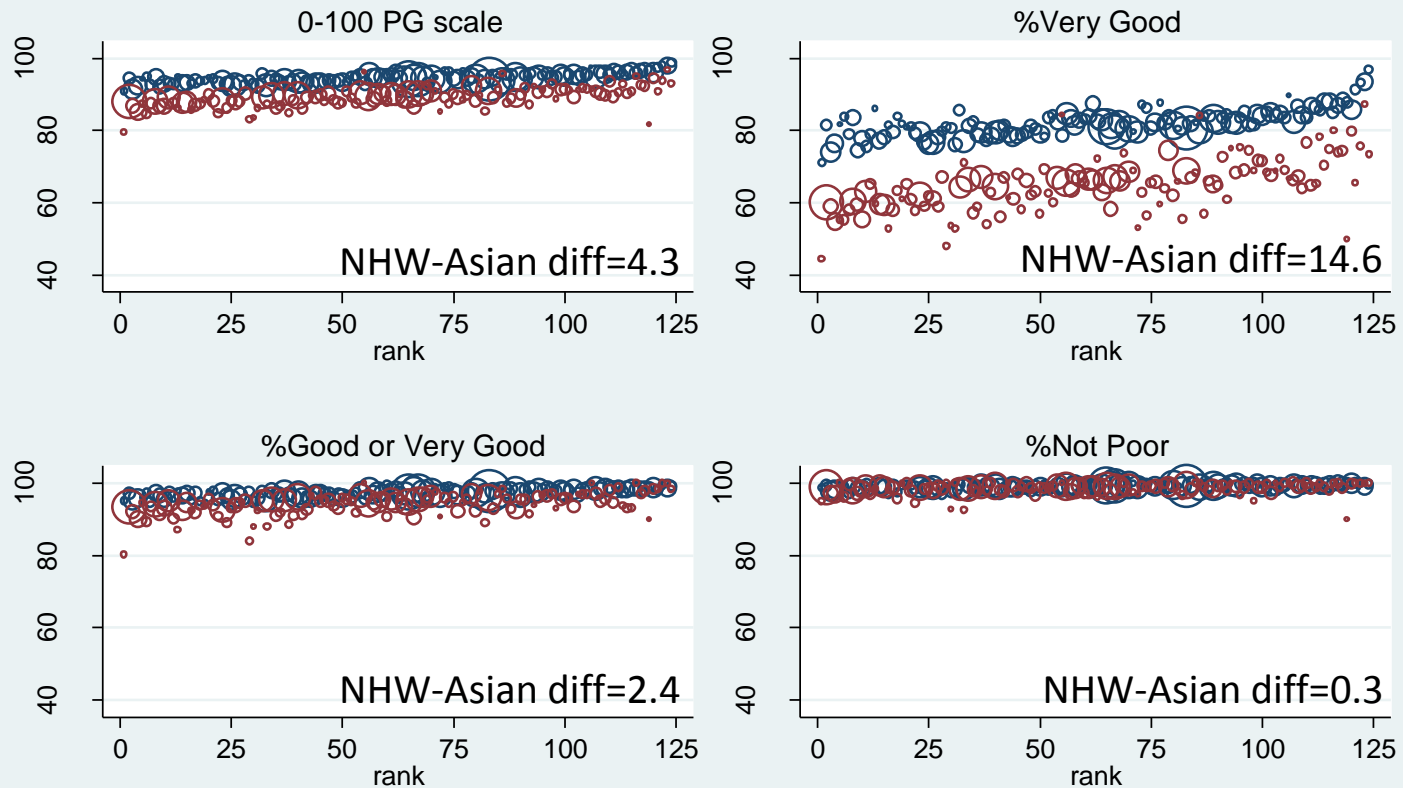
Differences were statistically significant ($P < .01$) except for those *italicized in gray.*

Result

Response effect

Racial gap varies widely with differing scoring methods

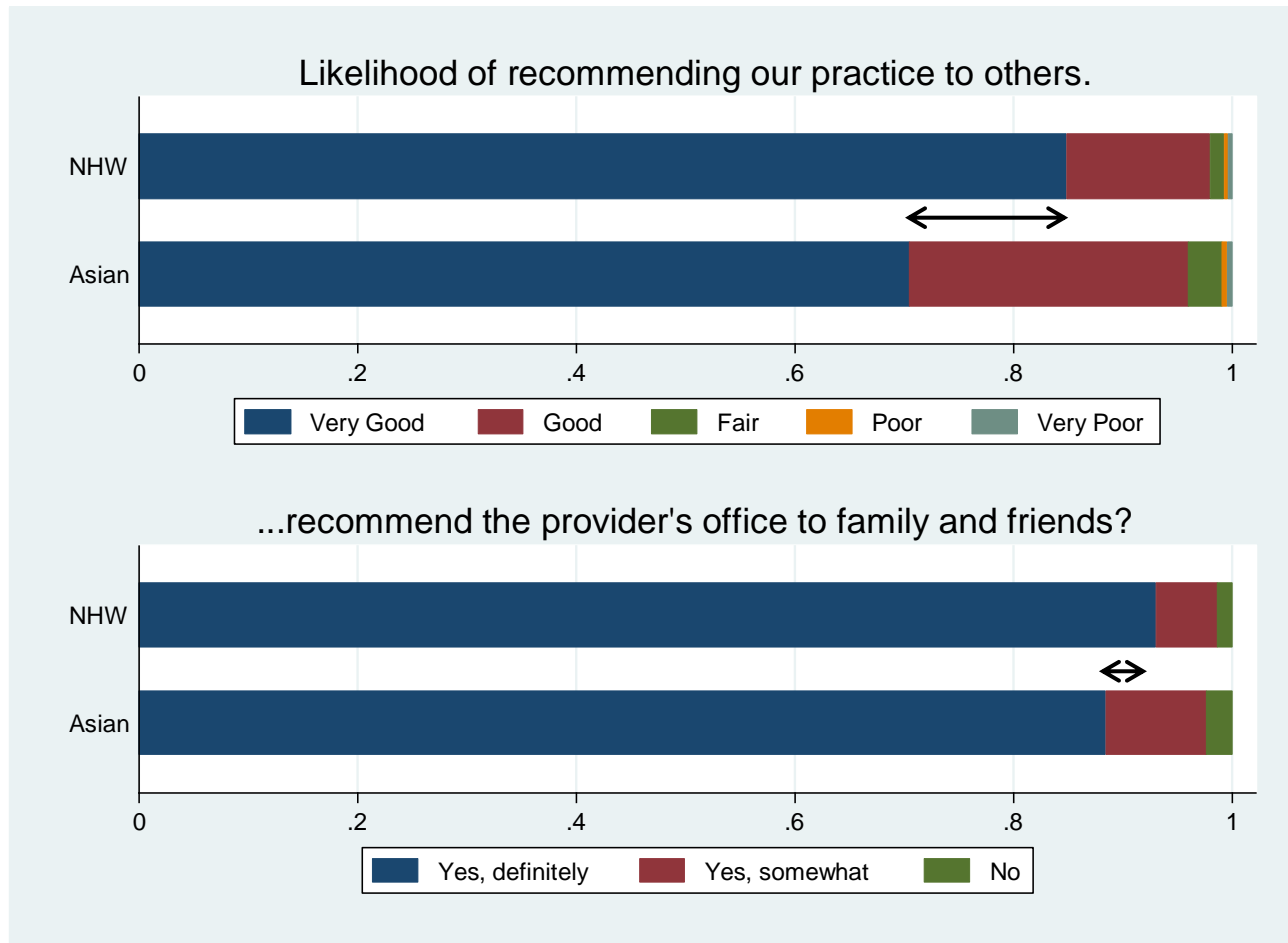
“Likelihood of Recommending the Practice”



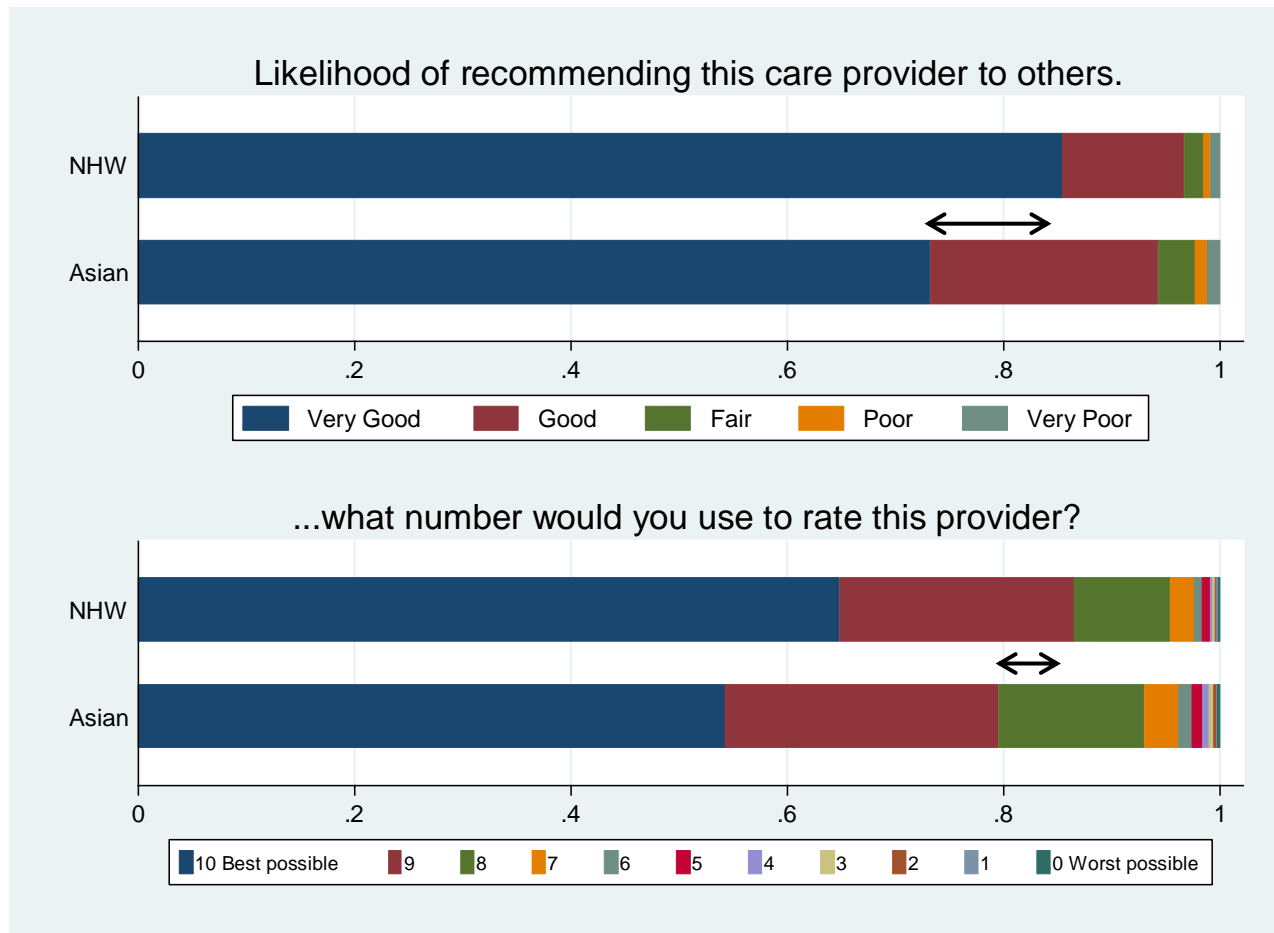
• NHW • Asian

Clinic/departments are ranked based on overall score of all responses regardless of R/E
Circle size is proportional to #survey responses, NHW and Asian respectively, of each clinic

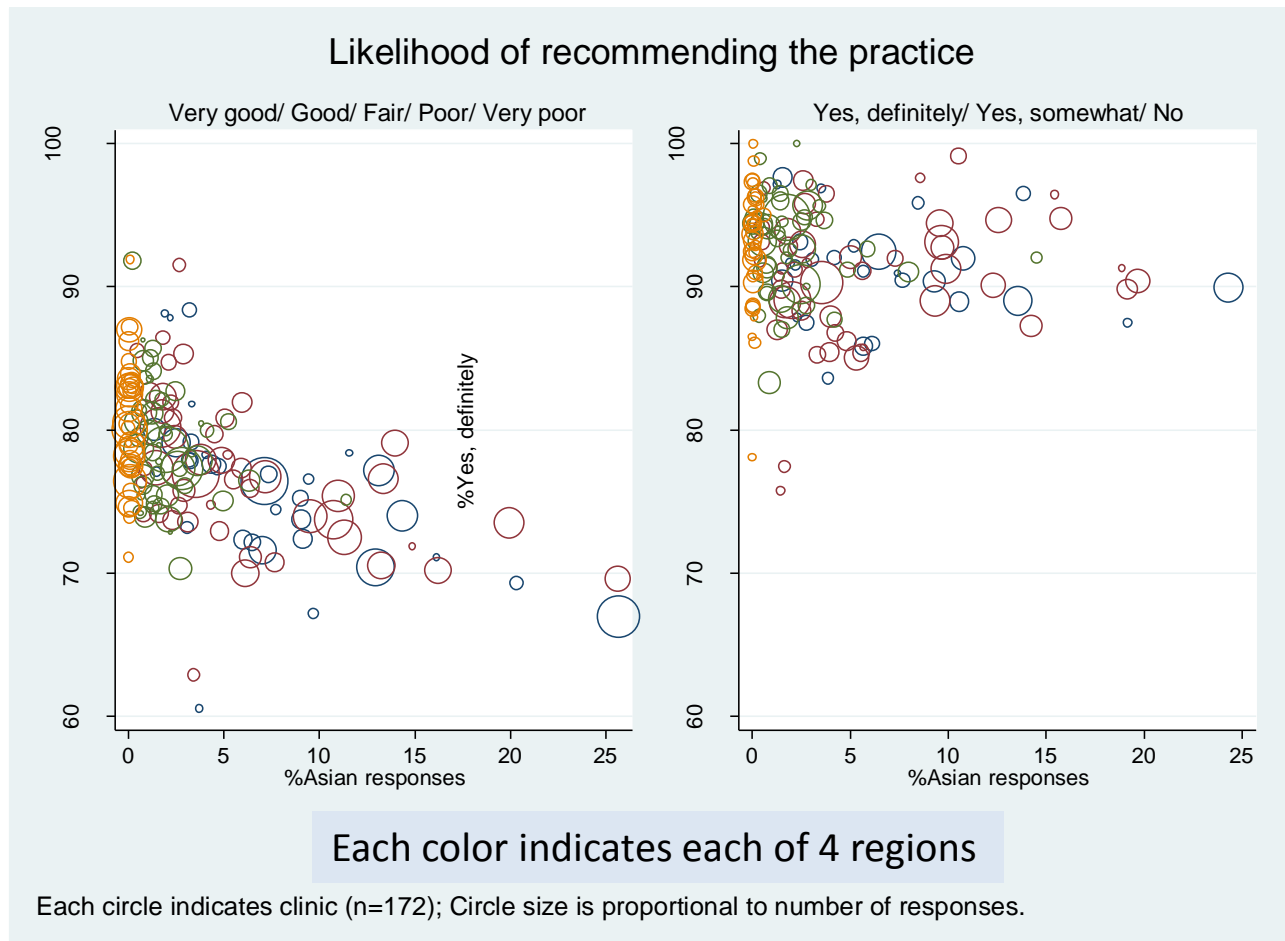
Answer scale matters: racial difference with 5- vs. 3-points scale



Answer scale matters: racial difference with 5- vs. 11-points scale



Provider scores: Asian response effect is mitigated with CG-CAHPS 3-points scale



Summary and conclusions: experience from PAMF

Race is the strongest among all demographic predictors

- The Asian effect is much larger than age/sex effect.
- Within each PAMF region, clinic, and provider, Asians score substantially lower than NHWs.
- PAMF regional difference is in part explained by racial composition. After adjusting for race, regional differences disappear or reverse.

Differential treatment is unlikely to be the driver of the racial difference

- Asian-NHW difference exist across ALL survey questions
- Asian-NHW gap is larger for aspects of care that apply equally to everyone (e.g., Moving through the visit, Billing) and is smallest in more subjective aspect of care
- Wait time satisfaction is much lower among Asians but neither their actual wait time nor their reported wait time explain the difference.

Scoring and answer scale matter

- Answer scale can make a big difference
 - Asians are much less likely to choose
 - 5 (Very Good) in 5-points scale (Very good, Good, Fair, Poor, Very Poor)
 - Asians are NOT much less likely to choose
 - 3 (Yes, definitely) in 3-points scale (Yes, definitely, Yes, somewhat, No)
 - 9 or 10 in 11-points scale (0 worst – 10 best)
- Scoring approach that weighs extreme values less is preferred to mitigate response effect

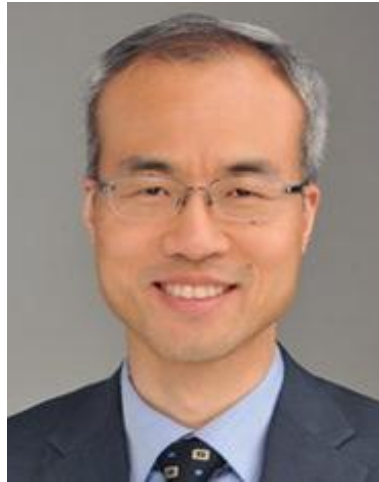
Appropriate scale, scoring or adjustment should be applied for provider assessment

- Provider score and rank change substantially with differing scoring and scales
- Providers with higher proportion of Asian patients fare better when using scales and scoring that are less sensitive to such response effects
- Response effect and heterogeneous patient composition should be considered or appropriately adjusted for when assessing provider performance

Collaborators & contributors



Minji Song,
Master student
CNU



Young Kyung Do,
MD, PhD
SNU



Nicole Johns, MPH
PAMFRI



Hal Luft, Ph.D.
PAMFRI

Fund: AHRQ K01 HS019815



Palo Alto Medical Foundation

Sutter Health

We Plus You

Thank you

Sukyung Chung
chungs@pamfri.org