Comparing Kaiser Permanente Members to the General Population: Implications for Generalizability of Research

Anna C Davis, PhD[1,2]; Jackson L Voelkel, MUS[1]; Carol L Remmers, PhD[1]; John L Adams, PhD[1,2]; Elizabeth A McGlynn, PhD[1,2]

ACD, 0000-0003-3833-472X

Corresponding Author
Elizabeth A McGlynn, PhD
Elizabeth.A.McGlynn@kp.org

Author Affiliations
[1] Kaiser Permanente Research and Quality Measurement, Pasadena, CA, USA

[2] Kaiser Permanente Bernard J Tyson School of Medicine, Pasadena, CA, USA

Abstract

INTRODUCTION: Insight into the characteristics of populations from which research samples are drawn is essential to understanding the generalizability of research findings. This study characterizes the membership of Kaiser Permanente and compares members to the population of the communities in which they live.

METHODS: This study is a descriptive comparison of population distributions for Kaiser Permanente members vs the general population within counties in which Kaiser Permanente operates. Kaiser Permanente data on demographics, membership, geographically linked census data, and chronic condition prevalence were compared with community data drawn from the US Census and the Behavioral Risk Factor Surveillance System.

RESULTS: Overall, Kaiser Permanente members were older (50% aged 40 or older compared to 45.8% of the general population) and more likely to be female (51.8% vs 50.5% of the general population). Distribution by race and ethnicity was similar for all Regions combined but varied somewhat within Regions. Distribution by neighborhood-linked income, education, and social vulnerability was similar between Kaiser Permanente and the community. Prevalence of 6 of 7 chronic conditions was higher in the community than in Kaiser Permanente, with differences ranging from 0.5% for depression to 7.7% for hyperlipidemia.

CONCLUSION: The demographic characteristics of Kaiser Permanente members are similar to the general population within each of the Kaiser Permanente Regions. Overall, the size and diversity of the Kaiser Permanente membership offers an effective platform for research. This approach to comparing health system members with the larger community provides valuable context for interpreting real-world evidence, including understanding the generalizability of research and of measures of system performance.

Introduction

Assessing the generalizability of research results and interpreting health system performance measures based on real-world data requires insight into the context within which the analysis is done. Observational studies contribute to understanding many dimensions of health and health services (eg, natural course of illness, effectiveness and comparative effectiveness of treatments, diffusion of new technologies, real-world experiences of people seeking health care services).[1-3] Compared to randomized clinical trials, observational studies have inherent biases and unmeasured confounders, but the insights obtained from these real-world studies offer important inputs to advancing health care policy and practice that cannot be obtained from randomized clinical trials,[4,5] which are seldom representative of their source populations.

Health care systems, especially managed care organizations, can leverage the substantial data generated in the usual course of health services delivery and extended longitudinal member relationships to address a variety of questions.[3,5,6] Managed care systems also have the advantage of a known population of enrollees from which results can be generalized. Whether observational studies are being conducted in a single system or a consortium of systems, it is useful to understand whether the population in those systems is like those living in the communities, states, and countries from which the observational study population is drawn. The Strengthening the Reporting of Observational Studies in Epidemiology criteria require describing the source population, including how this changed at various stages of eligibility or inclusion in a study as well as the generalizability of the results.[7] This contextual information is useful to potential research funders, journal editors and reviewers, and end users of the research.

Similarly, when health systems examine their performance on quality, service, or financial dimensions, they should understand how the populations they serve differ from other systems or geographies to which they may compare their results. Performance measures may be risk adjusted to account for variation due to differences in populations served, or they may be stratified by population characteristics of interest, such as race and ethnicity, geography, or health status. Both risk adjustment and stratification require the ability to adequately characterize the population on whom performance is reported.

This paper demonstrates an approach to setting the context for real-world studies and measures of health system performance. Using data from the 8 geographically dispersed Regions that constitute Kaiser Permanente, this study characterizes the membership of Kaiser Permanente on sociodemographic characteristics and prevalence of comorbidities and compares Kaiser Permanente members to the general population within the communities in which Kaiser Permanente operates.

Methods

SETTING
Kaiser Permanente is an integrated delivery system with 12.6 million enrollees in 8 Regions (Northern California, Southern California, Colorado, Georgia, Washington, Northwest [Oregon and parts of southwest Washington], Hawaii, and Mid-Atlantic States [District of Columbia and parts of Maryland and Virginia]). Kaiser Permanente comprises Kaiser Foundation Health Plan, Inc. and its subsidiaries; the not-for-profit Kaiser Foundation Hospitals, which operates 39 hospitals and over 720 other clinical facilities; and the Permanente Medical Groups, self-governed physician group practices that exclusively contract with the Kaiser Foundation Health Plan and its subsidiaries to meet the health needs of Kaiser Permanente's members.[8] Appendix A contains geospatial maps of the Kaiser Permanente Regions.

DESIGN, DATA SOURCES, AND STUDY POPULATIONS
Data on Kaiser Permanente membership were drawn from 3 sources: Kaiser Permanente's Member Month Mart, which contains monthly membership data for all Kaiser Permanente Regions; Kaiser Permanente's Geographically Enriched Member Sociodemographics datamart, which contains demographic and geographically linked census data based on member residential addresses for both current and historical timepoints; and Kaiser Permanente's Integrated Data Repository, which combines medical record data and claims from services rendered by outside medical practitioners to offer a complete view of member utilization and health conditions.

To align with comparator estimates for the general population, Kaiser Permanente members were included in all analyses if they were enrolled in December 2020. Demographic analyses included people of all ages, while chronic condition analyses were limited to those age 18 or older as of December 31, 2020.

Community demographic comparisons were constructed from the 2016–2020 5-year American Community Survey (ACS) estimates. The tidycensus package in R was used to extract the specific variables of interest from the US Census Bureau's Application Programming Interface.[9]

Chronic condition comparisons were based on state-level estimates from the Behavioral Risk Factor Surveillance System (BRFSS).[10] The 2020 BRFSS was used for all conditions except for hypertension and hyperlipidemia, which were most recently available from the 2019 BRFSS. The BRFSS contains self-reported data on condition history based on validated survey items; the methods for BRFSS have been reported elsewhere.[11,12]

Data were summarized by Kaiser Permanente Region and overall. For demographic analyses, the comparator population for each Kaiser Permanente Region was defined using the counties that intersected the Kaiser Permanente Region areas, and "Kaiser Permanente total" represented an aggregation of the regional county-based footprints. For chronic condition analyses, comparator populations were based on the whole US state(s) in which the Kaiser Permanente Region was located, and "Kaiser Permanente total" was a simple average of the state-level BRFSS estimates for the states in which Kaiser Permanente operates. For Kaiser Permanente Regions spanning more than 1 state, the BRFSS results were an average of the individual state rates; for Kaiser Permanente Regions located in a single state, the corresponding statewide BRFSS estimates were used. In the 2 Regions located in California, the overall state BRFSS estimate was used for both. Finally, both the demographic and chronic condition analyses provided US overall general population estimates for context.

Additional detail about the data sources and definitions for each of the variables included in the descriptive analysis can be found in Appendix A, Tables A2 and A3.

STUDY VARIABLES

Data on race and ethnicity were reported during enrollment or when interacting with the health system (86% of members) or were estimated (14%) using the Bayesian Improved Surname Geocoding (BISG) method.[13–15] To enable separation of Asian and Pacific Islander populations within Kaiser Permanente data, an extension of the BISG method that resulted in 7 race and ethnicity groups was implemented for this study; a description of this

BISG extension is provided in Appendix B. Members who reported Hispanic ethnicity were categorized as Hispanic for all races. ACS 2016–2020 5-year estimates were used for the comparator population. Some ACS categories were aggregated to match the available categories for the Kaiser Permanente membership.

Geographically linked census information on educational attainment, income, and social vulnerability index (SVI) was joined to Kaiser Permanente members based on the US Census Block Group or Tract in which they resided. The national-scale SVI was binned into quintiles.[16] These characteristics of neighborhoods were summarized for both Kaiser Permanente members and the comparator population. Educational attainment was summarized among people aged 25 and older to align with the way it is reported by the US Census; all other geographically linked data were summarized for all ages.

All Kaiser Permanente members were insured through employer-based coverage, Medicare, Medicaid, or individual insurance plans. General population uninsurance rates in the community were calculated for comparison.

The Centers for Medicare and Medicaid Services' Chronic Condition Warehouse method[17] was used to estimate chronic condition prevalence for Kaiser Permanente members based on International Classification of Diseases 10th Revision (ICD-10-CM) codes for chronic obstructive pulmonary disease, coronary artery disease, depression, diabetes, hyperlipidemia, and hypertension. A 5-year look-back period was used to capture all ICD-10-CM codes for each member from 2016 to 2020. In addition, the most recent height and weight data available from the medical record for each member was combined with ICD-10-CM codes during the 2-year period from 2019 to 2020 to identify prevalence of overweight or obesity. A Body Mass Index of 25 or greater was used to categorize all adults as overweight or obese to align with the state-level BRFSS estimates. Condition prevalence estimates for adult Kaiser Permanente members were compared to self-reported condition information for the general adult population based on the BRFSS.

STATISTICAL ANALYSIS

The analysis compares population distributions of demographic characteristics and prevalence of health conditions. Measures of statistical significance were not calculated due to very large sample

sizes and because the Kaiser Permanente population was nested within the comparator population.

SENSITIVITY ANALYSES
As a sensitivity analysis to assess whether the demographic comparisons would vary based on the method of defining the comparator geography, 2 additional versions of the general population comparator were constructed using 1) census tracts that intersect the Kaiser Permanente footprint in each Region; and 2) the entire state for each Kaiser Permanente Region. These are presented alongside the primary analysis defined at the county level (see Appendix C, Tables C1–C9).

A second sensitivity analysis was related to the chronic condition comparisons and examined the method of aggregating BRFSS data when Kaiser Permanente Regions spanned multiple states. In Appendix C, Table C10, the chronic condition results are presented with weighted (rather than simple) averages of the individual state BRFSS rates for the 3 geographies that spanned multiple states: Mid-Atlantic States, Northwest, and Kaiser Permanente overall. Weighting was based on the proportion of the total Kaiser Permanente membership located in each of the contributing states.

This study was determined to be exempt from Institutional Review Board review because it was not human subjects research; the data used in the analysis were deidentified, and the analysts had access to the data as part of their standard work.

Results

POPULATION SIZE
As of December 2020, the Kaiser Permanente California Regions together made up about 9 million of Kaiser Permanente's then 12.2 million members. Other Regions ranged in size from 257,000 members in Kaiser Permanente Hawaii to 768,000 in Kaiser Permanente Mid-Atlantic States (Table 1). Market share (the proportion of the general population covered by Kaiser Permanente) within the counties where Kaiser Permanente operates was 18% on average, ranging from 32.5% in Kaiser Permanente Northern California to 4.7% in Kaiser Permanente Georgia (Appendix A, Table A1).

DEMOGRAPHICS OF KAISER PERMANENTE OVERALL
A slightly larger proportion of the total Kaiser Permanente membership was female compared to the general population (51.8% vs 50.5%, Table 1). Fifty percent of Kaiser Permanente members were aged 40 or older, compared to 45.8% of the general population. The distribution of race and ethnicity among Kaiser Permanente members was like the general population of the areas Kaiser Permanente serves; Kaiser Permanente had a larger proportion of members who were non-Hispanic Asian (15.3% vs 12.2%) and a smaller proportion who were non-Hispanic White (43.3% vs 45.0%). Educational attainment and income for the neighborhoods in which Kaiser Permanente members live was comparable to the general population, although a smaller proportion of Kaiser Permanente members lived in communities that were classified as being the most (20.1% vs 22.7%) and least (17.0% vs 21.0%) vulnerable on the SVI. No Kaiser Permanente members were uninsured compared to an average uninsured rate of 7.4% in the Kaiser Permanente service areas. These findings were similar across the different methods of constructing the comparison geography (Appendix C, Table C9). A table presenting calculated ordinary differences between Kaiser Permanente and the general population in each Region is presented in Appendix C, Table C11, to assist readers with quickly assessing magnitudes of differences.

Comparing Kaiser Permanente members overall to the whole US population revealed patterns that underscore the utility of more focused geographic comparisons. A higher proportion of Kaiser Permanente members were female compared to the overall US population (51.8% vs 50.8%, Table 1). A higher proportion of Kaiser Permanente members were Hispanic or Latino (27.8% vs 18.2%) or Asian (15.3% vs 5.6%) compared to the US population. A lower proportion of Kaiser Permanente members were Black or African American (10.0% vs 12.2%) or White (43.3% vs 60.1%). The probability that Kaiser Permanente members had less than a high school education was higher than in the US (12.5% vs 11.5%); the probability that Kaiser Permanente members had some college education, a bachelor's degree, or a graduate or professional degree was higher than in the US overall. Kaiser Permanente members had a lower probability of having household incomes less than $75,000 and a higher probability of having household incomes $75,000 or greater than the US overall. A higher proportion of Kaiser Permanente members lived in communities that were the most socially vulnerable (20.1% vs 18.8% in the US overall), and a lower proportion of Kaiser Permanente members lived in communities that were the least vulnerable (17.0% vs 20.6%).

Variable (%)	Colorado		Georgia		Hawaii		Mid-Atlantic States		Northern California		Northwest		Southern California		Washington		Kaiser Permanente total		US total
	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	General population
N (millions)	0.6	4.9	0.3	6.3	0.3	1.4	0.8	9.1	4.4	13.6	0.6	3.7	4.6	22.9	0.7	6.4	12.2	68.4	329.5
Sex (%)																			
Female	52.4	49.8	54.6	51.6	50.4	49.8	53.0	51.3	51.3	50.3	51.9	50.5	51.6	50.4	53.4	49.9	51.8	50.5	50.8
Male	47.6	50.2	45.4	48.4	49.6	50.2	47.0	48.7	48.7	49.7	48.1	49.5	48.4	49.6	46.6	50.1	48.2	49.5	49.2
Race and ethnicity (%)																			
Hispanic or Latino	15.9	21.9	5.3	11.3	8.2	10.7	14.6	12.7	23.7	29.7	9.9	13.1	43.4	44.9	7.3	12.6	27.8	27.3	18.2
Non-Hispanic American Indian or Alaskan Native	0.4	0.4	0.2	0.1	0.2	0.1	0.2	0.2	0.4	0.3	0.6	0.6	0.3	0.3	0.8	0.9	0.4	0.3	0.6
Non-Hispanic Asian	4.1	3.5	6.1	5.9	32.4	36.8	13.7	8.8	22.1	19.8	6.9	5.3	12.1	12.8	9.0	9.9	15.3	12.2	5.6
Non-Hispanic Black	4.4	4.4	50.0	32.4	0.9	1.8	39.7	26.0	6.8	5.4	3.6	2.1	8.4	5.9	4.2	4.2	10.0	10.4	12.2
Non-Hispanic Pacific Islander	0.2	0.1	0.1	0.0	11.5	9.7	0.1	0.0	0.8	0.5	0.7	0.5	0.4	0.3	0.7	0.7	0.8	0.5	0.2
Non-Hispanic Multiracial	1.8	3.0	1.4	2.6	22.4	19.2	2.2	3.4	2.2	4.1	1.6	4.4	1.8	2.9	3.2	5.4	2.5	3.9	2.8
Non-Hispanic White	73.2	66.3	37.0	47.2	24.2	21.6	29.5	48.3	44.0	39.8	76.7	73.7	33.5	32.7	74.6	66.0	43.3	45.0	60.1
Age, y (%)																			
0–4	3.6	5.9	3.4	6.2	4.6	6.2	5.5	6.3	4.8	6.0	4.5	5.6	5.0	6.2	3.2	6.1	4.7	6.1	6.0
5–9	4.3	6.2	4.1	6.6	5.0	5.9	5.1	6.2	5.3	6.2	5.0	5.9	5.5	6.1	3.7	6.1	5.2	6.2	6.1
10–14	5.1	6.5	5.3	7.1	5.6	5.9	5.3	6.4	5.7	6.5	5.4	6.1	6.0	6.6	4.3	6.1	5.7	6.5	6.5
15–19	5.4	6.4	6.5	7.0	5.6	5.4	5.5	6.2	5.8	6.2	5.6	6.1	6.2	6.5	5.1	5.9	5.9	6.4	6.5
20–24	5.4	6.8	7.9	6.7	6.4	6.5	6.5	6.2	6.2	6.3	6.2	6.6	6.8	7.0	6.4	6.6	6.5	6.7	6.7
25–29	6.0	8.1	6.3	7.3	6.8	7.3	7.1	7.3	6.9	7.7	6.7	7.4	7.3	8.1	7.0	8.0	7.0	7.8	7.1
30–34	6.8	8.0	6.6	6.9	7.2	7.0	7.6	7.5	7.8	7.6	7.4	7.3	7.7	7.5	7.5	7.8	7.6	7.5	6.8
35–39	6.9	7.3	6.8	7.1	7.3	6.6	7.3	7.3	7.7	7.1	7.4	7.2	7.3	6.9	7.3	7.2	7.4	7.1	6.5
40–44	6.8	6.7	7.3	6.7	6.9	6.1	6.8	6.5	7.0	6.5	7.2	6.6	6.8	6.4	6.9	6.4	6.9	6.5	6.1

Table 1: Demographic characteristics of Kaiser Permanente members vs the general population in the counties in which Kaiser Permanente operates (*Continued*)

Table 1: *Continued*

Variable (%)	Colorado		Georgia		Hawaii		Mid-Atlantic States		Northern California		Northwest		Southern California		Washington		Kaiser Permanente total		US total
	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	General population
45–49	6.7	6.5	8.2	7.2	6.6	6.0	6.6	6.7	6.7	6.5	6.8	6.5	6.5	6.5	6.7	6.3	6.6	6.6	**6.3**
50–54	6.9	6.1	8.6	6.8	6.8	6.1	7.1	6.9	6.8	6.4	6.7	6.1	6.7	6.4	7.3	6.2	6.8	6.4	**6.4**
55–59	7.2	6.4	8.5	6.4	7.2	6.4	7.3	6.7	6.9	6.4	6.7	6.2	6.6	6.3	7.9	6.3	6.9	6.4	**6.7**
60–64	7.4	5.7	7.3	5.5	6.9	6.4	6.8	5.9	6.5	5.8	6.7	6.3	6.2	5.6	8.5	6.2	6.6	5.8	**6.2**
65–69	6.9	4.8	5.2	4.5	5.6	6.0	5.5	4.7	5.3	4.9	5.9	5.7	5.2	4.5	6.3	5.1	5.4	4.8	**5.3**
70–74	6.3	3.6	3.9	3.3	4.7	4.4	4.5	3.5	4.3	3.7	5.2	4.1	4.3	3.4	5.2	3.9	4.5	3.6	**4.1**
75–79	3.9	2.2	2.1	2.1	3.0	3.0	2.8	2.4	2.8	2.5	3.2	2.6	2.7	2.3	3.2	2.4	2.8	2.4	**2.8**
80–84	2.3	1.4	1.1	1.3	1.8	2.1	1.5	1.6	1.8	1.7	1.8	1.7	1.7	1.6	1.8	1.6	1.7	1.6	**1.9**
85+	2.1	1.6	0.9	1.2	2.0	2.8	1.2	1.7	1.7	2.0	1.5	1.9	1.5	1.9	1.8	1.8	1.6	1.8	**2.0**
Educational attainment (%)																			
Less than high school	7.2	7.8	9.0	10.2	7.9	7.5	9.3	8.7	11.2	13.3	8.5	8.4	16.9	17.7	6.7	8.0	12.5	12.5	**11.5**
High school graduate	19.2	20.4	24.1	23.8	26.6	27.3	20.0	20.0	18.9	18.8	20.3	21.6	22.1	21.0	19.6	21.0	20.6	20.8	**26.7**
Some college or associate's degree	28.5	28.9	28.0	26.9	31.2	31.6	23.6	23.1	29.0	28.3	33.6	33.8	29.5	28.5	31.9	32.6	29.2	28.4	**28.9**
Bachelor's degree	27.7	26.6	24.0	24.1	21.2	21.9	24.7	25.0	24.6	23.6	23.2	22.5	20.2	21.0	24.8	23.8	23.0	23.1	**20.2**
Graduate or professional degree	17.1	16.2	14.6	15.0	10.9	11.6	22.1	23.2	15.9	16.0	13.8	13.7	11.1	11.8	15.6	14.7	14.3	15.2	**12.7**
Household income in dollars (%)																			
0–24,999	10.7	13.4	13.8	15.7	11.7	13.0	10.6	11.1	11.8	13.9	12.7	15.5	14.7	16.1	11.3	13.5	12.8	14.3	**18.4**
25,000–34,999	5.9	6.9	7.4	8.1	6.2	6.3	4.9	5.2	5.3	6.3	7.3	7.9	7.0	7.4	5.9	6.6	6.1	6.8	**8.6**
35,000–74,999	26.2	27.8	30.1	29.3	24.4	25.8	22.8	21.9	22.3	22.8	28.6	29.3	27.2	26.4	26.5	27.2	25.1	25.7	**29.2**
75,000–99,999	14.1	13.6	13.9	13.2	13.8	13.7	13.2	12.5	12.3	11.7	14.1	13.6	13.6	12.7	13.9	13.5	13.2	12.7	**12.8**
100,000+	42.8	38.3	34.5	33.8	41.6	41.2	48.3	49.4	47.9	45.3	36.7	33.7	37.2	37.5	41.0	39.3	42.2	40.4	**31.0**
Social vulnerability index (%)																			

Table 1: *Continued*

Variable (%)	Colorado		Georgia		Hawaii		Mid-Atlantic States		Northern California		Northwest		Southern California		Washington		Kaiser Permanente total		US total
	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	General population
SVI 5 (most vulnerable)	7.5	12.2	14.5	16.1	15.5	13.8	8.8	8.4	16.7	23.4	14.1	16.0	31.9	35.3	9.8	13.4	20.1	22.7	18.8
SVI 4	11.6	13.9	23.4	20.9	21.1	19.5	18.2	14.6	19.0	18.4	18.7	22.4	23.8	21.3	16.5	17.6	20.9	18.8	19.5
SVI 3	11.9	13.2	15.1	16.1	30.1	27.8	20.0	17.3	23.3	20.5	23.6	22.3	18.4	15.9	22.6	21.8	21.1	17.9	20.1
SVI 2	21.0	20.4	20.3	19.4	18.5	21.2	22.5	21.8	24.1	21.4	25.7	23.6	14.5	14.4	27.3	25.1	21.0	19.3	20.9
SVI 1 (least vulnerable)	47.9	39.8	26.7	27.3	14.7	17.6	30.6	37.7	16.9	16.1	17.8	15.8	11.3	12.7	23.9	22.1	17.0	21.0	20.6
Health insurance (%)																			
Uninsured	0.0	7.3	0.0	12.7	0.0	3.9	0.0	6.6	0.0	5.1	0.0	6.3	0.0	8.4	0.0	6.0	0.0	7.4	8.7

Notes: Data for Kaiser Permanente members represent all members of the Kaiser Permanente Region as of December 2020. Data for the general population in each Region are an aggregation of people living in the counties in which Kaiser Permanente operates and are based on 2016–2020 5-year American Community Survey estimates. For Kaiser Permanente members, age, race and ethnicity, and sex are based on Health Plan data. For both Kaiser Permanente members and the general population, education, income, and social vulnerability index (SVI) are based on census-linked data and as such describe the characteristics of the neighborhoods in which people live. The distribution of educational attainment is reported among people aged 25 or older for both Kaiser Permanente members and the general population; all other data are reported for the total population.

SVI = social vulnerability index.

DEMOGRAPHICS OF THE INDIVIDUAL KAISER PERMANENTE REGIONS COMPARED TO THEIR LOCAL AREAS

While at a national level, demographic differences were very small between Kaiser Permanente members and the general population of the Kaiser Permanente service areas, in some Kaiser Permanente Regions there were larger differences in select demographic characteristics.

A higher proportion of Kaiser Permanente members were female than the general population in all Regions, but this difference was greater in Kaiser Permanente Colorado, Kaiser Permanente Georgia, and Kaiser Permanente Washington, where the percent female among Kaiser Permanente members was between 2.6 and 3.5 percentage points higher than in their respective general populations (Appendix C, Table C11). All Kaiser Permanente Regions skewed older than the general population, except Kaiser Permanente Hawaii, where the proportion of members who were age 65 and older was lower (17.1% vs 18.4% [Table 1]).

Greater variation was found in the distribution by race and ethnicity. In Kaiser Permanente Georgia, 50.0% of Kaiser Permanente members were Black compared to 32.4% of the general population (Table 1). In Kaiser Permanente Mid-Atlantic States, 39.7% of Kaiser Permanente members were Black compared to 26.0% of the general population. The Kaiser Permanente member populations in Kaiser Permanente Colorado and Kaiser Permanente Washington had the greatest differential in proportion of people who were White compared to the general populations of those areas, with differences of 6.8 percentage points and 8.6 percentage points, respectively (Appendix C, Table C11).

Differences in neighborhood-linked estimates of educational attainment were very small across all Regions. Income distribution estimated from underlying population data based on member location was also similar between the regional Kaiser Permanente member populations and their general population comparators, although in 6 of the 8 Regions, a greater proportion of the Kaiser Permanente population was estimated to be in the highest income level (differences ranging from 0.4 percentage points in Kaiser Permanente Hawaii to 4.5 percentage points in Kaiser Permanente Colorado, Appendix C, Table C11).

In Kaiser Permanente Colorado, Kaiser Permanente Northern California, Kaiser Permanente Northwest, and Kaiser Permanente Washington, a larger share

of the Kaiser Permanente member population lived in neighborhoods in the least vulnerable category of the SVI compared to their respective general populations (differences ranging from 0.8 to 8.2 percentage points [Appendix C, Table C11]). In the other Regions a smaller proportion of the Kaiser Permanente member population lived in these least vulnerable communities relative to their comparator populations (differences ranging from −0.6 to −7.1 percentage points). The proportion of the population living in neighborhoods in the most vulnerable category of SVI was greater among Kaiser Permanente members than the general population in Kaiser Permanente Hawaii (15.5% vs 13.8%) and Kaiser Permanente Mid-Atlantic States (8.8% vs 8.4% [Table 1]).

The demographic comparisons were similar across the different methods of constructing the geographic comparators. The biggest differences were in comparisons with the full state population, particularly for Kaiser Permanente Georgia and Kaiser Permanente Mid-Atlantic States (Appendix C, Tables C2 and C4). For example, 24.5% of the Georgia state population lived in the most vulnerable quintile of SVI compared to 16.1% using either the census tract or county-based comparator. For Mid-Atlantic States, 40.6% of the combined state populations had incomes greater than $100,000 compared to 49.4% of the county-based comparator.

CHRONIC CONDITIONS

The prevalence of 6 of the 7 conditions was higher in the general population than among Kaiser Permanente members (coronary artery disease, chronic obstructive pulmonary disease, depression, hyperlipidemia, hypertension, and overweight), while prevalence was higher among Kaiser Permanente members for diabetes (Table 2). This overall pattern was generally consistent across the Kaiser Permanente Regions. For Kaiser Permanente overall, the differences in condition prevalence between members and the general population ranged from 0.5 percentage points for depression to 7.7 percentage points for hyperlipidemia. Overweight or obesity was the most prevalent condition in Kaiser Permanente overall and in the general population of the states in which Kaiser Permanente operates (60.9% and 63.4%). Depression rates among Kaiser Permanente members were highest in Kaiser Permanente Northwest (23.9%), Kaiser Permanente Washington (21.2%), and Kaiser Permanente Colorado (19.4%); the state-level comparator populations demonstrate a similar pattern. A table presenting calculated ordinary differences between Kaiser Permanente and the general population in

Condition (%)	Colorado		Georgia		Hawaii		Mid-Atlantic states		Northern California		Northwest		Southern California		Washington		Kaiser Permanente total		US total
	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	Kaiser Permanente members	General population	General population
Chronic obstructive pulmonary disease	4.8	4.5	2.6	6.6	6.0	3.7	2.2	5.0	2.7	5.4	4.1	5.7	4.0	5.4	4.7	5.3	3.5	5.2	6.2
Coronary artery disease	2.2	2.9	3.0	4.6	2.6	2.9	1.9	3.3	2.1	3.1	2.6	3.5	1.6	3.1	3.8	3.7	2.1	3.4	4.0
Depression	19.4	18.4	15.0	17.2	10.4	12.7	12.1	17.6	15.9	14.1	23.9	22.3	16.9	14.1	21.2	23.4	16.8	17.3	19.6
Diabetes	9.6	7.6	12.2	11.6	12.2	10.9	10.9	9.6	10.8	9.8	10.1	9.1	12.5	9.8	8.9	8.6	11.3	9.7	10.6
Hyperlipidemia	27.4	29.9	24.0	32.4	27.0	30.0	18.9	31.9	21.7	29.9	17.5	31.1	26.5	29.2	20.5	31.9	23.4	31.1	33.1
Hypertension	22.6	25.8	29.2	34.8	23.4	30.8	22.3	31.7	23.2	27.8	23.1	30.5	23.1	27.8	23.8	30.3	23.2	30.2	32.3
Overweight (BMI 25+)	57.0	59.2	67.5	67.2	54.2	58.1	64.4	63.7	57.7	64.0	60.7	64.1	64.4	64.0	62.2	64.0	60.9	63.4	67.1

Table 2: Prevalence of select chronic conditions among adults for Kaiser Permanente members vs the general population in the states in which Kaiser Permanente operates

Notes: All data are reported for adults age 18+. Kaiser Permanente member estimates were based on clinical data from the electronic health record (diagnoses; height and weight) for all adult members within each Kaiser Permanente Region as of December 2020. General population estimates were based on state-level statistics from the Behavioral Risk Factor Surveillance System (BRFSS) and reflect self-reported condition status as of 2019 (hypertension and hyperlipidemia) or 2020 (all other conditions). For each Region, the general population comparator was based on BRFSS data for the whole state(s) in which that Kaiser Permanente Region is located. For Kaiser Permanente Regions in a single state, that state's estimates were used; for Regions spanning more than one state, the BRFSS results were a simple average of the individual state rates (the supplemental appendix offers population-weighted averages of BRFSS results for these Kaiser Permanente Regions as a sensitivity analysis). In the 2 Regions located in California, the overall state estimates were used for both. The "Kaiser Permanente total" columns report data for all Kaiser Permanente members and BRFSS results averaged across all states in which Kaiser Permanente operates. Finally, the "US total" column reports BRFSS estimates for the US (50 states and the District of Columbia).

BMI = body mass index.

each Region is presented in Appendix C, Table C12 to assist readers with quickly assessing magnitudes of differences.

Discussion

This paper demonstrates an approach to comparing demographic characteristics and 7 common chronic conditions of Health Plan enrollees to various geographic units of analysis for the purpose of setting context for observational research and for understanding data on performance of health systems. For example, a comparison of population characteristics such as this one was useful to Kaiser Permanente leaders seeking to understand and contextualize data on COVID-19 vaccination rates for Kaiser Permanente members compared to vaccine uptake information published by state health authorities for the general population.

This analysis found that Kaiser Permanente members were comparable to those living in the communities from which the membership was drawn, both overall and within each of the geographic Regions in which Kaiser Permanente operates. Some differences in member characteristics between the Kaiser Permanente Regions were apparent; these reflect differences in the communities in which Kaiser Permanente operates. In addition, this analysis confirms that the Kaiser Permanente population includes large numbers of individuals who span the range of income, education, and SVI subgroups.

Funders and journal editors ask researchers to reflect on the generalizability of their research particularly for observational studies or performance reports. There are 2 dimensions to generalizability: the population studied and the characteristics of the delivery system. Describing the population provides 1 perspective for understanding whether there are population-based differences that might inform conclusions about the external validity of a wide range of research and measurement activities.[18,19] The findings of this analysis suggest that research based on Kaiser Permanente members might be generalizable to the local areas served because the population demographics are very similar.

The proportional representativeness of the member population is of less concern for weighted, stratified, or adjusted analyses that can incorporate demographic characteristics.[20,21] In the context of such multivariate approaches, a more relevant question is whether there is adequate sample size in all subgroups of interest. The size and diversity of the Kaiser Permanente membership offers an effective platform for research. Although Kaiser Permanente members are often viewed as more advantaged because they have health insurance, this analysis demonstrates that differences between Kaiser Permanente members and the general population on sociodemographic characteristics were small and not consistent across Regions.

In this analysis, demographic characteristics and chronic condition prevalence were compared between Kaiser Permanente members, who are fully insured, and the general population, which includes some uninsured individuals. Most national data sets do not separately identify the uninsured, so this approach mirrors the way many comparative analyses are conducted. At a population level there were few substantial differences between Kaiser Permanente members and the general population, which is not surprising because the uninsured are a small proportion of the total population in most Regions. Further, reports and data from the National Health Interview Survey and other sources suggest that the uninsured population in the US in the post–Affordable Care Act era is diverse across a variety of characteristics, including income, education, and employment status.[22-24]

A second comparison was the prevalence of chronic conditions between Kaiser Permanente members and the general population. Because uninsured people have lower rates of access to health services and are less likely to receive preventive services than the insured, this might contribute to differences.[25,26] Lower rates of access might be associated with lower rates of diagnosed chronic conditions. Lower rates of preventive services use might be associated with higher rates of preventable chronic conditions. Summary statistics on health status of the uninsured often show lower rates of specific chronic conditions and a larger proportion who report fair or poor health.[24] This analysis demonstrated lower prevalence of chronic conditions among Kaiser Permanente members than in the general population. Although differences in insurance status might explain some of the differences, other social factors also likely contribute. However, most of the differences observed were small and in the direction one might expect. This should not be surprising given the relatively small proportion of the uninsured in the general population and reporting on the prevalence of disease rather than how well those chronic conditions are managed.

Limitations

Because the majority of Kaiser Permanente members were enrolled in the 2 California Health Plans (4.4 million and 4.6 million in Northern and Southern California, respectively), enterprise estimates for Kaiser Permanente were driven by the characteristics of the membership residing in California.

Data source differences create some limitations in making comparisons between Kaiser Permanente members and the communities in which Kaiser Permanente operates. Kaiser Permanente member demographic data were based primarily on directly collected information (age, sex, race and ethnicity) whereas community comparisons were based on census estimates. Use of geospatially linked income and education data is a common practice in health system–based research, where research often leverages secondary data. Although neighborhood-based information may not accurately reflect individual education or income, in large studies, characterizing the experience of aggregate groups using these methods is appropriate.

The prevalence of health conditions for the general population was based on self-report from the BRFSS survey, which was compared to diagnostic coding from Kaiser Permanente's electronic medical record or claims. Both sources are subject to potential issues with accuracy. Self-report data are subject to recall bias as well as issues related to health literacy, while diagnosis-based data reflect conditions coded during receipt of services and rely on access to care and other factors. In this analysis, Kaiser Permanente estimates were based on a 5-year look-back period to better approximate the framing of most BRFSS questions, which ask whether a "health professional has ever told you that you have" a specific condition. Appendix D summarizes studies on the conditions included here, which suggest that self-report is a reasonably accurate source of information, with expected performance characteristics varying with the prevalence of conditions.

The BRFSS currently uses a single definition for overweight and obesity, which does not reflect more recent research suggesting that lower thresholds may be appropriate for certain subgroups, such as individuals of Asian descent.[27] Particularly in the Hawaii and Northern California geographies where people of Asian descent make up a larger share of the population, rates of overweight and obesity may be understated for both Kaiser Permanente and the general population. In geographies where the proportion of individuals who are of Asian descent is different between the populations being compared, the differential in rate of overweight and obesity may be misstated as a result. As national data definitions are updated to reflect new insights, comparisons to Health Plan members can also be updated.

The health outcomes of patients of integrated health systems may be most generalizable to patients with access to similar care and services. Describing the processes of care delivery in a system may also be important to interpreting the generalizability of outcomes achieved in that system. Access to care, integration of primary and specialty care services, a unified comprehensive medical record, and extensive population health management activities are among the hallmark features of the Kaiser Permanente care delivery ecosystem. Although efforts to transform care are underway in many state Medicaid programs (eg, California's Section 1115 Waiver "California Advancing and Innovating Medi-Cal" (CalAIM) initiative),[28] patients of integrated health systems often experience access to coordinated care that is different from patients of the health care safety net or many privately insured patients who may coordinate their own care within preferred provider networks.

Conclusion

Kaiser Permanente members were like those living in the communities in which Kaiser Permanente operates, and the membership was sufficiently large and diverse to support research activities. This analysis is illustrative of an approach that health plans can take to understand and communicate the similarities and differences between their member populations and the populations of the communities they serve. An understanding of population comparability is essential for effective comparisons when data cannot be risk adjusted, such as was seen in the early stages of the COVID-19 pandemic in trying to understand the risks of infection, hospitalization, and death, and later in assessing vaccine uptake. Other health systems may wish to compile similar summary statistics to characterize their patients or members. Such transparency about the representativeness of populations from which quality measures, performance data, and research samples are drawn would improve understanding of the

interpretability and generalizability of findings they produce.

Supplementary Materials
Supplemental material is available at: https://www.thepermanentejournal.org/doi/10.7812/TPP/22.172#supplementary-materials.

REFERENCES
1. Barnish MS, Turner S. The value of pragmatic and observational studies in health care and public health. Pragmat Obs Res. 2017;8:49–55. DOI: https://doi.org/10.2147/POR.S137701
2. Black N. Why we need observational studies to evaluate the effectiveness of health care. BMJ. 1996;312(7040):1215–1218. DOI: https://doi.org/10.1136/bmj.312.7040.1215
3. Stoto M, Oakes M, Stuart E, Priest EL, Savitz L. Analytical methods for a learning health system: 2. Design of observational studies. eGEMs. 2017;5(1):29. DOI: https://doi.org/10.5334/egems.251
4. Fleurence RL, Naci H, Jansen JP. The critical role of observational evidence in comparative effectiveness research. Health Aff (Millwood). 2010;29(10):1826–1833. DOI: https://doi.org/10.1377/hlthaff.2010.0630
5. Stoto M, Oakes M, Stuart E, Savitz L, Priest EL, Zurovac J. Analytical methods for a learning health system: 1. Framing the research question. eGEMs. 2017;5(1):28. DOI: https://doi.org/10.5334/egems.250
6. Dahabreh IJ, Kent DM. Can the learning health care system be educated with observational data? JAMA. 2014;312(2):129–130. DOI: https://doi.org/10.1001/jama.2014.4364
7. Vandenbroucke JP, von Elm E, Altman DG, et al. Strengthening the reporting of observational studies in epidemiology (STROBE): Explanation and elaboration. Ann Intern Med. 2007;147(8):W163–W194. DOI: https://doi.org/10.7326/0003-4819-147-8-200710160-00010-w1
8. Kaiser Permanente. Who We Are: Fast Facts. Accessed September 27, 2022. https://about.kaiserpermanente.org/who-we-are/fast-facts
9. United States Census Bureau. American Community Survey 5-Year Data (2009–2020). Accessed March 15, 2023. https://www.census.gov/data/developers/data-sets/acs-5year.html
10. Centers for Disease Control and Prevention. Chronic Disease and Health Promotion Data & Indicators. Accessed June 4, 2022. https://chronicdata.cdc.gov/Heart-Disease-Stroke-Prevention/Behavioral-Risk-Factor-Surveillance-System-BRFSS-N/ikwk-8git/data
11. Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System. Accessed June 4, 2022. https://www.cdc.gov/brfss/
12. Silva NM. The behavioral risk factor surveillance system. Int J Aging Hum Dev. 2014;79(4):336–338. DOI: https://doi.org/10.1177/0091415015574184
13. Adjaye-Gbewonyo D, Bednarczyk RA, Davis RL, Omer SB. Using the bayesian improved surname geocoding method (BISG) to create a working classification of race and ethnicity in a diverse managed care population: A validation study. Health Serv Res. 2014;49(1):268–283. DOI: https://doi.org/10.1111/1475-6773.12089
14. RAND Corporation. RAND bayesian improved surname geocoding. Accessed August 8, 2022. https://www.rand.org/health-care/tools-methods/bisg.html
15. Elliott MN, Fremont A, Morrison PA, Pantoja P, Lurie N. A new method for estimating race/ethnicity and associated disparities where administrative records lack self-reported race/ethnicity. Health Serv Res. 2008;43(5 Pt 1):1722–1736. DOI: https://doi.org/10.1111/j.1475-6773.2008.00854.x
16. Centers for Disease Control and Prevention Agency for Toxic Substances and Disease Registry. CDC/ATSDR Social Vulnerability Index 2018 Database US. Accessed August 8, 2022. https://www.atsdr.cdc.gov/placeandhealth/svi/data_documentation_download.html
17. Centers for Medicare and Medicaid Services. Chronic conditions data warehouse. Accessed June 4, 2022. https://www2.ccwdata.org/web/guest/condition-categories-chronic
18. Kukull WA, Ganguli M. Generalizability: The trees, the forest, and the low-hanging fruit. Neurology. 2012;78(23):1886–1891. DOI: https://doi.org/10.1212/WNL.0b013e318258f812
19. St Sauver JL, Grossardt BR, Leibson CL, Yawn BP, Melton LJ, Rocca WA. Generalizability of epidemiological findings and public health decisions: An illustration from the Rochester Epidemiology Project. Mayo Clin Proc. 2012;87(2):151–160. DOI: https://doi.org/10.1016/j.mayocp.2011.11.009
20. Cook RD, Weisberg S. Applied Regression Including Computing and Graphics. New York, NY, USA: John Wiley & Sons; 2009.
21. Greenland S. Applications of stratified analysis methods. In: Rothman KJ, Greenland S, Lash TL, eds. Modern Epidemiology. Philadelphia, PA: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008:283–302.
22. Cha A, Cohen R. Reasons for being uninsured among adults aged 18–64 in the United States, 2019 (Publication No. 382). National Center for Health Statistics; 2020. Accessed March 14, 2023. https://www.cdc.gov/nchs/data/databriefs/db382-H.pdf
23. Insure the Uninsured Project. Snapshot: Remaining uninsured in California. 2021. Accessed March 29, 2023. https://www.itup.org/wp-content/uploads/2021/12/Final.-ITUP-Snapshot-2021.pdf
24. National Center for Health Statistics. Interactive summary health statistics for adults: National Health Interview Survey 2019–2021. Accessed March 14, 2023. https://wwwn.cdc.gov/NHISDataQueryTool/SHS_adult/index.html
25. Collins SR, Davis K, Doty MM, Kriss JL, Holmgren AL. Gaps in health insurance: An all-American problem. The Commonwealth Fund. 2006. Accessed March 29, 2023. https://www.commonwealthfund.org/publications/fund-reports/2006/apr/gaps-health-insurance-all-american-problem
26. Schoen C, Doty MM, Collins SR, Holmgren AL. Insured but not protected: How many adults are underinsured? Health Aff (Millwood). 2005;Suppl Web Exclusives(Suppl1):W5-289–W5-302. DOI: https://doi.org/10.1377/hlthaff.w5.289
27. World Health Organization Expert Consultation. Appropriate body-mass index for Asian populations and its implications for policy and intervention strategies. Lancet. 2004;363(9403):157–163. DOI: https://doi.org/10.1016/S0140-6736(03)15268-3
28. California Department of Health Care Services. CalAIM: Our journey to a healthier California for all. Accessed September 23, 2022. https://www.dhcs.ca.gov/CalAIM/Pages/calaim.aspx