Denominator Demonstration Project

Investigators:
Paul Drawz, MD, MHS, MS, University of Minnesota Medical School
Peter Bodurtha, MPP, Hennepin Healthcare Research Institute
Tyler Winkelman, MD, MSc, Hennepin Healthcare
Steve Waring, DVM, PhD, Essentia Health

Background
Reliable population-level estimates of disease prevalence and incidence require accurate population counts. For the Minnesota Electronic Health Record Consortium (MNEHRC) COVID-19 Project, we used population counts from the US Census American Community Survey (ACS) to report vaccination percentages. However, for some localities and demographic groups the percent vaccinated exceeded 100%, suggesting ACS counts may have limitations for estimating disease prevalence based on medical records, even when those medical records represent an entire state population.

This report compares population counts from two sources: the 2019 5-year ACS (January 1, 2015 through December 31, 2019) and patient registries at participating MNEHRC sites.

Methods:
Primary Aim: Compare ACS and MNEHRC counts

Sites created a patient data file for MNEHRC populations counts as follows:

1. All patients who had one of the following:
    a. Any encounter (1/1/2019 – 9/30/2022)
    b. Any lab (1/1/2019 – 9/30/2022)
    c. A vital sign (1/1/2019 – 9/30/2022)
    d. A vaccine (through MMWR Week 38, 2022)
2. MN residents only (State was listed as Minnesota or ZIP was unknown)
3. Sites report on their assigned patients after centralized deduplication procedures
4. Individuals who died before the end of the study period were excluded.

We compared population distributions by age, sex, race and ethnicity, and language.

Results:
Comparisons between the ACS and MNEHRC-derived population counts are presented in **Table 1** and **Table 2**. Overall, MNEHRC population counts were higher across all demographic features except for age 0-18 and multirace compared to ACS counts.

Counts stratified by age group (overall, 0-18, 19-44, 45-64, and 65+) are provided in Table 1. The MNEHRC to ACS ratio is provided in **Table 2**, further demonstrating overall higher counts when using the full EHR database compared to the ACS. Ratios where the MNEHRC counts were more than 1.5 times larger than ACS counts were shaded red and more than 1.2-1.4 times larger were shaded yellow. Boxes were shaded green when the MNEHRC count were more than 20% smaller than ACS counts.

Estimates for older adults (ages 65+) using MNEHRC counts are higher than ACS counts (Table 2). Black or African American Minnesotans were more likely to be represented in MNEHRC counts across all age groups compared to the ACS (Table 2). If the MNEHRC population distribution is more reflective of the Minnesota population, disparities in COVID-19 vaccination rates between African American and White, non-Hispanic Minnesotans woulc be much larger than previously reported.

Discussion

This report underscores the significance of population count decisions and the potential to over- or underestimate the prevalence of disease based on those decisions. Reliable estimates of disease prevalence depend on accurate counts of both the number of individuals with the outcome of interest (numerator) as well as the size of the population (denominator). The ACS is commonly used to estimate the size of a given population for a specific geographic region and was widely used to determine COVID-19 vaccination rates at the local level. Our comparisons show that the ACS may vary substantially compared with populations that seek healthcare in a given state, contract or transmit infectious diseases, or use scarce healthcare resources. Additional study is needed to determine the most appropriate population counts when working with EHR data across a defined geographical area in the United States.

Whether ACS or MNEHRC population counts best reflect the population most relevant to disease prevalence in Minnesota is beyond the scope of the current report. However, ACS data are known to underestimate the population, in part, due to differential response rates across populations of color. Health system EHR data also has limitations, including imperfect deduplication that counts individuals more than once when they use multiple health systems, unobserved out-migration, and the overrepresentation of people who are more likely to seek healthcare. Despite these shortcomings, it remains possible that EHR population counts are a better reflection of the population for the purposes of estimating condition prevalence.

When responding to public health crises, officials are interested in people who are at risk of using healthcare resources or transmitting disease, not just residents in the state who have responded to a Census survey. These are the people represented in health care data. For this reason, we believe additional study is warranted to determine whether MNEHRC-derived population better represents populations for the purpose of public health surveillance and, in many cases, results in more accurate population-level estimates of condition prevalence. The choice is not trivial for a public health response – it can be the difference between reporting no disparities in vaccination or large disparities in vaccination.

Table 1. Demographic Distribution of Minnesota Population Estimates Using ACS and MNEHRC Counts by Age

	Ages 65+		Ages 45-64		Ages 19-44		Ages 0-18		All Ages	
	ACS	**MNEHRC**	**ACS**	**MNEHRC**	**ACS**	**MNEHRC**	**ACS**	**MNEHRC**	**ACS**	**MNEHRC**
Gender										
Male	388,709	727,099	724,749	818,951	955,274	1,196,241	700,920	679,136	2,769,652	3,440,356
Female	469,254	853,003	730,018	867,176	923,275	1,315,731	671,179	648,014	2,793,726	3,666,240
Race										
White, non-Hispanic	757,067	1,252,626	1,270,587	1,248,757	1,416,446	1,625,580	949,846	780,410	4,442,456	4,907,665
Black or African American, nh	15,562	44,646	63,957	115,135	138,179	250,696	130,015	173,908	350,306	584,400
Native American, nh	4,216	7,378	11,991	14,626	20,025	26,072	17,253	16,960	53,353	65,038
Asian or Pacific Islander, nh	13,296	29,907	46,969	64,764	126,848	143,332	75,219	85,981	264,643	323,989
Multirace, nh	4,605	2,327	14,673	6,362	49,598	24,957	75,990	37,708	145,250	71,751
Hispanic	5,699	21,086	45,216	78,259	124,078	172,346	120,718	105,049	299,252	376,745
Other/unknown/missing,nh	2,570	226,574	1,374	124,146	3,375	291,834	3,058	144,887	8,118	833,717
Language										
English	876,676	1,247,266	1,349,326	1,427,743	1,648,161	2,120,754	870,373	2,120,754	4,751,504	5,937,367
Spanish	3,629	12,523	20,961	46,461	44,228	24,550	10,828	77,107	79,646	182,779
Somali	2,802	6,256	6,435	10,529	16,224	77,107	5,956	24,550	31,417	73,416
Total	857,963	1,582,406	1,454,767	1,686,930	1,878,549	2,513,728	1,372,099	1,328,355	5,563,378	7,163,305

Table 2. Ratio of MNEHRC to ACS Count by Age

	Age 65+	Age 45-64	Age 19-44	Age 0-18	Overall
	MNEHRC vs. ACS (Ratio)				
Gender					
Male	1.87	1.13	1.25	0.97	1.24
Female	1.82	1.19	1.43	0.97	1.31
Race					
White, non-Hispanic	1.65	0.98	1.15	0.82	1.10
Black or African American, nh	2.87	1.80	1.81	1.34	1.67
Native American, nh	1.75	1.22	1.30	0.98	1.22
Asian or Pacific Islander, nh	2.25	1.38	1.13	1.14	1.22
Multirace, nh	0.51	0.43	0.50	0.50	0.49
Hispanic	3.70	1.73	1.39	0.87	1.26
Other/unknown/missing,nh	88.16	90.35	86.47	47.38	102.70
Language					
English	1.42	1.06	1.29	2.44	1.25
Spanish	3.45	2.22	0.56	7.12	2.29
Somali	2.23	1.64	4.75	4.12	2.34
Total	1.84	1.16	1.34	0.97	1.28