Intersectional inequities and longitudinal prevalence estimates of opioid use disorder in Massachusetts 2014–2020: a multi-sample capture-recapture analysis

Jianing Wang,[a,*] Dana Bernson,[b] Elizabeth A. Erdman,[b] Jennifer Villani,[c] Redonna Chandler,[c] David Kline,[d] Laura F. White,[e] and Joshua A. Barocas[f,**]

[a]Biostatistics Center, Massachusetts General Hospital, Boston, MA, USA
[b]Office of Population Health, Massachusetts Department of Public Health, Boston, MA, USA
[c]National Institute on Drug Abuse, National Institutes of Health, Bethesda, MD, USA
[d]Division of Public Health Sciences, Department of Biostatistics and Data Science, Wake Forest University School of Medicine, Winston-Salem, NC, USA
[e]Biostatistics Department, School of Public Health, Boston University, Boston, MA, USA
[f]Divisions of General Internal Medicine and Infectious Diseases, Department of Medicine, University of Colourado School of Medicine, Aurora, CO, USA

The Lancet Regional Health - Americas 2024;32: 100709

Published Online xxx https://doi.org/10.1016/j.lana.2024.100709

Summary

Background As overdoses continue to increase worldwide, accurate estimates are needed to understand the size of the population at risk and address health disparities. Capture-recapture methods may be used in place of direct estimation at nearly any geographic level (e.g., city, state, country) to estimate the size of the population with opioid use disorder (OUD). We performed a multi-sample capture-recapture analysis with persons aged 18–64 years to estimate the prevalence of OUD in Massachusetts from 2014 to 2020, stratified by sex and race/ethnicity.

Methods We used seven statewide administrative data sources linked at the individual level. We developed log-linear models to estimate the unknown OUD-affected population. Uncertainty was characterized using 95% confidence intervals (95% CI) on the total counts and prevalence estimates.

Findings The estimated OUD prevalence increased from 5.47% (95% CI = 4.89%, 5.98%) in 2014 to 5.79% (95% CI = 5.34%, 6.19%) in 2020. Prevalence among Hispanic females doubled (2.46% in 2014 to 4.23% in 2020) and prevalence rose to nearly 10% among Black non-Hispanic males and Hispanic males from 2014 through 2019. Estimates for Black non-Hispanic females more than doubled from 2014 through 2019 (3.39% to 7.09%), and then decreased to 5.69% in 2020.

Interpretation This study is the first to provide OUD prevalence trend estimates by binary sex and race/ethnicity at a state level using capture-recapture methods. Using these methods as the international overdose crisis worsens can allow jurisdictions to appropriately allocate resources and targeted interventions to marginalised populations.

Funding NIDA.

Keywords: Opioid use disorder; Population size estimation; Capture-recapture; Healthcare based surveillance data; Health disparities

Introduction

The epidemiology and demographics of opioid-related overdoses have changed worldwide over the past several decades. Such changes are particularly noticeable in the U.S. where overdoses once attributable to prescription opioids like oxycodone and OxyContin have been largely replaced by those attributable to heroin and illicitly manufactured fentanyl.[1,2] These epidemiological changes disproportionately impact certain populations. Since 1999, opioid overdose deaths among females have[3] risen 1326%, compared to 901% for males.[4,5] Since 2013, opioid-related overdose deaths have risen

*Corresponding author. Massachusetts General Hospital, 5th Floor, Suite 560, 50 Staniford Street, Boston, MA, USA, 02114.
**Corresponding author. University of Colourado School of Medicine, 8th Floor, Academic Office 1, Mailstop B180, 12631 E 17th Ave, Aurora, CO, 80045, USA.
E-mail addresses: jwang145@mgh.harvard.edu (J. Wang), Joshua.Barocas@CUAnschutz.edu (J.A. Barocas).

Research in context

Evidence before this study

Before this study, a PubMed search from 2014 to 2023 using terms 'prevalence of opioid use disorder by sex and race/ethnicity', 'opioid use disorder prevalence by demography' highlighted a gap in evidence for demographic-specific prevalence accounting for underreporting in health administrative data. Opioid-related overdoses have evolved, shifting from prescription opioids to heroin and illicit fentanyl, affecting sociodemographic groups unevenly. In the U.S., since 1999, there has been a rise in opioid overdose deaths among females, outpacing that of males – a stark 1326% increase for females compared to 901% for males. Furthermore, since 2013, Black non-Hispanic and Hispanic populations in the U.S. have seen a disproportionate surge in opioid-related overdose deaths. While multifaceted, these disparities in opioid overdose largely emanate from systemic issues like differential access to health care and entrenched structural racism and sexism. Females, compared to males, face disparities in screening, treatment, and emergency interventions for opioid-related incidents. Similarly, compared to their White non-Hispanic counterparts, people of colour encounter numerous barriers to preventive and opioid use disorder (OUD)-specific care. Estimating the actual prevalence of OUD, a significant driver for opioid overdoses, is a complex endeavor. Traditional methods relying on surveys and administrative claims often overlook crucial population segments, especially marginalized groups and those distanced from healthcare systems. The crossroads of sex and race/ethnicity further complicates OUD prevalence estimation at the population level, underlining an urgent need for tailored interventions.

Added value of this study

This study offers in depth insights into the nuanced disparities of OUD prevalence in Massachusetts between 2014 and 2020. Massachusetts is a state that has been at the forefront of the U.S. overdose epidemic with fentanyl arriving earlier than the rest of the nation. Leveraging indirect estimation methods, which are increasingly recognized as superior to traditional direct estimation, particularly for stigmatized diseases like substance use, the findings revealed that the OUD prevalence for adults in the region fluctuated between 5.47% and 6.72%. Notably, the study highlights profound intersectional disparities, with Hispanic females exhibiting the most significant surge in prevalence rates, at times surpassing 10% in certain demographics such as Black non-Hispanic and Hispanic females, and Hispanic males. These insights highlight the limitations of direct estimation methods, given that the derived prevalence was nearly double that directly derived from observed cases. Beyond numbers, the study uncovers systemic discrepancies, with populations of colour being significantly underrepresented in surveillance data when compared to the White non-Hispanic population. This underscores potential structural barriers in healthcare accessibility, suggesting disparities that inhibit certain demographics from engaging with available care resources. The study also demonstrates the importance of comprehensive linked datasets that include data on the different touchpoints of individuals with OUD. These findings not only underscore the magnitude of the OUD epidemic but also illuminate the urgent need for tailored interventions, equitable resource allocation, and inclusive data collection methodologies to better serve the diverse populations at risk.

Implications of all the available evidence

The study indicates a pivotal shift in the opioid crisis within the U.S. Traditional epidemiological methods, based primarily on direct estimation, have fallen short in providing a comprehensive view of the magnitude of OUD prevalence, particularly among stigmatized and marginalized populations. This gap in understanding has, in turn, compromised the allocation of limited resources for effective intervention and prevention. Particularly concerning are the stark disparities emerging at the intersection of race/ethnicity and sex, highlighting systemic issues and barriers within healthcare and surveillance systems. Such evidence mandates an urgent reconsideration of the tools and methodologies we employ to study and combat OUD, emphasizing the necessity for indirect estimation techniques that can better capture the nuances and magnitude of the crisis.

disproportionately among the Black non-Hispanic and Hispanic populations.[6–8]

Within the U.S., Massachusetts has endured some of the highest overdose rates. With the arrival of fentanyl in Massachusetts in 2013, years before much of the rest of the U.S., overdose deaths have been higher than the national average since 2014.[9] Recent data show that the most dramatic increases in overdose deaths in the state are among Black non-Hispanic, Hispanic, and American Indian individuals.[10]

These sex and race/ethnicity inequities in opioid-related overdose are multifactorial. They are likely reflective of barriers to access to health care, structural racism, and sexism within the health care system. Compared to males, females are less often screened for opioid use disorder (OUD),[11] less likely to receive and/or be retained on medications for OUD (MOUD),[3,12] and less likely to receive naloxone by EMS for a suspected opioid-related overdose.[13,14] Compared to White Non-Hispanic people, people in Black non-hispanic, Others non-hispanic, and Hispanic groups (henceforth, these groups will be referred to as non-white) are less likely to have access to preventive health care and poorer access to emergency department (ED) and treatment for OUD.[15–17] If certain people are not being identified as at risk for opioid-related overdose and not receiving the

care they need, then it follows that opioid-related overdoses will increase disproportionately. The problem is that not being identified by the health care system makes the underlying population at risk seem smaller than it really is. Put another way: if one is not able to access care, then they do not get counted, thus artificially decreasing the size of the population.

The true size of the populations with OUD – a major contributing factor for opioid overdose –remains unknown. As noted above, this is largely due to limitations of how prevalence of stigmatized diseases such as OUD is estimated. Such estimates often rely on national surveys (e.g., National Survey on Drug Use and Health [NSDUH]) and administrative claims, which miss large swaths of the population, such as those who have limited access to healthcare systems and marginalized populations.[1,18] Also unknown is how the intersection of sex and race/ethnicity impacts the diagnosis and prevalence of OUD at the population level. A greater understanding of OUD prevalence among stratified population groups is important to scale targeted interventions that reduce morbidity and mortality appropriately.

Indirect estimation techniques such as capture-recapture are increasingly used to estimate the prevalence of diseases like OUD.[19–22] In cases where multiple surveillance systems are linked at the individual level, as in our study, the capture-recapture method exploits this linkage to estimate the true prevalence of a disease. This approach analyses the overlap in reporting across the systems. A larger estimated undetected population results from minimal overlap between data sources, indicating that many cases are likely unobserved. Conversely, substantial overlap suggests that the systems are effectively capturing individuals, leading to a smaller estimate of the undetected population. By this inference of the likelihood of non-detection, the capture-recapture method subsequently calculates the size of the population that never got detected. Though powerful, such a method has not been used for intersectional OUD prevalence estimates, which is key to reducing overdose deaths in vulnerable populations. As part of the HEALing Communities Study (HCS), a multi-state trial to reduce overdose deaths funded by the National Institute on Drug Abuse and the Substance Abuse and Mental Health Services Administration,[23] a workgroup was formed to use state-level healthcare surveillance data with the goal of helping communities understand the size of the population with opioid misuse and/or OUD for the purposes of planning and implementing evidence-based interventions. We performed this analysis as part of a larger effort within this group in the HCS.[21,24] We conducted a multi-sample stratified capture-recapture analysis to estimate the annual OUD prevalence in Massachusetts between 2014 and 2020 by sex and race/ethnicity. We hypothesized differences in OUD prevalence over time by sex and race/ethnicity groups, and that the differences in these prevalence estimates were in part influenced by the variation in the ability of the data sources in identifying OUD cases according to sex or race/ethnicity group. This analysis contributes to the worldwide literature on using linked datasets to address overdoses.

Methods

Data sources and OUD measures

We used the Massachusetts Public Health Data Warehouse (PHD) for this analysis. The PHD is a comprehensive database that contains more than 25 state administrative datasets from 2011 to 2020 that can be linked at the individual-level.[25] The datasets were originally from 28 sources and agencies and the database is only available upon request.[25] The database includes all Massachusetts residents who have at least one medical, pharmacy, or dental insurance claim in the state's All-Payer Claims Database (APCD).[26] A probabilistic methodology that was developed by MDPH links identifiers from each dataset to identifiers in the state's All Payer Claims Database. Linkage rates ranged from 70% to 99.8%, with an overall average linkage rate of 94.6%. This study protocol (Pro00038088) was approved by Advarra Inc., the HEALing Communities Study single Institutional Review Board.

For this analysis, we used seven data sources with sufficient information to define an OUD variable: the Massachusetts APCD medical claims; Bureau of Substance Addiction Services (BSAS) addiction treatment records; Acute Care Hospital Inpatient, Emergency Department, and Outpatient Observation discharge records (i.e., Case Mix); Registry of Vital Records and Statistics (RVRS) death certificates (Deaths); RVRS birth certificates (Births); Massachusetts Ambulance Trip Record Information System trip reports (MATRIS); and Prescription Monitoring Program records (PMP). Supplemental Table S1 describes the definitions used for OUD case identification within individual data sources. We combined the previously defined variables[20] and additional diagnosis codes for OUD informed by other studies[21,22] to identify the target population. Because APCD, Case Mix, and Births are all similar data sources (from "traditional" health care settings) – the cases identified from any of them are attributable to the diagnosis during the health care settings visits, assessing them as separate "captures" would result in dependencies and, thus, violate the independence assumption required by the capture-recapture (refer to the 'Generic Model Formulation' section in the Supplemental Material for more details about capture-recapture methodology).[27] We, therefore, combined these three data sources into one. Functionally, this means that if an individual was identified with OUD in any of these data sources in a given year they were counted as being captured from the combined data source in that year. We labelled the cases with

identifiable OUD status from each of the remaining individual data sources.

Inclusion and exclusion criteria

We included individuals identified as having OUD in any of the above-listed data sources between 2014 and 2020 (2014 was the first year in which all data sources were complete). We included individuals aged 18–64 years, as this age group has been most affected by opioid overdose deaths.[28] Given that this was also the target age group of the HCS, this ensures consistency with other states' analyses in the work group under the HCS.[21,22] Because of the administrative nature of the data, we were limited to sex and race/ethnicity categories that were standardized in the PHD. Sex was dichotomously defined as male or female. Race/ethnicity was defined as White non-Hispanic, Black non-Hispanic, Asian/Pacific Islander (PI) or other races non-Hispanic (which includes American Indian; hereafter Other non-Hispanic), or Hispanic. We excluded people without recorded sex and/or race/ethnicity, because these cases would have resulted in an unknown denominator in our stratified prevalence estimates.

Analytic approach

To build models for estimating the overall and stratified prevalence, we first identified persons with OUD in each data source as they met definitions in Supplemental Table S1 at any time within a given year and independently across years. The identified individuals from each data source were matched according to their matching identification number based on a cross-agency linkage methodology that was created by the Massachusetts Department of Public Health. We then developed contingency tables by aggregating the observed people with OUD (i.e., those who were identified in at least one data set using the OUD variables listed in Supplemental Table S1) to summarize the number of these individuals per year (shown in Supplemental Table S2). To account for the heterogeneity between sex and race/ethnicity groups, we adopted a stratified analysis by disaggregating the contingency table into eight group-specific tables and estimating the group-specific prevalence separately.[27] In doing this, we assumed that the populations were homogeneous and had a similar engagement with the data systems, satisfying the equal-catchability assumption required for capture-recapture analyses.[27] A graphical flow demonstrating the process from case identification to the generation of the contingency table is shown in Supplemental Figure S1.[29]

Next, we constructed a series of log-linear models[29,30] for 2014 to 2020 (see Supplemental Appendix for more description of model formulation). Throughout the model selection, we constructed separate models for each sex and race/ethnicity group by determining an appropriate parsimonious model representing the best data relationship. To ensure consistency and efficiency, the final model for each stratum was chosen through an automated selection process, considering Poisson or Negative Binomial distributional assumptions and different degree of data source interactions. In addition to the main data source effects, we allowed interactions between data sources to adjust for the data source dependencies. The final model for each stratum was chosen using the lowest Akaike Information Criterion (AIC). We computed empirical 95% confidence intervals (CIs) of aggregated estimates using parametric bootstrapping. All statistical analyses were performed using SAS version 3.81 (Enterprise Edition) (SAS Institute Inc., Cary, NC, USA) and R (R Core Team, 2023).

Small sample adjustment

Although stratification can help reduce the impact of heterogeneity and provide group-specific estimates, it reduces the group sample sizes, resulting in unrealistic estimates. This occurred in two groups: the Other non-Hispanic group in 2014 and the Hispanic group in 2019. For these two estimates, we adjusted the data before the model fitting by adding $(0.5)^{k-1}$ to each cell in the contingency table, where k is the number of data sources.[31] There are other methods to adjust the small samples prior to the automated model selection, but they have been illustrated to have similar performance through comprehensive simulation studies.[32]

Role of data sources and "capturability"

To understand how frequently different population groups were represented in each data source, we constructed two-way contingency tables of the counts of identified OUD cases by sex-race/ethnicity groups and individual data sources for each year. We hypothesized that the White non-Hispanic group would be more likely to be found in each data source compared to other racial/ethnic groups, regardless of sex. We performed a set of chi square tests on the contingency tables to compare the frequencies of appearing in a given data source between a pair of demographic groups. For each sex, we compared the Black non-Hispanic or Hispanic group against the White non-Hispanic group based on the hypothesis that the coverage of the data for White non-Hispanic group might differ from the other groups. In doing this, we determined whether different sex and race/ethnicity groups were more or less likely to be found in a data source.

Sensitivity analyses

We conducted a sensitivity analysis to evaluate the impact of using data sources with low detection probabilities. Specifically, we investigated the estimates without the Death dataset, which, from empirical estimates, had low capturability of the OUD cases. These were generated from the same automated model selection process as noted above. We evaluated the impact of

the prevalence estimation with and without the Death data source cases by assessing the overlap between 95% CIs for overall prevalence.

Role of the funding source

This research was supported by the National Institutes of Health and the Substance Abuse and Mental Health Services Administration through the NIH HEAL Initiative under award UM1DA049412. The funder played no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; nor the decision to submit the manuscript for publication. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. Drs. Villani and Chandler were substantially involved in UM1DA049412, consistent with their roles as Scientific Officers. The views expressed in this paper are those of the authors only.

Results

Prevalence of OUD in the overall adult population

We excluded 4–5% of identified individuals (equating to 5000–7000 individuals) from the analysis due to missing demographic information. Negative Binomial distribution hierarchical log-linear models were selected for all strata for each year (see Supplemental Appendix for the complete model selection process). The estimated prevalence of OUD in Massachusetts among persons ages 18 to 64 increased from 5.47% (95% CI = 4.89%, 5.98%) in 2014 to 5.79% (95% CI = 5.34%, 6.19%) in 2020 (Table 1). The annual trend of the estimated prevalence relative to the observed OUD cases from MA PHD data sources is shown in Fig. 1. The number of persons estimated to have OUD increased from 242,470 (95% CI = 216,968, 265,349) in 2014 to 258,258 (95% CI = 238,011, 276,253) in 2020, an overall increase of 15,788 persons with OUD. The increase from 2014 to 2020 was not statistically significant through a post-hoc Poisson regression analysis. The number of "known" persons increased modestly during this period (107,494–113,080), and the proportion of "known" persons with OUD to the total estimated did not vary substantially over time, varying between 40.2% (estimated in 2018) and 44.8% (estimated in 2016). Importantly, across all years, the total estimated OUD prevalence (combining known and unknown groups) was more than two-fold higher than the calculated prevalence using only the known population counts (Table 1, Fig. 1).

Prevalence of OUD by sex

Males consistently had higher known, unknown, and total prevalence estimates than females (Supplemental Table S3). Overall estimated prevalence among males increased from 6.77% (95% CI = 5.74%, 7.66%) in 2014 to 7.34% (95% CI = 6.48%, 8.09%) in 2020. In contrast, the estimated prevalence trend among females fluctuated more across years and ranged from 4.22% (95% CI = 3.68%, 4.70%) in 2014 to 4.30% (95% CI = 3.96%, 4.60%) in 2020. Among both males and females, the proportion of known persons to the total with OUD fluctuated over the study period but did not qualitatively differ between the two populations over time (Supplemental Figure S2).

Prevalence of OUD by race/ethnicity

Estimated prevalence increased among the Black non-Hispanic (5.43% in 2014 to 7.71% in 2020) and Hispanic (4.48% in 2014 to 6.15% in 2020) populations over time but was relatively stable among the White non-Hispanic population (6.09% in 2014 to 6.10% in 2020) and the Other non-Hispanic population (1.11% in 2014 to 1.51% in 2020) (Supplemental Table S4). While the Black non-Hispanic and Hispanic populations account for only 6.80% and 11.90%, respectively, of the general Massachusetts population (whereas the White non-

Year	Known OUD[a]	Prevalence of known OUD (%)[b]	Estimated unknown OUD			Total estimated OUD[c]			Estimated prevalence (%)[d]		
			Estimates	95% LB	95% UB	Estimates	95% LB	95% UB	Estimates	95% LB	95% UB
2014[e]	107,494	2.42	134,976	109,474	157,855	242,470	216,968	265,349	5.47	4.89	5.98
2015	112,796	2.53	141,440	114,433	164,549	254,236	227,229	277,345	5.71	5.11	6.23
2016	117,850	2.64	144,836	113,817	171,958	262,686	231,667	289,808	5.89	5.19	6.49
2017	120,818	2.70	173,463	142,358	201,622	294,281	263,176	322,440	6.57	5.88	7.20
2018	121,090	2.70	180,125	144,581	210,511	301,215	265,671	331,601	6.72	5.93	7.40
2019	119,570	2.67	159,546	130,490	185,378	279,116	250,060	304,948	6.23	5.59	6.81
2020	113,080	2.54	145,178	124,931	163,173	258,258	238,011	276,253	5.79	5.34	6.19

Notes: OUD, opioid use disorder; LB, lower bound; UB, upper bound. [a]Total known counts were obtained directly from MA PHD. [b]Prevalence of Known OUD is computed by dividing the Known OUD size by the general population size. [c]Total estimated OUDs are the sum of known and estimated unknown OUD population sizes, along with 95% CIs labeled as 95% LB and UB. [d]Estimated prevalence uses 2020 US census data for the general population size. (Yr. 2014 = 4,434,642, Yr. 2015 = 4,450,444, Yr. 2016 = 4,462,891, Yr. 2017 = 4,479,308, Yr. 2018 = 4,483,062, Yr. 2019 = 4,476,770, Yr. 2020 = 4,460,035). These sizes have excluded people <18-year-old or >64-year-old. The data was obtained from US census data. [e]The results for the years 2014 and 2019 are derived from the analysis of the small sample adjusted data. For all the rest years, the original data were used in the model fitting and selection.

Table 1: **Capture-recapture estimates of Opioid Use Disorder (OUD) prevalence in Massachusetts for residents aged 18–64 years, 2014–2020.**

Sex	Race/ethnicity[a]	Year	Known OUD	Estimated unknown OUD			Total estimated OUD[b]			Estimated prevalence (%)		
				Estimates	95% LB	95% UB	Estimates	95% LB	95% UB	Estimates	95% LB	95% UB
Male	White (Non-Hispanic)	2014	53,204	63,719	45,806	88,637	116,923	99,010	141,841	7.25	6.14	8.80
		2015	55,577	63,499	44,301	91,017	119,076	99,878	146,594	7.42	6.23	9.14
		2016	57,515	65,719	46,674	92,533	123,234	104,189	150,048	7.73	6.53	9.41
		2017	57,760	86,718	65,322	115,122	144,478	123,082	172,882	9.11	7.76	10.90
		2018	57,575	90,378	65,406	124,885	147,953	122,981	182,460	9.40	7.81	11.59
		2019	56,249	55,400	37,980	80,810	111,649	94,229	137,059	7.15	6.04	8.78
		2020	52,725	63,636	47,802	84,717	116,361	100,527	137,442	7.53	6.51	8.90
Male	Black (Non-Hispanic)	2014	3458	7868	4160	14,882	11,326	7618	18,340	7.65	5.15	12.39
		2015	3595	11,505	6934	19,089	15,100	10,529	22,684	10.00	6.97	15.03
		2016	3986	10,038	6574	15,326	14,024	10,560	19,312	9.11	6.86	12.54
		2017	4340	7072	4961	10,081	11,412	9301	14,421	7.26	5.92	9.18
		2018	4442	15,089	8792	25,897	19,531	13,234	30,339	12.26	8.31	19.05
		2019	4695	13,935	8731	22,242	18,630	13,426	26,937	11.55	8.32	16.70
		2020	4718	11,218	8365	15,044	15,936	13,083	19,762	9.80	8.05	12.15
Male	Others[a] (Non-Hispanic)	2014	685	2637	918	7579	3322	1603	8264	1.88	0.91	4.67
		2015	768	1959	853	4498	2727	1621	5266	1.49	0.89	2.88
		2016	840	2194	957	5032	3034	1797	5872	1.61	0.95	3.11
		2017	846	2199	1018	4749	3045	1864	5595	1.56	0.95	2.86
		2018	901	2938	1378	6261	3839	2279	7162	1.92	1.14	3.59
		2019	869	2245	1118	4508	3114	1987	5377	1.52	0.97	2.63
		2020	880	5096	2184	11,890	5976	3064	12,770	2.88	1.48	6.16
Male	Hispanic	2014	7529	7974	5471	11,622	15,503	13,000	19,151	6.58	5.52	8.13
		2015	8149	15,734	11,448	21,625	23,883	19,597	29,774	9.81	8.05	12.23
		2016	8946	13,778	9476	20,033	22,724	18,422	28,979	9.07	7.36	11.57
		2017	9602	15,879	11,474	21,976	25,481	21,076	31,578	9.85	8.15	12.21
		2018	9837	13,059	9649	17,675	22,896	19,486	27,512	8.60	7.32	10.33
		2019	10,100	13,947	10,347	18,799	24,047	20,447	28,899	8.87	7.54	10.66
		2020	9898	12,606	9624	16,512	22,504	19,522	26,410	8.13	7.05	9.54
Female	White (Non-Hispanic)	2014	37,250	46,338	34,404	62,411	83,588	71,654	99,661	4.98	4.27	5.94
		2015	39,149	40,035	30,606	52,368	79,184	69,755	91,517	4.74	4.18	5.48
		2016	40,473	41,942	28,867	60,938	82,415	69,340	101,411	4.97	4.18	6.12
		2017	41,640	49,297	34,918	69,598	90,937	76,558	111,238	5.52	4.65	6.75
		2018	41,586	46,614	35,015	62,055	88,200	76,601	103,641	5.40	4.69	6.34
		2019	40,783	44,025	30,621	63,296	84,808	71,404	104,079	5.24	4.41	6.43
		2020	38,232	37,130	28,342	48,644	75,362	66,574	86,876	4.71	4.16	5.43
Female	Black (Non-Hispanic)	2014	2097	2993	1574	5692	5090	3671	7789	3.30	2.38	5.05
		2015	2087	5085	2156	11,993	7172	4243	14,080	4.57	2.70	8.96
		2016	2327	5137	2706	9748	7464	5033	12,075	4.66	3.14	7.54
		2017	2473	4125	2211	7696	6598	4684	10,169	4.05	2.87	6.24
		2018	2574	4495	2456	8228	7069	5030	10,802	4.27	3.04	6.53
		2019	2666	9198	4407	19,197	11,864	7073	21,863	7.09	4.23	13.07
		2020	2523	7064	4106	12,154	9587	6629	14,677	5.69	3.94	8.72
Female	Others[a] (Non-Hispanic)	2014	368	323	171	610	691	539	978	0.38	0.29	0.53
		2015	397	343	176	667	740	573	1064	0.39	0.30	0.56
		2016	388	329	165	656	717	553	1044	0.37	0.28	0.53
		2017	436	351	189	652	787	625	1088	0.39	0.31	0.54
		2018	430	314	184	537	744	614	967	0.36	0.30	0.47
		2019	428	266	147	478	694	575	906	0.33	0.27	0.43
		2020	374	35	16	75	409	390	449	0.19	0.18	0.21
Female	Hispanic	2014	2903	3123	1839	5302	6026	4742	8205	2.46	1.93	3.34
		2015	3074	3280	1882	5718	6354	4956	8792	2.51	1.96	3.47
		2016	3375	5700	3349	9700	9075	6724	13,075	3.49	2.58	5.02
		2017	3721	7821	4240	14,426	11,542	7961	18,147	4.30	2.96	6.76
		2018	3745	7237	4128	12,688	10,982	7873	16,433	3.97	2.85	5.94

(Table 2 continues on next page)

Sex	Race/ethnicity[a]	Year	Known OUD	Estimated unknown OUD			Total estimated OUD[b]			Estimated prevalence (%)		
				Estimates	95% LB	95% UB	Estimates	95% LB	95% UB	Estimates	95% LB	95% UB
(Continued from previous page)												
		2019	3780	20,530	6503	64,819	24,310	10,283	68,599	8.65	3.66	24.41
		2020	3730	8393	5051	13,944	12,123	8781	17,674	4.23	3.06	6.17

[a]The Other (Non-Hispanic) group includes Asian/PI non-Hispanic, American Indian, and Other non-Hispanic populations. [b]Due to the rounding, the summation of sex and race groups might not be perfectly the same as the estimated overall population sizes.

Table 2: **Capture-recapture estimates of Opioid Use Disorder (OUD) prevalence in Massachusetts for residents aged 18–64 Years, 2014–2020, by sex and race/ethnicity.**

Hispanic population accounts for 70.80% of the general population), the estimated OUD prevalence was disproportionately high among people in non-white groups over the study period. In 2018, the OUD prevalence between all these groups did not qualitatively differ. By 2019, however, the prevalence in the Black non-Hispanic (9.28%) population was substantially higher than in the White non-Hispanic population (6.18%) and remained higher by more than 1% in 2020. The Hispanic population consistently had the highest estimated prevalence over the years, except for 2017. For people in non-white groups, a fluctuated proportion, between 20% and 48%, of the total OUD populations were observed in the data sources. The trend for the White non-Hispanic population, however, remained stable with proportions of known OUD to total OUD between 42% and 49% (Supplemental Figure S3). This indicates a consistent and higher degree of "capture" by administrative data sources for the White non-Hispanic populations compared with the other racial/ethnic groups.

Prevalence of OUD by sex and race/ethnicity strata

There were important findings at the intersection of sex and race/ethnicity. The known and estimated unknown OUD population sizes across demographic groups over the years are shown in Fig. 2. Among females, the White non-Hispanic population had the highest estimated prevalence in 2017 (5.52%, 95% CI = 4.65%, 6.75%) but the Black non-Hispanic population had the highest estimated prevalence in 2019 (7.09%, 95% CI = 4.23%, 13.07%) (Table 2). By 2020, the estimated prevalence for these two groups was comparable (4.71%, 95% CI = 4.16%, 5.43% for White non-Hispanic; 5.69%, 95% CI = 3.94%, 8.72% for Black non-Hispanic). Generally, White non-Hispanic and Other non-Hispanic females had a stable total estimated prevalence across the seven years whereas the estimated

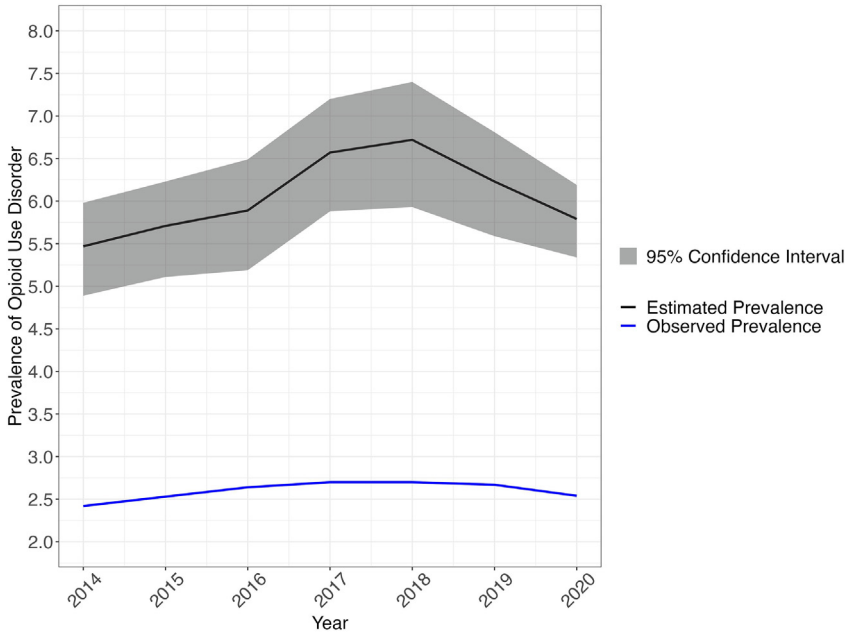

Fig. 1: **Longitudinal trend of the total estimated prevalence compared to the observed case prevalence**. Notes: This is a line graph of the prevalence of OUD by year among adults ages 18–64 in Massachusetts. The blue line indicates the prevalence of observed cases—those directly counted in the dataset. The black line and gray shades indicate the estimated total prevalence (sum of observed and unobserved) and parametric bootstrapped 95% CIs.

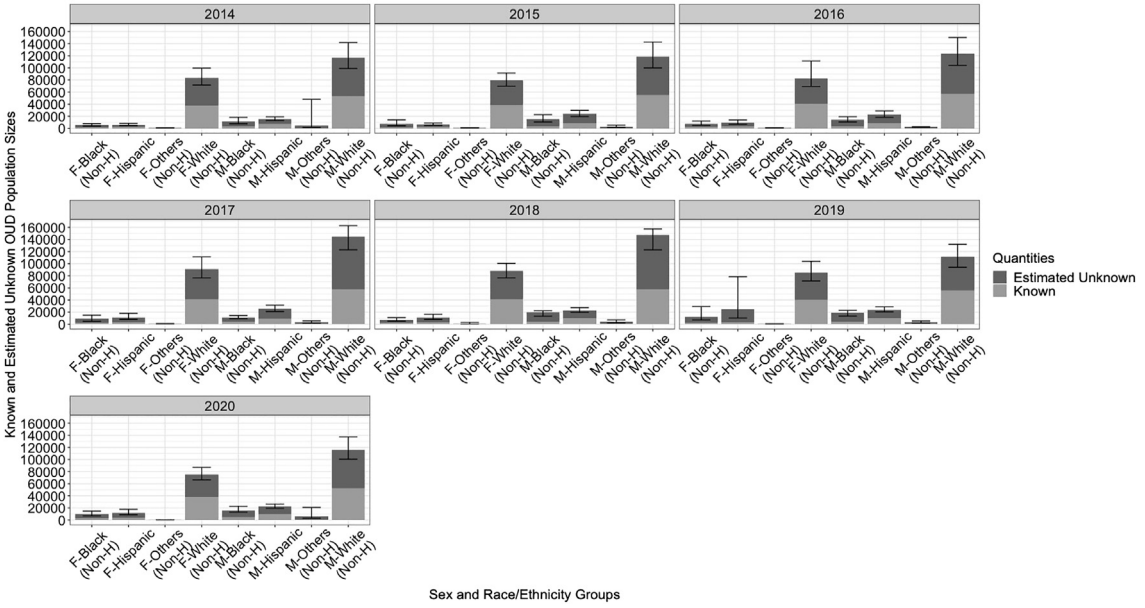

Fig. 2: **Known and capture-recapture estimates of OUD sizes in Massachusetts for residents aged 18–64 years, 2014–2020, by sex and race/ethnicity**. Notes: This is a panel of histogram plots showing the known OUD sizes (light gray colored), estimated unknown OUD sizes (dark gray colored), and 95% confidence intervals for each sex and racial group in a given year. 'F' indicates Female; 'M' indicates Male; 'H' indicates Hispanic. The Others non-Hispanic group include Asian/Pacific Islander non-Hispanic, Alaska Native/American Indian or Other non-Hispanic populations.

prevalence among Hispanic and Black non-Hispanic females nearly doubled over the study period (2.46% in 2014 to 4.23% in 2020 for Hispanic female, 3.30% in 2014 to 5.69% in 2020 for Black non-Hispanic females). Except for 2014 and 2017, Black non-Hispanic males had the highest prevalence among males. Among White non-Hispanic males, OUD prevalence estimates remained between 7 and 9% across the years. We also note different proportions of unknown OUD to total OUD for each population. For both the male and female White non-Hispanic populations, this ratio remained qualitatively unchanged from 2014 to 2020. The proportion of unknown OUD to total OUD population steadily increased for Black non-Hispanic and Hispanic females and were consistently higher for people in non-white groups compared to White non-Hispanic groups for both sex populations (Supplemental Figure S5).

Role of data sources and "capturability"

The contingency table of identified cases by demographic strata and data sources is visualized in Supplemental Figure S4. The p-values from chi-square tests for race/ethnicity groups within each sex group are displayed. In the years 2014 and 2017–2019, the observed capture patterns were not statistically significantly different between racial groups. However, both Black non-Hispanic and Hispanic females significantly differed from White non-Hispanic females, indicating a difference in interactions and capturability by the administrative data sources between these groups.

Sensitivity analyses

In sensitivity analyses that did not include Death data (i.e., overdose death data), we found that the total known counts of OUD cases were not significantly reduced compared to the main analysis. There were, however, slightly higher prevalence estimates with wider confidence intervals derived in the selected models (Supplemental Table S5). The 95% confidence intervals covered the estimates when we included Death data.

Discussion

This study aimed to develop OUD prevalence estimates for adults in Massachusetts from 2014 to 2020 and provide sex and race/ethnicity stratified estimates to understand intersectional disparities. OUD prevalence in Massachusetts adults during this period was estimated to be between 5.47% and 6.72%, with the most significant increase observed in Hispanic females. OUD prevalence estimates surpassed 10% in specific years for Black non-Hispanic and Hispanic females, and Hispanic males. These findings are among the first estimates of OUD prevalence using indirect estimation for sex and race/ethnicity and have global implications.

While traditional epidemiological methods have extensively reported drug use trends, our study's use of indirect estimation methods helps characterize the scope of OUD. It is well documented that direct estimation for stigmatized diseases like substance use is no longer adequate. Our prevalence estimates are nearly twice those calculated directly from observed cases in

the PHD and higher than those reported by NSDUH.[18] Insufficiently understanding the scope of the problem can lead to inadequate resource allocation for substance use treatment, harm reduction, and preventive services. How do we expect to narrow the consistently wide and ever-growing treatment gap without a thorough understanding of the size of the problem?

Narrowing this treatment gap is challenging due to the complex and demographically changing epidemic.[7,11] It is unsurprising that disparities exist at the intersection of race/ethnicity and sex regarding OUD. Opioid overdoses are increasing faster for people in non-white groups than the White non-Hispanic population both in Massachusetts and nationally. Since OUD is a risk factor for overdose, it logically follows that we would find similar patterns. However, the scope of the problem, especially among Hispanic females and for Black non-Hispanic males and females, is surprising. This signals structural barriers needing increased resources for these groups.

Another important disparity to address is that populations in non-white groups were consistently underrepresented in surveillance data compared to the White non-Hispanic population. For example, although the prevalence of OUD in the Black non-Hispanic and White non-Hispanic populations were comparable, only 20%–35% of the Black non-Hispanic population was identified in the PHD data, whereas nearly 50% of the White non-Hispanic population was identified. This indicates that 50% of the White non-Hispanic population is known to the administrative system while only 20–35% of the Black non-Hispanic population with OUD is. These differences were also noted in the intersectional population analyses (Fig. 2). Such patterns of the unknown proportions suggest several possible issues. First, it is possible that touchpoints within the system are easier to access for certain segments of the population—making it more likely that they are found. Alternatively, it might suggest an over-use or under-use of the system (and resources therein) by certain segments of the population. Both might be influenced by geographical barriers, knowledge or awareness of OUD, difficulty affording medical care, difficulty in accessing culturally appropriate care, or other access-related factors that result in people not showing up in these data sets. In either case, our findings should prompt systems and health providers to explore the causes and consequences of these patterns of "engagement" and to develop and test efforts that enhance equity.

There are limitations to the study. First, we assume that the behaviour of interacting with the data sources is consistent within a stratified population group. Second, race is a social construct, and some of the data sets used administratively determined (rather than self-reported) race. Due to small numbers of events, American Indian and Alaska Native residents were not analysed separately even though they have the highest rates of fatal opioid-related overdoses in MA.[33] Additionally, the racial categories in the PHD are restrictive and may not be fully encompassing of how an individual self-identifies, highlighting a need for updated categorization methods. The identified cases with missing demographic information were excluded from the analysis; however, these cases may represent a particularly vulnerable population affected by OUD. Third, although individuals who identify as non-binary sex are at higher risk for many diseases, we were only able to look at binary sex (male/female).[34] Significant portions of datasets likely used self-reported measures, but others used administrative methods. This variability is a limitation, potentially introducing ambiguities. Fourth, log-linear capture-recapture models can be sensitive to sparse data, evidenced by a sharp increase in prevalence for Black non-Hispanic and Hispanic groups. We accounted for sparse data by using established methods. However, simulation studies may be useful to quantify the sensitivity and inflation of the estimation of the unknown sizes. Fifth, we assumed that opioid-related overdose deaths were due to opioid overdose rather than overdose from other substances when opioids were present. Though MA DPH has attempted to validate this assumption through linkage with the medical examiner toxicology data, further investigation is suggested due to potential inaccuracies when multiple substances are present, as it would impact the data quality. In the current implementation of MAPHD, the Death data cannot distinguish between purposeful co-use of substance and drug contamination. Lastly, the methods used for defining OUD have undergone a careful discussion among clinical and public health experts, though there was not a formal clinical validation.

We found that longitudinal OUD prevalence estimates in Massachusetts by sex and race/ethnicity were substantially higher than using direct estimation methods, and mirrored trends for death involved with opioid-related overdose. Populations in non-white groups are suffering disproportionately from OUD and from overdose. While the estimates themselves are unsettling and should be used to inform resource allocation, engagement efforts, and service planning, our analysis demonstrates that there is much that we are missing with regard to the overdose epidemic. These methods are reproducible at nearly any geographic level and could help communities worldwide understand the scope of stigmatized diseases such as OUD.

Contributors

All of the authors made substantial contributions to the article. Wang, Barocas, Chandler, White, and Kline conceived the study, including the conceptualization and methodology investigation, and led its execution. All authors collaborated on the study design, analysis, and creation of tables and figures. All authors were involved in manuscript preparation and review. Bernson, Erdman, Barocas, and Wang had full access to all the raw data in the study and took responsibility for the integrity of the data and the accuracy of the data analysis. All authors were involved in manuscript preparation and review for manuscript submission. Barocas and Wang had final responsibility for decision to submit the manuscript.

Data sharing statement

The individual participant data will not be available. The definition of the Opioid Use Disorder and the non-stratified contingency table with small cell suppression will be available with the publication. Other than the main text and supplementary materials, no other documents will be available. The raw data will only be available upon request.

Declaration of interests

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Dr. White reports grant support for the research presented in this manuscript, which was paid to her respective institution. Dr. Kline reports grant support from National Institute on Drug Abuse for the research presented in this manuscript (R01DA052214). Dr. Barocas also reports membership on the scientific advisory committee for eMed. All other authors have declared that no financial support was received for the research of this manuscript and that there are no financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this paper.

Acknowledgements

Trial Registration: ClinicalTrials.gov Identifier: NCT04111939.

This research was supported by the National Institutes of Health and the Substance Abuse and Mental Health Services Administration through the NIH HEAL Initiative under award UM1DA049412 with additional supports from the National Institute on Drug Abuse [DP2DA051864, R01DA052214] and National Institute of General Medical Sciences [R35GM141821]. Drs. Chandler and Villani were substantially involved in UM1DA049412, consistent with their roles as Scientific Officers. They had no substantial involvement in the other cited grants. This study protocol (Pro00038088) was approved by Advarra Inc., the HEALing Communities Study single Institutional Review Board. We wish to acknowledge the participation of the HEALing Communities Study communities, community coalitions, community partner organizations and agencies, and Community Advisory Boards and state government officials who partnered with us on this study. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health, the Substance Abuse and Mental Health Services Administration or the NIH HEAL Initiative.℠

We also acknowledge the Massachusetts Department of Public Health for creating the unique, cross-sector database used for this project and for providing technical support for the analysis.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.lana.2024.100709.

References

1 Reuter P, Caulkins JP, Midgette G. Heroin use cannot be measured adequately with a general population survey. *Addict Abingdon Engl*. 2021;116(10):2600–2609.
2 Manchikanti L, Singh VM, Staats PS, et al. Fourth wave of opioid (illicit drug) overdose deaths and diminishing access to prescription opioids and interventional techniques: cause and effect. *Pain Physician*. 2022;25(2):97–124.
3 Huhn AS, Berry MS, Dunn KE. Review: sex-based differences in treatment outcomes for persons with opioid use disorder. *Am J Addict*. 2019;28(4):246–261.
4 Injury Facts. *Overdose death by select drugs, United States, 1999-2020*; 2023. [cited 2023 Feb 14] Available from: https://injuryfacts.nsc.org/home-and-community/safety-topics/drugoverdoses/data-details/.
5 Hedegaard H, Minininet. *Overdose death by select drugs, United States, 1999-2020* [cited 2023 Feb 14]; 2023. Available from: https://injuryfacts.nsc.org/home-and-community/safety-topics/drugoverdoses/data-details/usechs/products/databriefs/db428.htm.
6 Goodman-Meza D, Friedman J, Kalmin MM, et al. Geographical and socioeconomic disparities in opioid access in Mexico, 201515afety-topics/drugoverdoses/data-details/usechsta. *Lancet Public Health*. 2021;6(2):e88–e96.
7 Milam AJ, Furr-Holden D, Wang L, Simon KM. Health data disparities in opioid-involved overdose deaths from 1999 to 2018 in the United States. *Am J Public Health*. 2021;111(9):162727s/d.
8 Jordan A, Mathis M, Haeny A, Funaro M, Paltin D, Ransome Y. An evaluation of opioid use in black communities: a rapid review of the literature. *Harv Rev Psychiatry*. 2021;29(2):108–130.
9 *Drug Overdose Mortality by State*. Drug Overdose Mortality by State; 2022. [cited 2023 Aug 25] Available from: https://www.cdc.gov/nchs/pressroom/sosmap/drug_poisoning_mortality/drug_poisoning.htm.
10 Current opioid statistics. Current opioid statistics | Mass.gov. [cited 2023 Aug 25] Available from: https://www.mass.gov/lists/current-opioid-statistics.
11 Barbosa-Leiker C, Campbell ANC, McHugh RK, Guille C, Greenfield SF. Opioid use disorder in women and the implications for treatment. *Psychiatr Res Clin Pract*. 2021;3(1):3ni.
12 Dennis BB, Martin LJ, Naji L, et al. Sex-specific risk factors and health disparity among hepatitis C positive patients receiving pharmacotherapy for opioid use disorder: findings from a propensity matched analysis. *J Addict Med*. 2022;16(4):e248–e256.
13 Substance Abuse and Mental Health Services Administration. *Center for behavioral health Statistics and quality. Treatment episode data set (TEDS): 2004-2014*. Natl Admiss Subst Abuse Treat Serv; 2016. Available from: https://www.samhsa.gov/data/sites/default/files/2014_Treatment_Episode_Data_Set_National_Admissions_9_19_16.pdf.
14 Bettano A, Jones K, Fillo KT, Ficks R, Bernson D. Opioid-related incident severity and emergency medical service naloxone administration by sex in Massachusetts, 2013–2019. *Subst Abuse*. 2022;43(1):479–485.
15 Phan MT, Tomaszewski DM, Arbuckle C, et al. Racial and ethnic disparities in opioid use for adolescents at US emergency departments. *BMC Pediatr*. 2021;21(1):252.
16 Andraka-Christou B. Addressing racial and ethnic disparities in the use of medications for opioid use disorder. *Health Aff*. 2021;40(6):920–927.
17 Rosales R, Janssen T, Yermash J, et al. Persons from racial and ethnic minority groups receiving medication for opioid use disorder experienced increased difficulty accessing harm reduction services during COVID-19. *J Subst Abuse Treat*. 2022;132:108648.
18 McCance-Katz EF. *The national survey on drug use and health*; 2019. Available from: https://www.samhsa.gov/data/sites/default/files/reports/rpt29392/Assistant-Secretary-nsduh2019_presentation/Assistant-Secretary-nsduh2019_presentation.pdf.
19 Keyes KM, Rutherford C, Hamilton A, et al. What is the prevalence of and trend in opioid use disorder in the United States from 2010 to 2019? Using multiplier approaches to estimate prevalence for an unknown population size. *Drug Alcohol Depend Rep*. 2022;3:100052.
20 Barocas JA, White LF, Wang JN, et al. Estimated prevalence of opioid use disorder in Massachusetts, 2011-2015: a capture-recapture analysis. *Am J Public Health*. 2018;108(12):1675–1681.
21 Thompson K, Barocas JA, Delcher C, et al. The prevalence of opioid use disorder in Kentucky's counties: a two-year multi-sample capture-recapture analysis. *Drug Alcohol Depend*. 2023;242:109710.
22 Doogan NJ, Mack A, Wang J, et al. Opioid use disorder among Ohiohiodisorder in Kentuon: prevalence estimates from nineteen counties using a multiplier method. *Am J Epidemiol*. 2022;191:2098–2108.
23 El-Bassel N, Jackson RD, Samet J, Walsh SL. Introduction to the special issue on the HEALing communities study. *Drug Alcohol Depend*. 2020;217:108327.
24 Doogan NJ, Mack A, Wang J, et al. Opioid use disorder among Ohiohioties Study. *Drug Alcohol Depend*. 2020;217:108327.
25 Bharel M, Bernson D, Averbach A. Using data to guide action in response to the public health crisis of opioid overdoses. *NEJM Catal*. 2020;1(5) [cited 2022 Jul 12] Available from: https://catalyst.nejm.org/doi/full/10.1056/CAT.19.1118.
26 *Public Health Data Warehouse (PHD) Technical Documentation*; 2022 [cited 2023 Apr 18]. Available from: https://www.mass.gov/info-details/public-health-data-warehouse-phd-technical-documentation.
27 Tilling K, Sterne JA. Capture-recapture models including covariate effects. *Am J Epidemiol*. 1999;149(4):392–400.
28 Lippold KM. Racial/ethnic and age group differences in opioid and synthetic opioidoidemiol. *MMWR Morb Mortal Wkly Rep*. 1999;149(4):392–400 [Internet]. 2019;68. [cited 2023 Jun 8] Available from: https://www.cdc.gov/mmwr/volumes/68/wr/mm6843a3.htm.

29 Cormack RM. Log-linear models for capture-recapture. *Biometrics*. 1989;45(2):395395oid.

30 Manrique-Vallier D, Ball P, Sadinle M. Capture-recapture for casualty estimation and beyond: recent advances and research directions. In: Carriquiry AL, Tanur JM, Eddy WF, eds. *Statistics in the public interest*. Cham: Springer International Publishing; 2022. p. 15–31. (Springer Series in the Data Sciences). [cited 2022 Nov 18] Available from: https://link.springer.com/10.1007/978-3-030-75460-0_2.

31 Evans MA, Bonett DG. Bias reduction for multiple-recapture estimators of closed population size. *Biometrics*. 1994;50(2):388388ct.

32 Kurtz ZT. *Local log-linear models for capture-recapture*. Carnegie Mellon University; 2014 [cited 2022 Aug 18]. Available from: https://kilthub.cmu.edu/articles/thesis/Local_Log-Linear_Models_for_Capture-Recapture/6720452/1.

33 Commonwealth of Massachusetts. *Opioid-Related Overdose Deaths, All Intents, MA Residents – Demographic Data Highlights*; 2022 [cited 2023 Apr 18]. Available from: https://www.mass.gov/doc/opioid-related-overdose-deaths-demographics-december-2022/download.

34 Newcomb ME, Hill R, Buehler K, Ryan DT, Whitton SW, Mustanski B. High burden of mental health problems, substance use, violence, and related psychosocial factors in transgender, non-binary, and gender diverse youth and young adults. *Arch Sex Behav*. 2020;49(2):645–659.