University of Nebraska Medical Center
BREAKTHROUGHS FOR LIFE

University of Nebraska Medical Center
DigitalCommons@UNMC

Theses & Dissertations

Graduate Studies

Summer 8-13-2021

Perinatal Periods of Risk: Examination Of Data Quality & Inclusion Criteria, New Unbiased Reference Groups, And A Nationwide County-Level Analysis

Carol S. Gilbert
University of Nebraska Medical Center

Tell us how you used this information in this short survey.

Follow this and additional works at: https://digitalcommons.unmc.edu/etd

 Part of the Maternal and Child Health Commons

Recommended Citation

Gilbert, Carol S., "Perinatal Periods of Risk: Examination Of Data Quality & Inclusion Criteria, New Unbiased Reference Groups, And A Nationwide County-Level Analysis" (2021). *Theses & Dissertations*. 560.
https://digitalcommons.unmc.edu/etd/560

This Dissertation is brought to you for free and open access by the Graduate Studies at DigitalCommons@UNMC. It has been accepted for inclusion in Theses & Dissertations by an authorized administrator of DigitalCommons@UNMC. For more information, please contact digitalcommons@unmc.edu.

PERINATAL PERIODS OF RISK: EXAMINATION OF DATA QUALITY & INCLUSION CRITERIA, NEW UNBIASED REFERENCE GROUPS, AND A NATIONWIDE COUNTY-LEVEL ANALYSIS

By

Carol S. Gilbert

A DISSERTATION

Presented to the Faculty of the University of Nebraska Graduate College in partial fulfilment of the requirements for the degree of
Doctor of Philosophy

Medical Sciences Interdepartmental Area
(Preventive and Societal Medicine)

Under the Supervision of Professor Eleanor Rogan, PhD

University of Nebraska Medical Center
Omaha, Nebraska
July 28, 2021

Supervisory Committee :

John W. Sparks, MD

Melissa K. Tibbits, PhD

Christopher S. Wichman, PhD

William M. Sappenfield, MD, MPH
(University of South Florida)

ACKNOWLEDGEMENTS:

I thank my husband Ed Taylor, my mother Jean Eden and her spouse Lorene Ludy, my children Ian Taylor and Michael Taylor and their spouses Naomi Thyden and Natalie Nieman, and my grandson Liam Black Eyes. They have all been very patient and supportive. I also thank Bill Sappenfield for his mentorship for many years and the useful way he thinks, and Pam Xaverius and Chad Abresch for being the sparks that got this ball rolling, and for giving good advice and encouragement along the way.

ABSTRACT

Perinatal Periods of Risk: examination of data quality & inclusion criteria, new unbiased reference groups, and a nationwide county-level analysis

Carol Gilbert, PhD

University of Nebraska Medical Center, 2021

Supervisor: Eleanor G. Rogan, PhD

Records of births, infant deaths, and fetal deaths are compiled by the US Vital Records System and used to monitor population health and guide health policy. The Perinatal Periods of Risk Approach (PPOR) relies on vital records data to address fetal and infant mortality in US cities. It uses reference groups to estimate preventable mortality by risk period. To avoid biased analyses due to poor data quality for small and early infant and fetal deaths, an expert committee recommended that PPOR analyses exclude fetal deaths delivered at gestational age (GA) <24 weeks, and infant deaths and live births with birthweights (BW) <500 grams. Poor data quality has hindered creation of a national reference group and national analysis since then. Improvements in data systems and increasing survival of very preterm infants may have led to better data quality and the possibility of reducing the PPOR exclusion limits and excluding fewer deaths.

We developed a method to quantify underreporting by week of gestation (from 20 to 31 weeks) and by birthweight (in 100-gram intervals below 1500 grams, adjusting for state-level health-related factors. We found that differential reporting remains substantive at GA<24 weeks, and reporting requirements and health differences do not account for it. We cannot recommend lowering the original PPOR fetal GA limit. We found that a fetal BW limit is redundant once a gestation limit is in place but that an infant death BW limit of 400 or 500 grams is still needed. We then assessed the quality of data elements needed for creating reference groups and

formed a subset of states from which to draw a nearly unbiased set of national reference groups for use by communities. Based on a new national reference group and a study population of 100 large counties, we summarized a national analysis using percentile charts for six components of preventable mortality. The charts allow communities to compare their outcomes nationally as well as determining which component is locally predominant.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

PPOR	Perinatal Periods of Risk
GA	Gestational Age
BW	Birthweight
VLBW	Very Low Birthweight (less than 1,500 grams)
SUID	Sudden Unexplained Infant Death
MHP	Maternal Health and Prematurity Period of Risk
MHP BW	The birthweight distribution component of MHP
MHP SM	The birthweight-specific mortality component of MHP
MC	Maternal Care Period of Risk
NC	Newborn Care Period of Risk
IH	Infant Health Period of Risk
IH SUID	SUID component of IH
IH OTHER	Component of IH due to all causes other than SUID
US	United States of America

INTRODUCTION

The US Vital Records System gathers and organizes information about births, deaths, and fetal deaths from all states and territories. This electronic data contains detailed medical and social information used to monitor population health and conduct research to guide health policy. Although the vital records system has improved considerably over the last century (Brumberg, Dozor, & Golombek), data quality problems remain. Under-reporting of early *fetal deaths* has been documented at state, county, and hospital levels (Goyal, DeFranco, Kamath-Rayne, Beck, & Hall, 2017; Heck, Schoendorf, & Parker, 1999; Ramsay & Santella, 2011). Underreporting of very low birthweight (VLBW) *infant deaths* has also been documented (Heck et al., 1999; Kleinman, 1986). Although many states routinely follow up on births of very small infants to ensure that a death is recorded if appropriate (McCarthy, Terry, Rochat, Quave, & Tyler, 1980), others may not. Underreporting of deaths can vary by population (e.g. hospital, county, or state). This "differential" reporting can bias findings and misdirect policy makers (Kleinman, 1986). Other important data quality issues include inaccurate or unknown data elements (missing data) for cases that *are* reported (E. C. W. Gregory, Martin, Argov, & Osterman, 2019) . Missing information is more common among fetal and infant deaths than among live births (Gould, Chavez, Marks, & Liu, 2002), and more common among early gestation and low birth weight births and deaths.

The causes of data quality problems include variations in policies, procedures, and training. One cause of differential reporting is differing reporting requirements. Although all US states have long required reporting of all live births and all infant deaths, fetal death reporting requirements vary by state. Requirements are specified in terms of gestational age (GA) and birthweight (BW) (E. C. Gregory, Driscoll, Anne, 2014; Kowaleski, 1997). Fetal death reporting requirements may also affect reporting practices for very early births and infant deaths due to

misreporting of infant deaths as fetal deaths (Ramsay & Santella, 2011). Reporting practices are also affected by training of physicians and hospital staff, and burial requirements and costs (Lumley, 2003; Melnik, Guldal, Schoen, Alicandro, & Henfield, 2015). Infant deaths must be linked to their birth certificates so that birth information is available, and linkage rates still vary (NCHS, 2016).

The Perinatal Periods of Risk Approach (PPOR) is a widely used methodology for investigating and addressing high fetal and infant mortality rates that relies on vital records data (Peck, Sappenfield, & Skala, 2010). A brief overview of the analytic methods used in PPOR is provided in the introduction to Chapter 3. Many local health departments and their community partners use PPOR to inform and motivate social and health systems changes (Besculides & Laraque, 2005; Prince, Young, Sappenfield, & Parrish, 2016; Xaverius, Salas, Kiel, & Woolfolk, 2014). The approach includes analytic and community engagement methods and is designed for use in the context of other community assessments and health improvement efforts. Key strengths of PPOR's analytic methodology include the ability to estimate preventable mortality, the inclusion of fetal deaths, and the ability to use existing data to identify locally important causes and risk factors (CityMatCH, 2012; Sappenfield, Peck, Gilbert, Haynatzka, & Bryant, 2010a).

In PPOR analysis, communities estimate preventable mortality by subtracting rates in a reference population from corresponding rates in a study population. As with any comparison, PPOR estimates can be biased if there is differential reporting. To illustrate, we imagine a county with two large birthing hospitals. One serves a predominantly Black population and records fetal deaths at all GA, while the other serves a mostly White population and does not record most early GA fetal deaths. In this county, the Black/White gap would be artificially inflated.

To reduce bias in comparing rates internationally, the World Health Organization exclude fetal deaths at <28 weeks of gestation, the range in which reporting is essentially complete (Smith et al., 2018). Standard fetal death tabulations by the National Center for Health Statistics within the US exclude fetal deaths at GA <20 weeks. PPOR users are advised to exclude fetal deaths at GA < 24 weeks **and** BW <500 grams and exclude live births and infant deaths at BW<500 grams. However, all of these "cutoffs" exclude a large proportion of deaths, reducing the information available for understanding community health. Recent improvements such as fetal autopsy recommendations (Nijkamp et al., 2017) adoption of the 2003 birth certificate revision (Martin, 2014), and nearly universal electronic reporting systems (NAPHSIS, 2021; Westat, 2016) may have reduced differential reporting to the extent that lower PPOR cutoffs are justified.

Chapters 1 and 2 investigate the quality of recent (2012-2016) US vital records data, including fetal deaths, births, and linked infant deaths. We assess differential reporting of fetal and infant deaths by examining their distributions across states by week of gestation (from 20 to 31 weeks) and by birthweight (in 100-gram intervals below 1500 grams). We quantify the impact of underreporting on bias for general comparisons across US states. We estimate bias for different GA and BW cutoffs using interquartile range, with the 50 states plus Washington DC as units of analysis. We adjust for state-level characteristics using quantile regression.

Chapter 3 builds on the findings in the first two chapters and creates an updated national analysis. The reference population is mentioned above as key to estimating preventable mortality in the PPOR approach. In recent years, some communities wishing to use PPOR cannot create high-quality local or state reference groups because of data quality

problems, small numbers, administrative limitations, or high local rates. A national reference group has not been produced since the 2000-2002 data years, primarily due to data quality problems (Christiansen-Lindquist et al., 2017) that cause bias in reference group rates and estimates of preventable mortality. This chapter assesses the quality of data needed to create reference groups and estimate the resulting bias. We create a set of relatively unbiased reference populations by restricting the pool of states from which to select cases. A second and related problem addressed in Chapter 3 is that communities using PPOR are not aware of the findings of other communities. Findings from the national analysis will alert communities to unusual patterns, enabling them to better interpret their local findings.

Our Chapter 3 analysis introduces two innovations. First, we include six major components of excess mortality (six PPOR outcomes) rather than the four period-specific mortality rates presented when PPOR was developed. Second, we create percentile charts for the six components based on a study group of 100 large US counties. We use these charts to summarize national, county-level findings and describe patterns of absolute and excess (preventable) mortality. The percentile charts will also allow individual communities to easily compare themselves to US counties, by component.

CHAPTER 1. AN ASSESSMENT OF UNDER-REPORTING OF EARLY FETAL DEATHS AND ITS IMPACT ON BIAS OF RATE COMPARISONS IN THE US

CHAPTER 1 INTRODUCTION

The US Vital Records System records births, deaths, and fetal deaths for use in public health monitoring and to guide policy. Under-reporting of very early gestational age (GA) and low birthweight (BW) fetal deaths has been documented at state, county, and hospital levels (Ramsay & Santella, 2011). Underreporting of cases in some places but not in others (differential reporting), missing data elements, and misreported elements can bias comparisons and mislead policymakers.

The Perinatal Periods of Risk Approach (PPOR) is a widely used methodology for investigating and addressing high fetal and infant mortality rates that relies on vital records data (Sappenfield et al., 2010a). PPOR estimates preventable mortality by subtracting rates in a reference population from corresponding rates in a study population, and differential reporting can bias these estimates. Therefore, PPOR users are advised to exclude fetal deaths at less than 24 weeks gestation and exclude fetal deaths at less than 500 grams weight at delivery. This practice omits a large proportion of fetal deaths.

In this chapter, we assess differential reporting across US states graphically and quantitatively to determine whether these exclusion criteria (cutoffs) can be lowered so that more deaths can be included.

CHAPTER 1 METHODS

Vital records micro-data files were obtained from the National Center for Health Statistics (NCHS) by special request via the National Association for Public Health Statistics and Information Systems (NAPHSIS). For this study, live births, fetal deaths, and linked infant deaths

for the years 2012-2016 were used. IRB approval was obtained through the University of Nebraska Medical Center (**#** 667-18-EP). Analyses were completed using SAS 9.4.

Reporting Groups

We initially identified fetal death reporting requirements using tables in CDC "Guides to fetal death data" from 2002, 2006, 2014-2016 (E. C. Gregory, Driscoll, Anne, 2014), and a 1997 CDC Publication (Kowaleski, 1997). We accessed state websites to clarify, verify & correct the CDC information. We classified the 50 states plus Washington DC (here collectively referred to as "states") into four groups based on fetal death reporting requirements that were in effect during 2012-2016. Reporting requirements for two states changed during this period, and, assuming that practice changes are likely to lag behind reporting requirements, we classified those states according to their original requirements. We assumed that the group of states requiring reporting of all fetal deaths most closely represents the true joint BW and GA distribution of early fetal deaths, and used that group as the benchmark for estimating "coverage" of other reporting requirements.

Data Quality

In the national data files, GA is recorded by weeks of completed gestation from 2-47 weeks for fetal deaths and from 17-47 weeks for live births. BW is reported in grams from 0 to 8165 grams (18 pounds) for fetal deaths and from 227-8165 grams for live births. We assessed data quality for fetal deaths, calculating the percent of reported cases with unknown BW and unknown GA, nationwide and by state and reporting group. We used the NCHS-provided GA data element defined as obstetric estimate (OE) of gestation or last menstrual period (LMP) if OE is unknown. In states that require reporting of all fetal deaths, standard tables produced by the Centers for Disease Control and Prevention (the CDC) presume that cases with unknown GA

are under 20 weeks unless their BW is >= 350 grams. We followed this practice after verifying its appropriateness based on the joint distribution of BW and GA in this dataset.

Underreporting by Gestational Age

Assessment of underreporting (or under registration) is ideally done by comparing to a more complete data source, such as medical records or census data (Shapiro, 1950). The US Census does not record fetal deaths, and medical records are not available for nationwide analysis. Therefore, for fetal deaths, underreporting has been assessed by comparing GA and BW-specific mortality ratios for different reporting requirements and population groups (Goyal et al., 2017; Tyler et al., 2012; Wingate & Alexander, 2006) . If states adhered perfectly to their reporting requirements, then states requiring reporting of all fetal deaths, and states requiring reporting at 20 or more weeks, would have similar distributions at 20 or more weeks. Differences could be due to random variation, underlying health differences, and differences in reporting practices, i.e. “differential reporting”.

To estimate bias due to differential reporting by week of GA, we first described the distribution of fetal mortality across US states by week of GA from 20 through 31 . To do this, we calculated GA-specific fetal mortality ratios (FMR) as the number of fetal deaths at the given week of GA, per thousand live births plus fetal deaths, for each state. This denominator approximates the population at risk of becoming a fetal or infant death. Our analysis does not include elective abortions or pregnancies that were not recorded. We graphed the results using boxplots, assessing sensitivity to factors known to affect GA distribution by comparing these to boxplots for Black and singleton subpopulations. We displayed GA-specific FMR for each of the four reporting groups as line graphs.

We initially used Poisson confidence intervals around state GA-specific FMRs to assure that variation exceeds what can be attributed to “random” error. We then adjusted for

underlying health differences using quantile regression, which does not require an assumed distribution or homoscedasticity (Rodriguez & Yao, 2017). We used the QUANTSELECT procedure in SAS version 9.4, with stepwise model selection. With states as units of analysis, and GA-specific FMR as the dependent variable, we separately optimized models to estimate the 25^{th} and 75^{th} percentiles. Categorical independent variables considered by the selection algorithm were GA, US Census region, and state number of births (as quartiles). Continuous variables considered were state infant mortality rate, percent of the population that was uninsured (average of 2012-2016 ACS data), percentages of births that occurred in metropolitan areas, were paid for by Medicaid (2016 CDC Wonder), and were twins or higher order pregnancies, and the percentages of births to women with the following characteristics: a Bachelor's degree or more, aged 35 or older, aged 19 or younger, White non-Hispanic, Black non-Hispanic, and Hispanic. Interactions of each variable with GA were also considered. Continuous variables were centered on their national means.

We quantified the potential bias caused by differential reporting by comparing fetal mortality for a place that reports most fetal deaths (like the 75^{th} percentile state) to a place that reports relatively few (like the state at the 25^{th} percentile). We chose these percentiles to reduce the influence of states at reporting extremes, and because the difference (the Interquartile range, or IQR) is a standard measure of dispersion. This is meant to be a "typical" amount of bias for an inter-state comparison. We assumed complete reporting at >=28 weeks of gestation and that the average IQR at 28-31 weeks (b_{avg}) represents the baseline true variation in FMR among states. We estimated bias in comparing the two places by summing IQR *in excess* of the baseline $\Sigma(y_{75} - y_{25} - b_{avg})$ below 28 weeks. We varied the lower limit (cutoff), from the traditional NVSS tabulation limit of 20 weeks, up to the PPOR "cutoff" of 23 weeks. Finally, we re-estimated bias using IQRs adjusted according to the quantile regression models selected.

A decision to reduce the GA exclusion limit recommended for PPOR analyses from 24 weeks to 23 weeks (for example), required evidence that there was NOT substantive under-reporting at 23 or more weeks. Because we do not wish to further restrict communities from using data about their earliest fetal or infant deaths, we did not consider recommending even higher cutoffs that exclude more deaths.

Underreporting by Birthweight

To examine the variation among states by BW we did the same graphical analysis as for GA, categorizing birth weight in 100-gram intervals and using the same denominator. We imputed unknown BW based on known GA according to the median BW for that GA. The median was calculated based on all fetal deaths in our study population for which both BW and GA were known. We compared boxplots and line graphs with and without imputing unknown BW. We quantified the potential bias as for GA, with 1000-1499 grams as the range of assumed complete reporting and assessed exclusion criteria at 500 grams and below. We then applied the GA exclusion criteria and re-examined the need for a BW cutoff based on boxplots and line graphs. We did not use regression to adjust for health differences due to the high percentage imputed BW, which would cause underestimates of variance.

A decision to reduce the recommended BW exclusion limit required the same standard of evidence as for GA but would also depend on the BW limit needed to limit bias for infant deaths .

CHAPTER 1 RESULTS

During the five-year period 2012-2016, in the 50 states and Washington DC, there were 19,797,470 live births and 265,537 fetal deaths, of which 118,312 occurred at 20 weeks gestation or more. Using fetal deaths plus live births as a denominator, the overall FMR was 13.2. Excluding <20 weeks fetal deaths the FMR was 5.9, consistent with published summary data (Hoyert Dl Fau - Gregory & Gregory).

Reporting Groups

Reporting groups are summarized in Table 1. Five "Gold Standard" states (Georgia, Hawaii, New York, Rhode Island, and Virginia) require reporting of all fetal deaths and provided the benchmark GA and BW distribution. The most common fetal death reporting requirement, shared by 27 states, was 20 or more weeks. This "20-weeks" group included Oklahoma, which changed its limit from 20 to 12 weeks during 2015. Based on the benchmark distribution, the 20-week requirement would capture only 10.6% of all fetal deaths. Our third group included 14 states that required reporting of fetal deaths if they met either the 20 weeks GA or a BW requirement (350 to 500 grams). This "350 or 20" group would capture 10.7 to 11.0% of all fetal deaths. Our fourth group included five states with requirements based only on BW. They were the only states that did **not** require reporting of all fetal deaths at 20 or more weeks. Three states (Arkansas, Montana, and Delaware) reported at 350 grams or 20 weeks *if birth weight was unknown*, which would capture 9.4% of fetal deaths. Kansas and New Mexico required reporting at 350 and 500 grams, respectively (capturing 8.5% and 6.1% of fetal deaths) and both instituted the 20-week requirement during 2014.

Table 1. Summary of fetal death reporting groups, by US Census Region.

States, numbers of fetal deaths, and marginal distribution of fetal deaths are listed. Fetal deaths at <20 weeks gestation are excluded.

	Census Region				
Fetal Death Reporting Group	Northeast	Southeast	Midwest	West	Total
"Gold Standard" Group (5 states) reporting all fetal deaths.	NY, RI N=8,212	GA, VA N=8,162		HI N=516	16,890 (14%)
"20-weeks" Group (27 states), reporting fetal deaths at 20 or more weeks of gestation.	CT, ME, NJ, PA, N=8,940	AL, FL, MD, NC, OK, TX, WV N=28,803	IA, IL, IN, MN, ND, NE, OH, SD N=16,019	AK, CA, CO, NV, OR, UT, WA, WY N=21,313	75,075 (63%)
"350 or 20" group (13 states + DC), reporting if 20 weeks OR 350 (or 400) grams.	MA, NH, VT N=2,026	DC, KY, LA, MS, SC, TN N=11,287	MI, MO, WI N=6,814	AZ, ID N=3,028	23,155 (20%)
"Grams only" group (5 states), reporting based on BW (or GA if BW is unknown).		AR, DE N=1,484	KS N=1,061	MT, NM N=647	3,192 (3%)
Total	19,178 (16%)	49,736 (42%)	23,894 (20%)	25,504 (22%)	118,312 fetal deaths

Data Quality

Overall, 5.0% of fetal deaths had unknown GA, 52.6% had unknown BW, and 4.6% had both unknown. Employing the CDC presumption described above reduced the percent of fetal deaths with unknown GA to .3% overall. Validity of the CDC presumption is supported by (1) among Gold group fetal deaths, 88% of those with known GA were delivered at less than 20 weeks, and (2) 77% of deaths with BWs below 350 grams had GA below 20 weeks. Unknown BW varied by state and reporting group and was more common among early GA fetal deaths and Gold group states (Table 2). After excluding fetal deaths with known or presumed gestation of less than 24 weeks, 5.5% had unknown BW (16.1% in the Gold Group), and 17 states had more than 5% of cases with unknown BW. Imputing unknown BW from known GA leaves 4.6% with unknown BW overall (7.6% in the Gold group), two states with more than 10% missing (GA, HI), and one state with more than 5% missing (VA).

Table 2 Percent of fetal deaths with unknown birthweight, by gestational age, among all fetal deaths, and among the group of "Gold" states that require reporting of all fetal deaths.

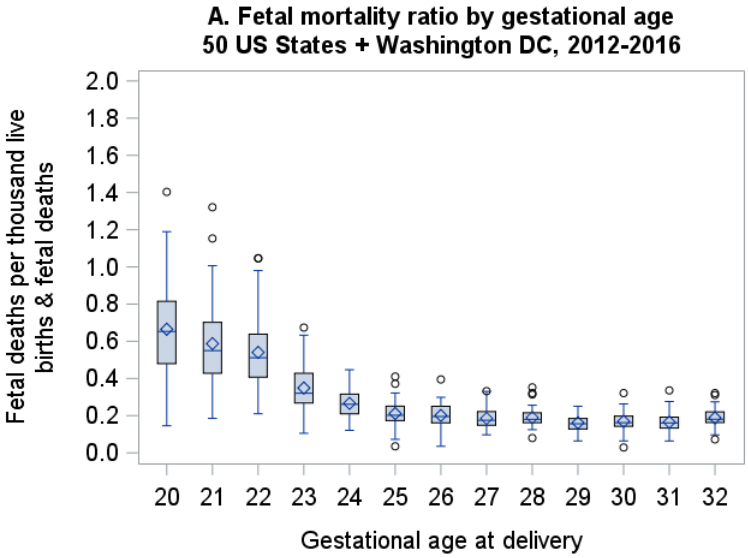
% with unknown birthweight	Gestational age at delivery				
	0 to 19 Weeks (known or presumed)	20-23 Weeks	24-31 Weeks	32 or more weeks	Total fetal deaths missing BW
Among all fetal deaths	89.4%	9.9%	6.5%	4.7%	52.6%
Among Gold states	91.7%	23.3%	18.6%	14.1%	83.9%

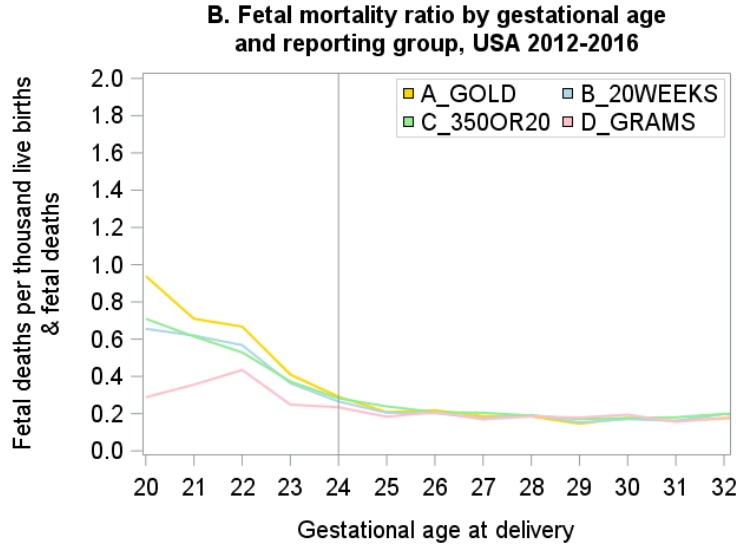
Underreporting by Gestational Age

State FMRs are shown as boxplots in Figure 1, by GA. In general, FMRs are constant from 31 down to 25 weeks of gestation, then increase with decreasing gestation. Of concern is the biologically implausible finding that fetal mortality appears to increase sharply for some states but not for others. Poisson confidence intervals (not shown) indicated that many differences between states are significant. For example, at 23 weeks, the confidence interval for the state with the highest rate does not overlap with confidence intervals for 37 other states. Figure 1, on the right, shows FMRs by fetal death reporting group and provides evidence that differences in reporting *requirements* contribute substantially to the observed variation at early GAs. The Gold group (the yellow line) begins to diverge from the other groups at about 24 weeks, and diverges further with decreasing GA. Below 26 weeks, FMRs in the "grams only" reporting group (the pink line) are lower than other groups. The "20-week" and "350 or 20" groups, which together account for 83% of fetal deaths (see Table 1), are similar throughout, however their ratios are lower than "Gold" ratios, indicating they may not be reporting all fetal deaths at 20 or more weeks of gestation. The differences between reporting groups are substantive: overall FMRs in the Gold standard and 20-week groups were 6.7 and 5.8, respectively: a difference of .9 deaths per thousand live births and fetal deaths. The components due to deaths below 24 weeks were 2.9, and 2.2, a difference of .7. This means the <24-week fetal deaths contribute 77% of the overall difference between reporting groups. These patterns persist for lower risk singletons and higher risk African Americans (figures not shown). A “bump” at 22 weeks may reflect differing beliefs regarding viability in that period (Allen, Donohue Pk Fau - Dusman, & Dusman; El-Metwally, Vohr B Fau - Tucker, & Tucker; Ramsay & Santella, 2011).

Figure 1 Fetal Mortality Ratios by Gestational Age USA 2012-2016

In the boxplot graph above, the unit of analysis is US States and the outcome is fetal deaths at the given GA per 1,000 live births plus fetal deaths. The heights of the boxes represent gestational age-specific interquartile range (IQR) across states; whiskers show minimum and maximum states; means are diamonds, and outliers are open circles. On the right, the same data is plotted by fetal death reporting group. Variation among states increases substantively below 24 weeks gestation, and much of this appears to be due to differences in reporting.

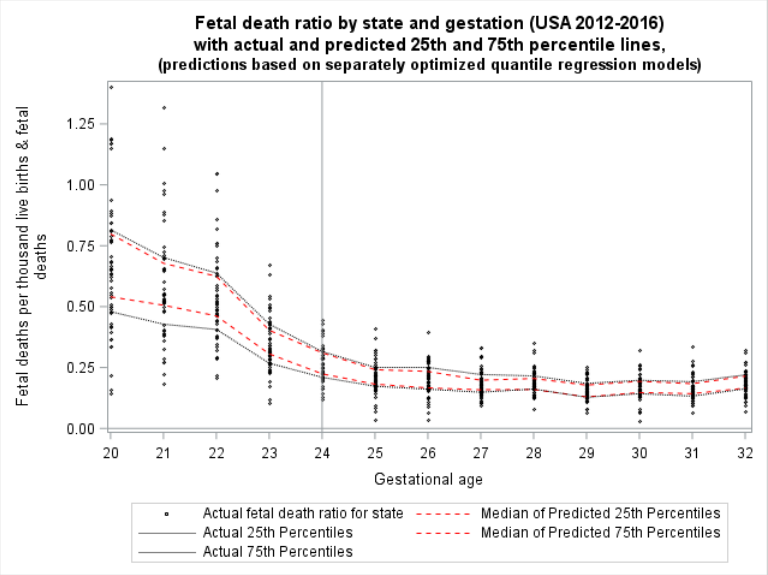




Adjusting for health differences (see Figure 2) attenuated the IQR but preserved the pattern of increasing variation with decreasing gestation. Also preserved is the fact that IQR begins to increase above the PPOR 24-week limit, not below, as would be required to reduce the limit. The dashed red lines in Figure 2 are adjusted estimates for the 25th and 75th quantiles. Variables selected by SAS PROC QUANTSELECT to predict the 25th percentile included state infant mortality rate, % Black, and %Black interaction with GA. The model for the 75th percentile included those, plus US Census Region, %Metropolitan, %Medicaid, and %multiple pregnancies. The adjusted R1 statistics, analogous to R-squared in least-squares regression, were .42 (for the 25th percentile) and .70 (for the 75th), indicating that the models were predictive but did not fully account for variation in mortality.

Figure 2 FMR by state and gestational age, crude and adjusted

Actual (black, solid lines) and model-adjusted (red, dashed) 25th and 75th percentiles. The models included GA as indicator (class) variables. Variation among states increases with decreasing GA, and that this effect persists although it is somewhat attenuated when adjusted for health-related factors.



Estimated bias due to variation across states is shown in Table 3 for various GA exclusion criteria. The baseline IQR at 28-31 weeks, assumed to represent true health differences, was .056. The estimated bias, the sum of the IQRs from 20 to 27 weeks *in excess of* the baseline, is 0.90, which is 15% of the national FMR. Raising the cutoff decreases this bias. The 24-week cutoff recommended for PPOR decreased the bias to .12, 2% of FMR. Adjustment for health-related factors reduced the estimated bias for the 20-week cutoff to .62 (11% of the national FMR), and for the 24-week cutoff to .08 (1% of the national FMR).

Table 3. Estimated bias in comparing fetal mortality in two places reporting at the 75th and 25th percentiles, respectively, under varying exclusion criteria.

The top row uses the NCHS practice of including fetal deaths at 20 or more weeks gestation. The bottom row uses the original PPOR recommendation, excluding <24-weeks cases. We assumed universal complete reporting at 28 weeks and above and subtracted the baseline average difference at 28-31 weeks. Using crude rates and including 20 or more weeks, the hypothetical comparison would be biased by .9, a substantive 15% of the national fetal mortality rate (5.9). Adjusting for health differences would reduce the bias to 11%. The PPOR 24-week cutoff would further reduce the estimated bias to .08, 1% of the national rate.

	Crude		Adjusted	
GA range	IQR minus Baseline	% of 5.9	IQR - Baseline	% of 5.9
20-27	0.90	15%	0.62	11%
21-27	0.62	10%	0.42	7%
22-27	0.40	7%	0.26	4%
23-27	0.22	4%	0.15	2%
24-27	0.12	2%	0.08	1%

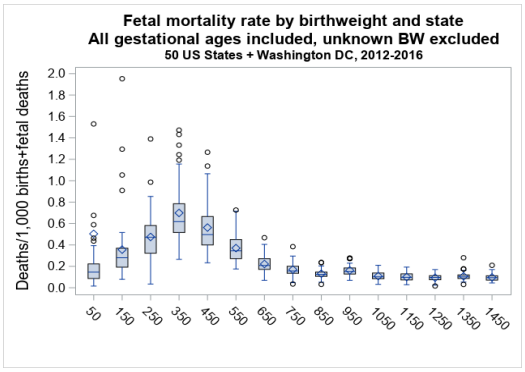
Underreporting by Birthweight

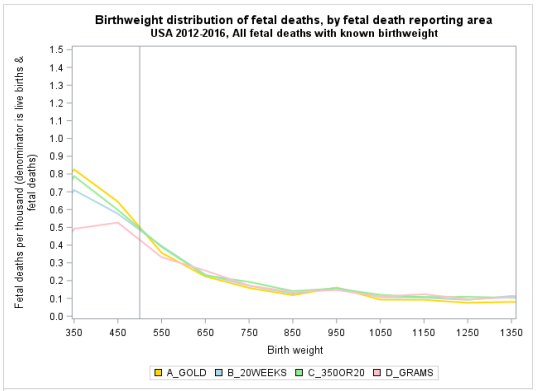
The high percentage of fetal death cases with unknown BW (Table 2) is a potential limitation for assessing underreporting by BW, so this analysis is repeated with BWs imputed as described in the methods section. Figure 3 shows boxplots for states and line graphs for reporting groups analogous to Figure 1. The top row excludes cases with unknown BW, while the bottom row imputes them. As with GA, the variation in state FMRs increases with decreasing BW, beginning in the 600 to 699-gram category. Reporting groups also begin to diverge below 600 grams. Adding imputed BW accentuates these effects. By reporting group, the "Gold" states diverge from other states more sharply as BW decreases. Imputing BW primarily affects <500-gram cases, since 97% of cases with unknown BW have GA<24 weeks (actual or presumed).

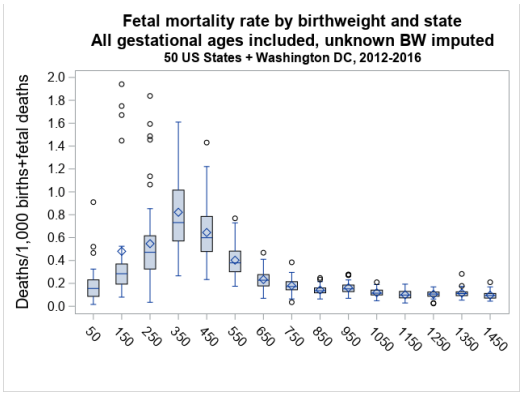
Figure 3 Fetal mortality ratios by state and reporting group

B nts.

Graphs on the top row exclude cases with unknown BW; graphs on the bottom row impute them.







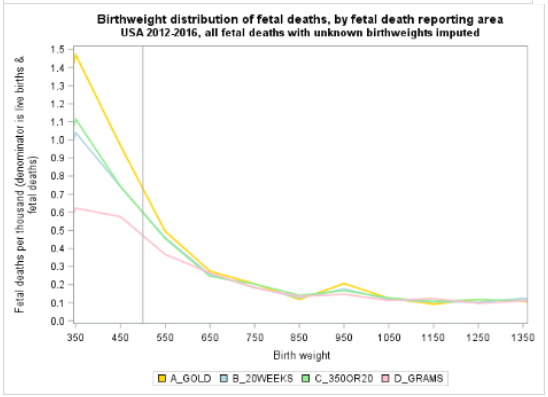
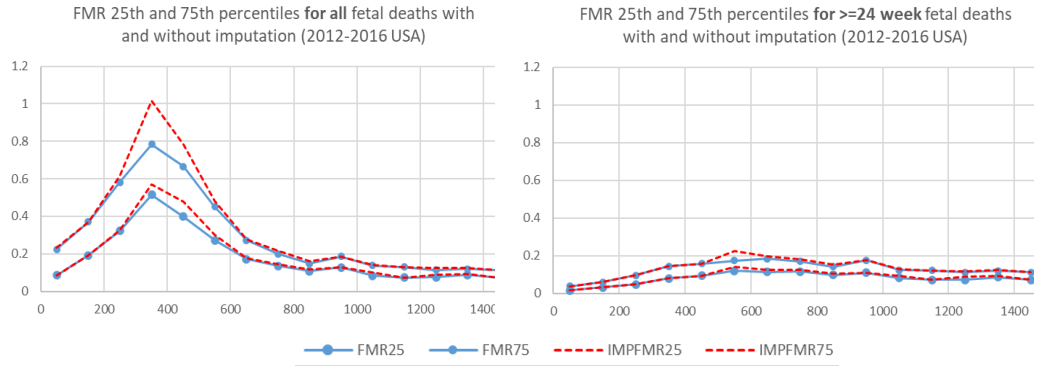


Figure 4 State 25th and 75th percentiles of fetal mortality ratios by birthweight, USA 2012-2016.

On the left, all fetal deaths are included. The blue solid line is FMRs with unknown birthweights excluded. On the right, fetal deaths under 24 weeks are excluded. Dashed red lines are FMRs with unknown birthweights imputed from known gestational age included. The dashed are percentiles with and without imputation with all fetal deaths included. The outcome is fetal deaths at the given BW range per 1,000 live births plus fetal deaths.



We then excluded fetal deaths at less than 24 weeks gestation based on our conclusions from Table 3. Figure 4 shows the effect of the GA exclusion, using only percentile lines for simplicity. The graph on the left includes all fetal deaths and is the same shape as the box plots in Figure 3. The graph on the right excludes the <24 weeks GA fetal deaths, and we see that both the effects of reporting and of imputing missing BW are greatly reduced.

Bias estimates based on the graphs in Figure 3 are shown in Table 4, with different BW exclusion criteria (analogous to Figure 1 and Table 3 by GA). Without the GA exclusion, bias is substantive but decreases to .24 as the exclusion limit is raised to the PPOR-recommended 500 grams. When <24 weeks fetal deaths are excluded, the bias is reduced even further, to .17. In fact, only an additional 1% of fetal deaths weigh less than 500 grams and would be affected by the 500-gram BW cutoff. As expected from Figure 3, excluding <24-week fetal deaths makes the four reporting groups almost indistinguishable (figure not shown). Thus, we conclude that a lower BW limit is not necessary if a GA exclusion is in place. However, communities may wish to use the 500-gram BW exclusion for consistency with infant death restrictions, which will likely not exclude many additional deaths. As expected from Figure 3, excluding <24-week fetal deaths makes the four reporting groups almost indistinguishable (figure not shown).

Table 4 Estimated bias in comparing fetal mortality ratios in places reporting at the 75th and 25th percentiles, under varying exclusion criteria.

Excluding The top row represents including all fetal deaths and would bias comparisons (differences) between the two places by 1.39 fetal deaths per thousand live births and fetal deaths. The bottom row represents the original PPOR recommendation to exclude deaths of infant born at less than 500 grams. This exclusion reduced the bias to .24. Omitting fetal deaths at less than 24 weeks reduced the maximum bias to .17, more than excluding <500 grams.

	All fetal deaths; unknown BW imputed		<24 weeks excluded; unknown BW imputed	
Lower limit	IQR - Baseline	% of 5.9	IQR - Baseline	% of 5.9
0 grams	1.39	24%	0.17	3%
100 grams	1.29	22%	0.19	3%
200 grams	1.15	20%	0.20	3%
300 grams	0.91	15%	0.19	3%
400 grams	0.50	9%	0.16	3%
500 grams	0.24	4%	0.17	2%

CHAPTER 1 DISCUSSION

This study updates and expands on the original studies that led to the published PPOR exclusion criteria. (Peck et al., 2010). We add information about the impact of reporting requirements and adjusting for health effects that contribute to variation in GA- and BW-specific FMRs. We introduce simple methods for comparing reporting requirements and for quantifying bias in FIMR differences to facilitate evaluation of exclusion criteria.

This analysis confirms previous findings that state reporting requirements affect FMRs, even in GA and BW ranges where requirements agree (Goyal et al., 2017; Kirmeyer, 2006). Reporting requirements are specified to capture (cover), different percentages of fetal deaths. Compared to states that require reporting of all fetal deaths, other reporting requirements are designed to capture only 6.1% to 11.0% of fetal deaths.

Our study found strong evidence of underreporting. Even at >=20 weeks gestation, where all but 5 states required reporting, variation across states indicates that requirements are not strictly followed. Similar to what has been previously noted (Kirmeyer, 2006), there is a lag between requirements and practices. States that require reporting of all fetal deaths appear not to report some in the 20-21 week range, and states that require reporting at GA>=20 weeks do not appear to match Gold state reporting levels until 24 weeks or later. These differences are substantive and will bias comparisons, including the estimation of excess mortality in PPOR analysis. The 24-week minimum GA recommended for PPOR analysis reduces this bias to about 1% of the national >20 weeks FMR.

We found that the fetal death BW cutoff is unnecessary if the 24-week GA cutoff is employed, although communities may wish to continue using the BW cutoff for consistency with infant deaths.

Limitations of this study include data quality problems, especially missing BW and GA. We addressed these by presuming that cases with unknown GA in states that report all fetal

deaths are <20 weeks, substituting median BW for GA for cases with unknown BW, and excluding fetal deaths at <24 weeks. A second limitation is that our analysis was ecological and at the state level. Comparisons within a state may be less biased because of common reporting requirements and collection/verification procedures or may be more biased due to county-level (Williams & Magsumbol, 2010), and hospital-level (Ramsay & Santella, 2011) differences.in reporting practices. Our study lacked information about state vital records follow-up, data cleaning, and & verification practices, which undoubtedly affect completeness of reporting.

In conclusion, we recommend excluding fetal deaths at <24 weeks gestation to reduce bias. Excluding fetal deaths at <500 grams excludes few additional cases and provides consistency with infant death BW exclusion criteria. Because known sources of variation operate at county, hospital, and physician levels, it is likely that differential reporting also exists within states.

CHAPTER 2 AN ASSESSMENT OF UNDER-REPORTING OF EXTREMELY LOW BIRTHWEIGHT INFANT DEATHS AND ITS IMPACT ON BIAS OF RATE COMPARISONS IN THE US

CHAPTER 2 INTRODUCTION

Records of US births, deaths, and fetal deaths are collected by local and state governments and used to monitor population health and guide health policy. Reporting problems have been documented in the US, especially for very low birthweight (VLBW) infant deaths. One such problem is underreporting, in which some infant deaths are not reported at all, which can bias findings and misdirect policy makers (Kleinman, 1986). Some states have standard processes for following up on births of very small infants to ensure that a death certificate is filed if appropriate (McCarthy et al., 1980). Misreporting of infant deaths as fetal deaths has also been documented (Ramsay & Santella, 2011). Among reported infant deaths, missing and incorrect information is more common than among live births (Gould et al., 2002), which can affect infant mortality rates for subgroups. Although all US states have long required reporting of all live births and all infant deaths, fetal death reporting requirements vary by state and are specified in terms of gestational age (GA) and birthweight (BW) (E. C. Gregory, Driscoll, Anne, 2014; Kowaleski, 1997). Fetal death reporting requirements may affect reporting practices for very early births and infant deaths. Reporting practices may also depend on training of physicians and hospital staff, and burial requirements and costs (Lumley, 2003; Melnik et al., 2015).

The Perinatal Periods of Risk Approach (PPOR) is a widely used methodology that relies on vital records data for investigating and addressing high fetal and infant mortality rates. In PPOR analysis, communities estimate preventable mortality by subtracting rates in a reference population from corresponding rates in a study population. Like other comparisons, PPOR estimates can be biased if there is differential reporting. To avoid bias, PPOR guidelines advise omitting births and infant deaths at BWs of less than 500 grams (Sappenfield et al., 2010a).

Recent improvements such as adoption of the 2003 birth certificate revision (Martin, 2014), and nearly universal electronic reporting systems (NAPHSIS, 2021; Westat, 2016) may have reduced differential reporting to the extent that lower PPOR cutoffs are justified.

This chapter investigates the quality of recent (2012-2016) US vital records data, including births and linked infant deaths. It estimates underreporting of very small infant deaths and its impact on bias for comparisons across US states.

CHAPTER 2 METHODS

Vital records micro-data files were obtained from the National Center for Health Statistics by special request via the National Association for Public Health Statistics and Information Systems (NAPHSIS). Live births and fetal deaths occurring during the years 2012-2016, and period linked infant deaths occurring during the same years, were used for this analysis. IRB approval was obtained through the University of Nebraska Medical Center (**#** 667-18-EP). Analyses were completed using SAS 9.4 and Microsoft Excel. The American Community Survey was the source of population un-insured rates for states, and CDC Wonder was the source for state percent of births paid by Medicaid. CDC "Guides to fetal death data" (E. C. Gregory, Driscoll, Anne, 2014) and state and legal websites were the source of information about state reporting requirements. States were assigned to one of four groups based on BW and GA reporting criteria as described elsewhere . The groups were five "Gold Standard" states that require reporting of *all* fetal deaths, two groups of states that require reporting at 20 or more weeks, and five states that required reporting based only on BW.

Data quality

In the national data files, GA is recorded by weeks of completed gestation from 2-47 weeks for fetal deaths and from 17-47 weeks for live births. BW is reported in grams from 0 to 8165 grams (18 pounds) for fetal deaths and from 227 (1/2 pound) to 8165 grams for live births. We used the detailed BW variable provided in the NCHS dataset, and GA variable provided in the NCHS datasets that uses obstetric estimate of GA, or last menstrual period (LMP) if the obstetric estimate is unknown. We categorized BW from 227 to 299, then in 100- gram intervals, 300-399, 400-499 etc. In graphs we plotted the categories by their center point. The actual center point for the lowest category is 263, but we use 250 in graphs for simplicity. We categorized states according to their fetal death reporting requirements (often laws), as described elsewhere . We assessed data quality for live births and infant deaths, calculating the

percent of reported cases with unknown BW and GA, nationwide and by state and fetal death reporting group.

Underreporting

We examined underreporting of infant deaths in two ways. First, to compare risk of death for different BWs, we calculated BW-specific infant mortality rates (IMR) as the number of deaths among infants born in the given BW category per thousand of births in that category. Second, to describe the impact of underreporting on infant mortality rates, we calculated BW-specific infant death ratios (IDR) using the same numerator, and a denominator consisting of all live births. The IDRs describe the distribution of infant deaths across BWs, and the sum of IDRs across all BW categories is the standard infant mortality rate, the total number of infant deaths divided by the total number of births in the same year.

Underreporting – Birthweight-specific mortality

We calculated and graphed BW-specific IMR for the nation and the four reporting groups. The probability of death is known to increase with decreasing BW. Because very few infants born weighing less than 400 grams have survived their first year, the IMR in those BW categories should be approximately 1000 deaths per 1000 live births (Bell & Zumbach, 2011). We graphed BW-specific IMRs for the 50 states plus Washington DC as a line graph, then as a boxplot. We calculated the percent of states that reported survival rates of 5% or more, and the survival rates reported by the 25% of states with the lowest IMRs at <500 grams.

Underreporting – Birthweight distribution

We created line graphs of BW-specific IDRs to examine the BW distribution of infant deaths for the nation, the four fetal death reporting groups, and the 50 states plus Washington DC. We added boxplots for the states to better display their distribution for each BW category. The upper and lower limits of box plot "boxes" are the 25th and 75th percentiles among the states, so their difference is the IQR, a common measure of dispersion. We hypothesized that

dispersion would increase with decreasing BW if some states continue to report all infant deaths, while others report fewer at lower birthweights.

Because studies have documented misreporting of infant deaths as fetal deaths (Ramsay & Santella, 2011), we investigated the impact of fetal death on the IDR curves. If infants that die are recorded as fetal deaths this could explain apparent underreporting of small infant deaths. To do this, we created a second national BW curve of FIMR by adding fetal deaths >=20 weeks gestation to both the numerator and the denominator of the national IMRs. Very few states report fetal deaths at <20 weeks gestation, and national fetal mortality rates are based on >=20 weeks. From the CDC infant death user guides we estimated the number of unlinked deaths in the neonatal category, and estimated the potential increase in IDR due to these, nationwide, by BW.

Because underlying health differences among states might lead to differences in BW distribution, we adjusted for health-related factors at the state level using quantile regression. This method does not require an assumed distribution or homoscedasticity (Rodriguez & Yao, 2017). We used the QUANTSELECT procedure in SAS version 9.4, with stepwise model selection. With states as units of analysis, and BW-specific IMR as the dependent variable, we separately optimized models to estimate the 25^{th} and 75^{th} percentiles. Categorical independent variables considered by the selection algorithm were BW category (forced in), US Census region, fetal death reporting group, and state number of births (as quartiles). Continuous variables considered were percent of the population that was uninsured (average of 2012-2016 ACS data), percentages of births that occurred in metropolitan areas, were paid for by Medicaid (2016 CDC Wonder), and that were twins or higher order pregnancies, and the percentages of births to women with the following characteristics: a Bachelor's degree or more, aged 35 or older, aged 19 or younger, White non-Hispanic, Black non-Hispanic, and Hispanic. Interactions of

each variable with BW category were also considered. Continuous variables were centered on their national means. We graphed 25th and 75th percentiles predicted by the selected models and overlaid the actual 25th and 75th percentiles.

Estimation of Bias

To quantify the potential bias caused by differential reporting at each BW category, we compared IDRs for a hypothetical place that reports most infant deaths (like the state at the 75th percentile) to another hypothetical place that reports relatively few (like the state at the 25th percentile). This measure is designed to estimate a "typical" amount of bias for an inter-state comparison. To assess the whether the PPOR exclusion criteria of <500 grams could be lowered without substantively biasing comparisons, we compared overall bias using different exclusion criteria. We assumed complete reporting at >=1000 grams and that the average IQR at 1000-1399 grams (b_{avg}) represents the baseline true variation in IDR among states. We estimated bias in comparing the two hypothetical places by summing IDRs *in excess* of the baseline $\Sigma(y_{75} - y_{25} - b_{avg})$ from the exclusion cutoff up to 999 grams. We tested exclusion cutoffs of 500, 400, 300, and 227 grams. Finally, we re-estimated bias using IQRs adjusted according to the predicted values of the selected quantile regression models.

A decision to reduce the BW exclusion limit recommended for PPOR analyses below 500 grams required evidence that there was NOT substantive under-reporting at 400-499 grams (for example). Because we do not wish to further restrict the use of data by communities, we did not consider recommending even higher cutoffs that would exclude more deaths.

CHAPTER 2 RESULTS

During the five-year period 2012-2016, in the 50 states and Washington DC, there were 19,797,470 live births, 116,059 of whom died before one year of age, for an average IMR of 5.9. The original PPOR cutoff of 500 grams would exclude 27,721 infant deaths (23.9%) from our analysis. Lowering the limit to 400 would mean excluding 14.6%, a difference of 9.3 percentage points and an additional 10,778 infant deaths nationwide. Reducing the limit to 300 would add another 10,582 deaths. There were 265,537 fetal deaths, of which 118,312 occurred at 20 weeks gestation or more.

Data quality

Birth weight was unknown for .03% of live births, with no more than 0.18% in any state. Infant deaths, the subset of live births who die within their first year of life, had poorer data quality than live births: 1.02% had unknown GA and .58% had unknown BW. No state had more than 2.81% of infant deaths with unknown GA or 2.25% with unknown BW. For fetal deaths, and 5.0% had missing GA. Presuming that, in states that report all fetal deaths, cases with missing GA are <20 weeks, unknown GA was reduced to 0.3%. Although it did not affect our analysis, 52.6% of fetal deaths had missing BW, and excluding known or presumed <20 weeks GA reduced this percentage to 7.2%. There were 868 unlinked deaths in the five-year period we studied, and an estimated 80% of those were neonatal. Assuming these are distributed equally across BW categories below 1500 grams, they would add about 53 deaths to each category.

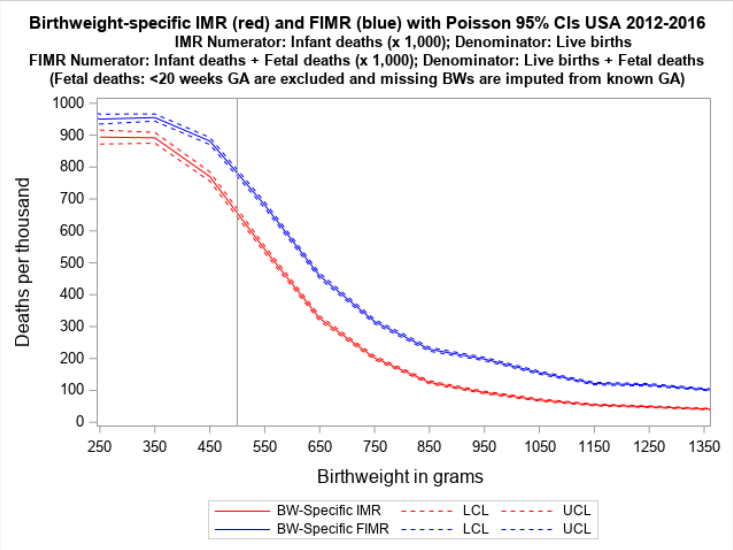
Underreporting – Birthweight-specific mortality

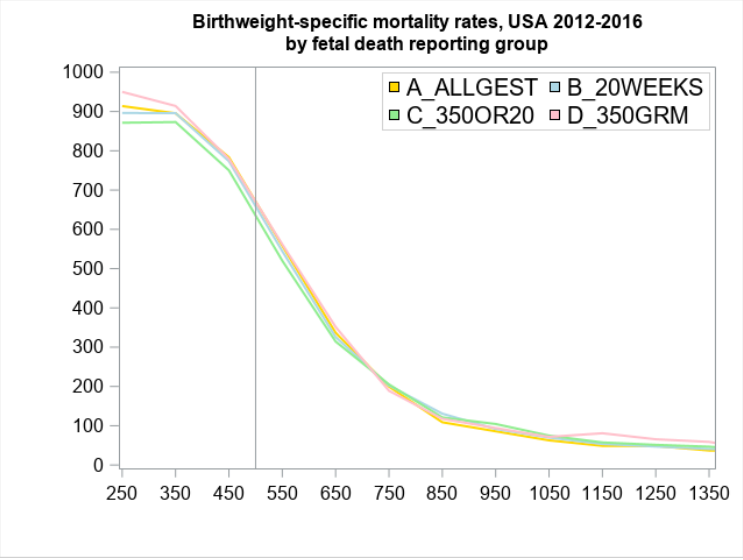
Risk of death for infants generally increases with decreasing BW as expected. However, as shown in Figure 5, the rate of increase slows between the 450 midpoint (400-499 grams) and the 350 midpoint (300-399 grams) and plateaus BW 227-399at an IMR of about 900 (90% mortality) for BW 227-399. This would mean that 10% of infants born at less than 400 grams

survive their first year, which is not plausible. Including fetal deaths in both numerator and denominator raises the plateau to about 95%, which may mean that misreporting of some infant deaths as fetal deaths could account for some unreported infant deaths. Actual mortality of live births below 400 grams is known to be close to 100%. We estimated that unlinked infant deaths could add 4 to 8 deaths per thousand in the BW range of interest, not enough to account for the gap. Below, in Figure 5, we see that differences between reporting groups are small in general but widen slightly below 400 grams. The group of five states with reporting requirements based only on BW had the highest rates below 400.

Figure 5 Birthweight-specific IMR nationwide and by fetal death reporting group.

Above is the curve of national rates (red, below) with the FIMR curve (blue, above). IMR is deaths of infants born in the given birthweight category, divided by live births in the given birthweight category. FIMR adds fetal deaths at 20 or more weeks gestation to both numerator and denominator of the IMR. Below are IMR curves for each fetal death reporting requirement.

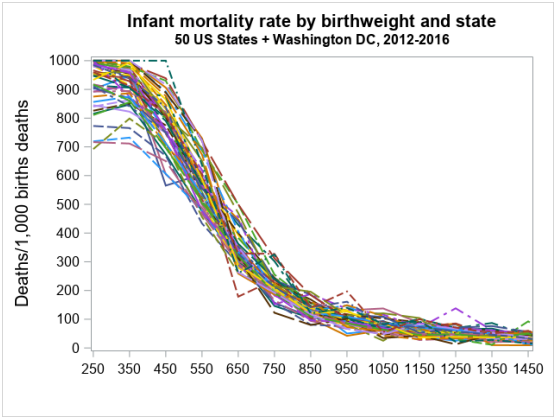


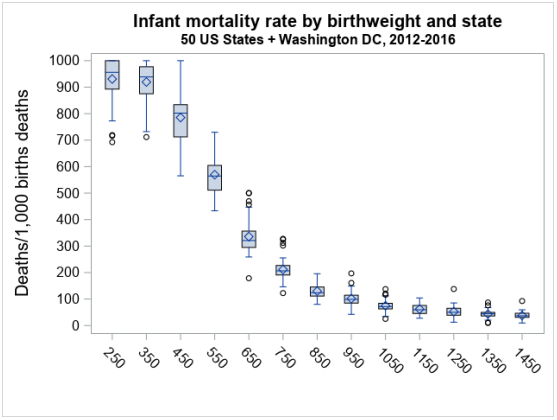


In Figure 6, we can see that the IMR curves in the 300-399 range for *four states* are below 800, implausibly implying that more than 20% of their <400-gram births survive their first year. There are 31 states with mortality rates below 900 in the 300-399 range, and 13 in the 200-299 range. Also notable is that variation or dispersion across states increases with decreasing BW, beginning with the 600-699 category, which we explore further via IDR in the next section. In summary, Figures 5 and 6 show strong evidence of underreporting below 400 grams, but are inconclusive for the 400-499 category.

Figure 6 Birthweight-specific IMR by state.

On the left is a line graph for each state, and on the right is a boxplot depicting the distribution across states at each birthweight category. Variation among states also increases with decreasing birthweight.



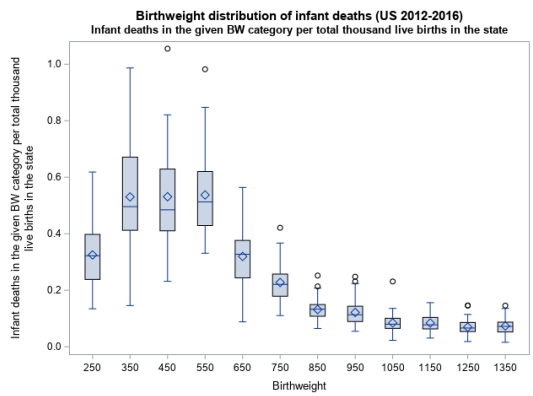


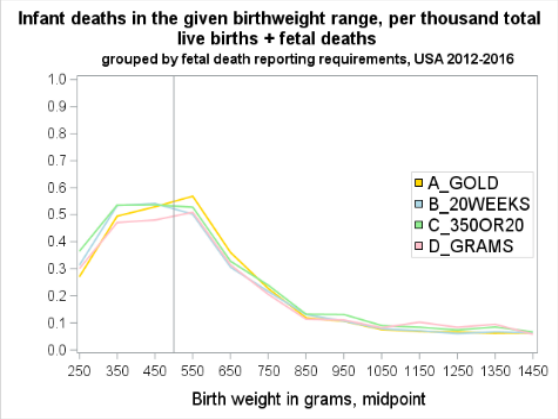
Underreporting -- Birthweight distribution

To assess variation among states we use infant death ratios, which are components of the overall mortality rate. In Figure 7, the graph on the left shows increasing ratios with decreasing BW until the ratio levels off below 600 grams. Note that the sharp decrease at the far left is at least partly due to this being a smaller range (227-299 grams) and is not the focus of our investigation. Of concern is the increase in *variation* among states as BW decreases. On the right in Figure 7, we see that fetal death reporting requirements may account for some of this variation as "Gold states" and states with only a BW requirement diverge from the majority of states that have a 20-weeks requirement. We estimate that unlinked deaths would add no more than .003 deaths per thousand to any category.

Figure 7 Birthweight distribution of the infant mortality rate.

On the left are boxplots of Infant Death Ratios by birthweight for the 50 US states plus Washington DC shows variation across states increasing below 800 grams. On the right is the same data by fetal death reporting group, showing only a slight increase in variation with decreasing birthweight.

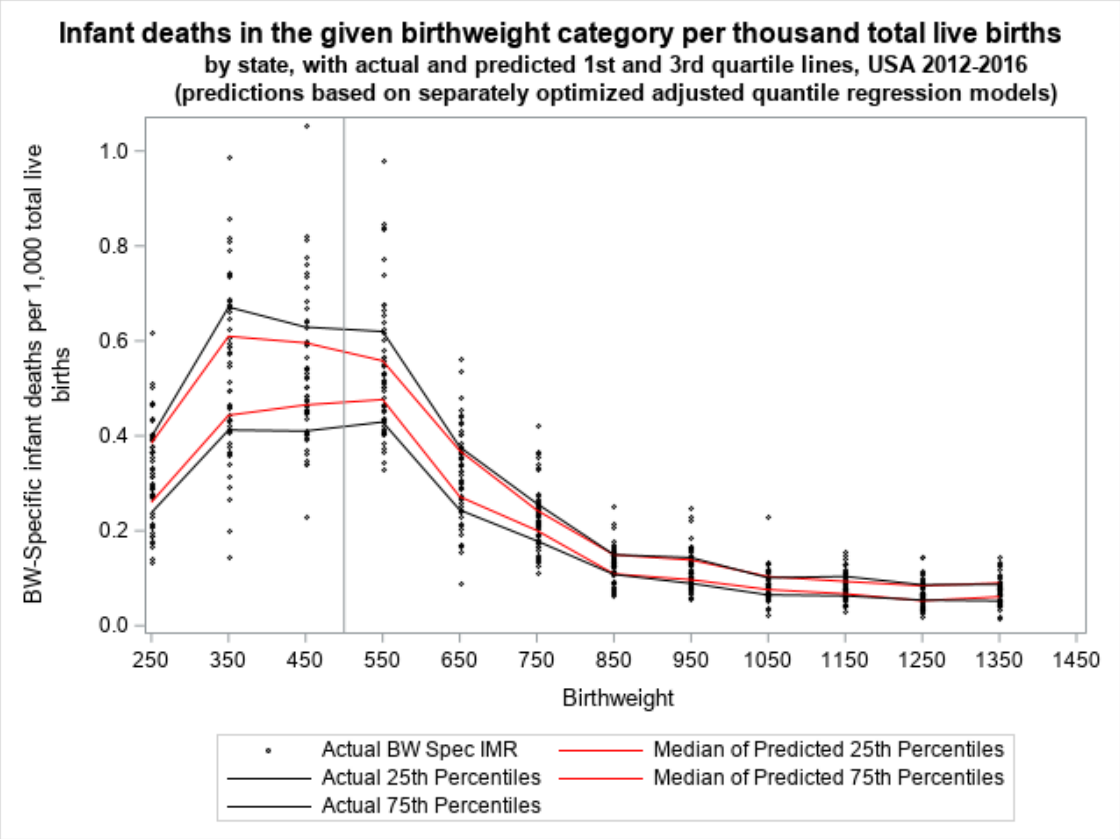




Results of adjusting for health differences among states using Quantile regression modeling are shown in Figure 8. Variation among states increases with decreasing GA, and that this effect persists when adjusted for health differences. Adjustment decreases the IQR below 800 grams but has little effect above 800 grams. The selected model for the 25th percentile included BW category, the state percent of mothers aged less than 20, the state percent of mothers who were Black non-Hispanic, and the interaction between BW category and the percent of mothers who were Black. The selected model for the 75th percentile included those, plus the number of births in the state (as quartiles), and the percent uninsured. The adjusted R1 values were .66 and .75, respectively, indicating that the models explained much, but not all, of the variation in the data.

Figure 8 Variation across states by birthweight, using crude and adjusted IQR.

Each point is a US state; the black lines are the actual 25th and 75th percentiles and their shape matches the shape of the boxplots in Figure 3. The red lines are predicted 25th and 75th percentiles based on quantile regression models.



Estimation of Bias

Bias estimates are shown in Table 5 based on crude and adjusted IQRs for varying exclusion criteria or "cutoffs". The estimated bias is the sum of the IQRs from 227 to 999 grams *in excess of the baseline*. With no exclusion the bias estimate is 0.85, 14% of the national IMR. Raising the cutoff decreases this bias. For the 500- gram cutoff used in PPOR, the bias is .32, 5% of the IMR. Adjusting for health reduces the bias estimate to .18, or 3% of the national IMR. Bias disappears with an 800-gram cutoff. We conclude that there is strong evidence of underreporting below 400 grams, and that differential reporting across states continues up to 800 grams. Omitting 400 to 499-gram infants further reduces but does not eliminate bias.

Table 5 Estimated bias in comparing infant mortality ratios in places reporting at the 75th and 25th percentiles, using crude and adjusted Interquartile range, under varying exclusion criteria.

	Crude bias estimate based on actual IDRs (baseline .036)		Regression-adjusted estimates based on model predicted IDRs (baseline .029)	
Lower limit	IQR - Baseline	% of 5.9	IQR - Baseline	% of 5.9
227 grams	0.85	14%	0.50	8%
300 grams	0.72	12%	0.40	7%
400 grams	0.50	8%	0.26	4%
500 grams	0.32	5%	0.16	3%
600 grams	0.16	3%	0.11	2%
700 grams	0.07	1%	0.04	1%
800 grams	0.02	0%	0.02	0%

CHAPTER 2 DISCUSSION

Lower than expected BW-specific mortality rates at less than 400 grams strongly support excluding deaths delivered at less than 400 grams to reduce bias due to underreporting. Variation of BW distribution across states would support an even higher cutoff. Our bias analysis indicated that reporting differences are substantive above 400 grams. However, there is a tradeoff. The difference in bias between the 400-gram and 500-gram lower limits is only 2 to 3% but excluding 400-499 gram infants omits an additional 9.3% of deaths from analysis.

Based on the graphs and on the fact that it was not selected as an explanatory variable in the regression analysis, it appears unlikely that fetal death reporting requirements explain infant death underreporting. Misreporting of infant deaths as fetal deaths, unlinked deaths, and other deaths with unknown BW probably contribute small amounts. Based on wide variation in BW-specific IMR in the 200 to 399-gram range, it is likely that verification and data cleaning processes are followed in some states and not others. This may account for most of the underreporting of infant deaths.

This study was mainly limited to information found in vital records data, and it was a state-level analysis. Having a “gold standard” such as clinical records to compare to would be the best way to judge completeness of reporting. An analysis at county and state levels could address other sources of variation and would be useful to communities.

CHAPTER 3. REFRESHING THE PERINATAL PERIODS OF RISK: A NEW REFERENCE GROUP AND NATIONWIDE COUNTY-LEVEL ANALYSES

CHAPTER 3 INTRODUCTION

Overview of PPOR Analytic Steps

A brief overview of the PPOR analytic steps is included in the following three paragraphs to provide context for the analyses in this chapter. The first phase of analysis divides a community's fetal and infant deaths into four categories, called periods of risk (Sappenfield et al., 2010a). The periods were defined by weight at birth and age at death, with dividing lines chosen such that each period presents a different set of likely causes. Communities calculate period-specific mortality rates and compare them to corresponding rates in a *reference population* using simple subtraction. The difference, the "excess mortality rate," is an estimate of *preventable* mortality. PPOR's ability to estimate preventability at the population level is important because it is usually not possible to determine whether an individual death was preventable from vital records data.

The reference population is an actual population of infant and fetal deaths typically selected based on their mothers' educational attainment, race/ethnicity, and age. Education is used primarily as a proxy for socioeconomic status (SES), as it is associated with both a woman's current SES and that of her family of origin, both of which potentially affect health (Braveman, 2014; Singh & Yu, 2019; Weinberg et al., 2019). Race, a powerful social construct in the United States, is a proxy for the effects of racism, such as multi-generational access to housing, education, health care, and wealth (Vilda, Hardeman, Dyer, Theall, & Wallace, 2021; Wallace, Crear-Perry, Richardson, Tarver, & Theall, 2017). Maternal age affects health through both social and biological mechanisms. Teen mothers are more likely to be poor and less educated, and to lack stable housing and jobs (Hodgkinson, Beers, Southammakosane, & Lewin, 2014;

Roth, Hendrickson, Schilling, & Stowell, 1998). Mothers in their 40s may have more resources, but they are more likely to have health problems such as hypertension, diabetes, and obesity (Khoshnood, Wall, & Lee, 2005; Woodall AM, 2020). The rationale for using a reference group to estimate preventability is a statement of justice. If one segment of the population has reached a particular health goal, all should be able to reach that goal (Peck et al., 2010). Communities are encouraged to choose a reference group that serves as an acceptable target and helps to make a compelling case in their local context. Many communities have used state or local versions of the original published national reference group: White non-Hispanic mothers aged 20 or older with at least one year of post-high school education. Variations have included using other racial/ethnic groups, omitting race as a criterion, requiring a bachelor's degree, or omitted older mothers.

PPOR analysis continues with a second phase that investigates period(s) with large excess mortality (Sappenfield, Peck, Gilbert, Haynatzka, & Bryant, 2010b). Many communities use confidence intervals or significance tests to help select periods to investigate, especially with small numbers of deaths. Phase 2 begins by further dividing excess mortality by causal mechanism. Then it focuses on known risk factors for the predominant causes. Phase 2 analysis uses vital records and other data sources (e.g., Census) to identify *which* of the many known causes of poor birth outcomes are likely contributors to preventable mortality *in the study population*. Other information sources such as Fetal Infant Mortality Reviews (ncfrp.org), scientific literature, and the expertise of local community members help direct the analyses and inform action. The analysis ends with community decisions based on findings.

National Reference Groups

National reference groups are needed in PPOR for several reasons. They facilitate national comparisons, and their large numbers and stable rates reduce the variance of excess

mortality estimates. State and local reference groups may be preferred and may also reduce bias due to differential reporting (Christiansen-Lindquist et al., 2017; Gould, 1999; Kowaleski, 1997). However, many communities cannot create them due to small numbers of deaths, data quality limitations, or administrative barriers. Any community can use national reference groups. We provide national Black reference groups for reasons outlined in the Methods section. Our county-level analysis provides a perspective that separate community analyses, with their different reference groups and timeframes, cannot. It also allows communities to compare their findings, *by component*, to other US counties. A national perspective is valuable: An unusually poor outcome in one community may mean that solutions in place elsewhere could be implemented locally, while an unusually good outcome may alert policymakers to local practices that work well and should be disseminated.

National Analysis

Although tabulations of excess mortality for the four periods of risk for US counties were made for the 1995-2002 data years, the distributions of key outcomes for the *second* phase of analysis were never provided. The most important of these uses Kitagawa decomposition (Kitagawa, 1955) to divide excess mortality in the Maternal Health and Prematurity Period into two portions: one due to BW distribution and the other due to birthweight-specific mortality (Sappenfield et al., 2010b). Another critical Phase 2 analytic step divides the infant health period by underlying cause of death. Many communities using PPOR find that SUID (Sudden Unexplained Infant Death) is so predominant relative to other causes that they divide this period according to whether the cause of death is SUID or not. SUID includes ICD-10 Codes R95 (Accidental suffocation and strangulation in bed), R99 (Other ill-defined and unspecified causes of mortality), and W75 (Sudden infant death syndrome – SIDS) (Shapiro-Mendoza et al., 2014). Many SUID deaths are related to unsafe infant sleep conditions

such as prone sleep position, unsafe sleep surfaces, and bedsharing. This chapter reports on six components of excess mortality: the four periods of risk, with two divided according to the above Phase 2 analytic steps.

This chapter's study objectives are (1) to assess the quality of vital records data elements needed to complete the initial steps of PPOR and create reference groups, (2) create a set of unbiased, nationally representative reference populations, (3) describe the six components of PPOR excess mortality for the US and selected subpopulations, and (4) describe the distribution of the six components across large US counties.

CHAPTER 3 METHODS

We obtained vital records micro-data files from the National Center for Health Statistics by special request because our analysis required county identifiers. This cross-sectional study used all live births, fetal deaths, and linked infant deaths for the years 2014-2016 with maternal residence at the time of birth or delivery in the 50 US states or Washington DC. IRB approval was obtained through the University of Nebraska Medical Center (**#** 667-18-EP). Analyses were completed using SAS 9.4.

Data Quality

We assessed the quality of two sets of data elements: those needed to assign cases to a Period of Risk and those needed to create reference groups. The first set includes birthweight (BW) for live births, BW and gestational age (GA) for fetal deaths, and BW, age at death, and cause of death for infant deaths. PPOR guidelines call for excluding fetal deaths at less than 24 weeks gestation, and all cases born weighing less than 500 grams (Sappenfield et al., 2010a). These criteria were designed to reduce bias due to differential reporting (Kowaleski, 1997). **Excluded cases** essentially form a fifth risk period that is not involved in comparisons and thus

does not contribute to bias. We classify cases that do not have the BW or GA information needed to classify them as either excluded or belonging to one of the four risk periods as having **“unknown status."** Such cases contribute to bias because rates of missing information are much higher among deaths (the PPOR numerator) than among births (most of the PPOR denominator). In its standard tabulations, the CDC makes a presumption that effectively imputes *unknown* GA to *GA<20* in states that require reporting of all fetal deaths (E. C. Gregory, Driscoll, Anne, 2014). We followed this practice and classify these cases as "excluded." We refer to cases that meet the BW and GA criteria as **"PPOR-valid."** For births, deaths, and fetal deaths, we calculated the percentages of cases with unknown BW, unknown GA, unknown status, as well as the percentages of excluded and PPOR-valid cases.

The second set of data elements, needed to create reference groups, includes maternal age, race/ethnicity, and educational attainment. We counted a case as **"qualified"** for inclusion in the reference group if it met all reference group criteria and **"disqualified"** if it failed to meet one or more criteria. Unknown data elements in this set only contribute to bias if the cases are PPOR-valid, and the impact of missing data for one element depends on the values of the others. For example, a case with unknown maternal education would not contribute to bias if it were disqualified based on maternal age or race/ethnicity, but would contribute if it were otherwise qualified. We counted a case as **"unclassifiable**" if it could not be qualified nor disqualified. We calculated the percent of unclassifiable cases nationwide, by state, and by county. We estimated the bias caused by unclassifiable deaths by assuming they would qualify for the reference group in the same proportion as classifiable cases did, by risk period. Because of substantive and uneven estimated bias caused by missing information for fetal deaths, we chose to omit states from the national reference group pool if more than 30% of their fetal deaths were unclassifiable.

Reference Groups

We selected reference groups from the remaining pool of states. We began by using the traditional PPOR reference group criteria, choosing infants and fetal deaths of White non-Hispanic mothers aged 20 and older with at least some post-high school education. We further restricted the population by successively excluding women with less than a bachelor's degree and women aged 40 and older, creating two additional White reference groups with lower rates. In recent years, most communities have used reference groups based on the traditional or bachelor's degree criteria to study their whole population or a subgroup. However, among the many communities using PPOR to address high *Black* infant mortality rates, *some* prefer a Black reference group. Although Black reference groups have higher rates and produce smaller excess mortality estimates, they produce more compelling results for some communities. Therefore, we created three analogous Black reference groups.

The four periods of risk, with two divided based on Phase 2 analysis form the six components of excess mortality used in this analysis. They are described in detail elsewhere (Sappenfield et al., 2010a, 2010b) and summarized in Table 6. First, we used the most restrictive reference group to calculate excess mortality for the four risk periods. We then followed standard PPOR methods to calculate period-specific excess mortality rates by subtracting reference group rates from study group rates. Finally, we subdivided the IH period by cause of death group and used Kitagawa decomposition to subdivide MHP excess mortality. We calculated the six components of PPOR excess mortality for the US, the four US Census regions, and selected maternal demographic subpopulations. We produced Poisson confidence intervals for reference group rates using SAS PROC GENMOD.

County-Level National Analysis

For the county-level analysis, we selected counties based on the denominator to avoid biasing the study sample toward high-mortality counties. We included counties with 25,000 or more PPOR-valid fetal deaths plus live births. We displayed PPOR findings for a selected county using the standard four-period grid and a pie chart showing the six components of excess mortality. Finally, we summarized the distribution of study sample counties by graphing components of excess mortality as percentiles.

Table 6 Defining the four periods of risk and six components of excess mortality:

1. MHP Maternal Health and Prematurity	Fetal deaths at 24 or more weeks of gestation weighing 500-1499 grams at delivery, plus Infant deaths weighing 500-1499 grams at birth. < 1500 grams is termed "Very Low Birth Weight" or VLBW.
MHP_BW Birthweight Distribution	Portion of MHP excess mortality due to birthweight distribution, i.e. more "babies being born way too small" in the study population than in the reference population
MHP_SM Birthweight-specific Mortality	Portion of MHP excess mortality due to birthweight specific mortality, i.e. "the babies born way too small" in the study population are not surviving as well as those in the reference population
2. MC Maternal Care	Fetal deaths delivered at 24 or more weeks of gestation and weighing 1,500 grams or more. "Larger stillbirths"
3. NC Newborn Care	Infants (born alive) and weighing at least 1,500 grams who died on or before their 27th day.
4. IH Infant Health	Infants (born alive) and weighing at least 1,500 grams who died between 28 and 364 days of age.
IH_SUID Potentially sleep-related	Portion of IH deaths with underlying cause of death coded as R95 (SIDS), W75 (Suffocation or strangulation in bed), or R99 (Ill-defined)
IH_OTHER	Portion of IH deaths with underlying cause of death coded as any other cause.

CHAPTER 3 RESULTS

Data Quality

In the 2014-2016 period, 11,912,448 live births, 155,476 fetal deaths, and 69,410 infant deaths (linked to their birth record) occurred among US residents. The infant mortality rate (IMR) was 5.83 infant deaths per thousand live births. Table 7 summarizes the three data files, with numbers and percentages of unknowns *contributing to bias* shaded in grey. The 3.54% of fetal deaths with unknown GA were mostly presumed to be less than 20 weeks gestation (following a CDC standard practice) (Kirmeyer, 2006) and so were excluded from PPOR analysis. This step reduced the unknown gestation to 0.24%. Although over half of fetal deaths were missing BW after excluding cases with GA<24 weeks, only 1.75% of fetal deaths could not be classified into a risk period. Among live births, 0.17% were excluded from PPOR analysis because they weighed less than 500 grams, and 0.04% because BW was unknown. Among infant deaths, 24% were excluded because they weighed less than 500 grams, and 0.6% because their BW was unknown. No infant death cases were missing age at death or cause of death.

Nationwide, 1.4% of fetal and infant deaths did not have the information needed to classify them into a period of risk, and this percentage exceeded 5% in only 12 of the 294 counties large enough to do PPOR (7.6% was the maximum). We excluded 74% of fetal deaths and 24% of infant deaths by implementing standard PPOR restrictions needed to assure unbiased comparisons.

Table 7 Missing birthweight and gestational age, and their impact on classifying cases as either "excluded" or belonging to a risk period. Measures that reflect data quality are shaded.

		Live Births	Fetal Deaths	Infant Deaths
	All	**11,912,448**	**155,478**	**69,410**
	N GA Unknown	9,626	5,498	629
	% GA Unknown	0.08%	3.54%	0.91%
with CDC Presumption*	**N GA Unknown**	NA.	368	NA.
	% GA Unknown	NA.	0.24%	NA.
	N BW Unknown	4,951	77,947	405
	% BW Unknown	0.04%	50.13%	0.58%
For PPOR	**Unknown status**	0.04%	1.75%	0.58%
	Excluded**	0.17%	74.28%	23.99%
	N PPOR-Valid	**11,887,728**	**37,255**	**52,356**
	% PPOR-Valid	99.79%	23.96%	75.43%

*Standard CDC tabulations presume that in states that require reporting of all fetal deaths, fetal deaths with unknown birthweights were delivered at <20 weeks gestation. This means they are not included in most tabulations, and that they would be excluded from PPOR analyses.

**To assure comparability across states, counties, and hospitals, PPOR analysis excludes fetal deaths at <24 weeks gestation and fetal deaths, infant deaths, and live births at <500 grams birth weight. Cases not excluded are referred to as "PPOR-Valid".

The quality of data elements needed to classify cases into (or out of) the traditional PPOR reference group is summarized in Table 8. Less than 2% of live births and infant deaths had unknown maternal age, race/ethnicity, and education, but among fetal deaths, 22.8% were missing race/ethnicity, and 15.62% were missing educational attainment. Reference group status could not be determined for a total of 28.5% of fetal deaths using the most restrictive reference group (which was worst-case in terms of bias). This unclassifiable percentage dropped to 8.5% for the PPOR-valid subset. MHP and MC reference group rates were biased downward by an estimated 3.4% and 6.4%, respectively. To reduce MHP and MC bias, we created a pool of states from which to select reference populations by omitting eight states with high percentages of unclassifiable fetal deaths (34% to 88%, in RI, CT, WV, NH, NY, MA, VA, and CO). The remaining subset of 42 states plus Washington DC accounted for 84% of US live births and forms the "reference pool". The reference pool has better data quality than the national dataset. Among PPOR-valid births in the reference pool, reference group status could not be determined for 4.56% of fetal deaths, and MHP and MC bias was reduced to 2.3% and 3.1%, respectively.

Table 8 Missing information needed for reference group status: maternal age, race/ethnicity, and educational attainment

	Live Births	Fetal Deaths	Infant Deaths
Among All Events			
Age Unknown	0.01%	0.55%	0.06%
Race/Ethnicity Unknown	0.83%	22.80%	1.86%
Education Unknown	1.78%	15.62%	1.83%
Unclassifiable*	2.57%	28.49%	3.77%
Among Select Groups			
Unclassifiable *among PPOR-valid	2.56%	8.46%	3.31%
Unclassifiable* among PPOR-valid with 8 states** omitted	1.57%	4.56%	2.68%

* Unclassifiable: the case could not be classified as either *qualified for* the traditional reference group based on all three demographic characteristics or *disqualified* based on one or more characteristics.

**The eight states where more than 30% of fetal deaths could not be classified as either 'disqualified from', or 'qualified for' the reference group were RI, CT, WV, NH, NY, MA, VA, and CO.

Reference Groups

Period-specific feto-infant mortality rates (FIMR) for the nation and six proposed national reference groups are shown in Table 9, with columns for the MHP, MC, and NC periods, and both components of the IH period. The last column is the "overall" FIMR, the sum of the four period-specific rates, including the two IH components. The two components of the MHP period are defined only in terms of *excess* mortality, using Kitagawa decomposition. Counts and rates for the national reference groups and selected state reference groups, including those needed for Kitagawa decomposition, are provided in Appendix A. The group based on traditional criteria is labeled W1. Below it are increasingly restricted white reference groups W2 and W3 and corresponding Black reference groups B1 through B3. As with Black IMR, the overall Black reference group FIMRs are roughly double their White counterparts. We can examine the disparity using the PPOR approach by treating W1 as the reference group for B1. Overall excess mortality was 10.29 - 5.29 =5.00. The MHP difference (4.61-1.85=2.67) accounts for 55% of that gap. The predominance of the MHP period is consistent with the previously reported importance of very low BW (<1,500 gram) deaths in the US Black/White IMR disparity (Iyasu, Becerra, Rowley, & Hogue, 1992). MC contributes 18%, and IHSUID 13%.

Further restricting reference group criteria changes both the excess rates and the relative importance of the components. For the White reference group, requiring a bachelor's degree reduces the overall FIMR by 17%, from 5.29 in W1 to 4.37 in W2. The bachelor's degree requirement reduces rates slightly in most components but reduces the SUID rate by 57%. For the Black reference group, the bachelor's degree has a similar effect, except that the MC rate increases slightly. This may reflect findings that higher socioeconomic status is accompanied by more stress among Black mothers and does not benefit Black birth outcomes in the US as much

as it benefits Whites (Assari, 2020; Ekeke, Mendez, Yanowitz, & Catov, 2020; Fishman et al., 2020; Sims & Coley, 2019). Omitting women over age 40 has a small effect generally. However, for Black women, it decreases the MC rate slightly, from 2.39 to 2.29 (4.2%), and the NC rate by 4.9%, consistent with findings that Black women over age 40, in particular, have an increased risk of poor birth outcomes (Brisendine, Rice, Goldfarb, & Wingate, 2020; Metz, 2020; Schummers et al., 2019). The SUID rate for W2 is 0.19, consistent with international best-case SUID rates (Müller-Nordhorn et al., 2020), and is not further reduced by omitting mothers age 40 and over.

These reference groups represent large portions of the US population, and accordingly, confidence intervals are narrow. W1 and W3 include 37% and 20% of the pool of states from which the reference populations are drawn, respectively. The most restrictive Black reference group includes only 2% of the reference pool, but 14% of the pool's Black subpopulation. Poisson confidence intervals for the overall rates range from +/-.075 (+/-1.4%) for the traditional group to +/-.394 (+/-4.6%) for the most restrictive Black reference group.

Table 9 PPOR rates for the nation and six potential reference groups.

The denominator for all rates in a population is the number of live births plus fetal deaths in that population. All cases are PPOR-valid, i.e., all weighing less than 500 grams at birth or delivery and fetal deaths delivered at less than 24 weeks gestation are excluded. All reference groups (B1-W3) exclude Hispanic ethnicity and teens.

PPOR-Valid numbers and rates		**MHP**	**MC**	**NC**	**IH Other**	**IH SUID**	**Overall FIMR (sum)**
USA, 50 states plus Washington DC		2.83	1.93	1.17	0.82	0.76	7.51
Reference pool, 8 states omitted		2.89	1.98	1.22	0.86	0.81	7.76
Reference Groups, 8 states omitted	W1 White, age 20+, some college credit or more.	1.85 (1.8, 1.89)	1.41 (1.37, 1.45)	1.00 (0.97, 1.04)	0.59 (0.56, 0.61)	0.44 (0.42, 0.47)	5.29 (5.22, 5.37)
	W2 White, age 20+, bachelor's degree or more	1.56 (1.51, 1.62)	1.25 (1.21, 1.30)	0.90 (0.86, 0.94)	0.47 (0.44, 0.50)	0.19 (0.17, 0.21)	4.37 (4.28, 4.46)
	W3 White, age 20-39, bachelor's degree or more	1.53 (1.48, 1.59)	1.22 (1.17, 1.27)	0.87 (0.83, 0.91)	0.46 (0.43, 0.49)	0.19 (0.17, 0.21)	4.27 (4.18, 4.36)
	B1 Black, age 20+, some college credit or more.	4.61 (4.46, 4.77)	2.31 (2.20, 2.42)	1.25 (1.17, 1.33)	1.01 (0.94, 1.09)	1.11 (1.03, 1.19)	10.29 (10.06, 10.52)
	B2 Black, age 20+, bachelor's degree or more	4.09 (3.83, 4.36)	2.39 (2.19, 2.60)	1.05 (0.93, 1.19)	0.87 (0.76, 1.00)	0.47 (0.39, 0.57)	8.87 (8.49, 9.26)
	B3 Black, age 20-39, bachelor's degree or more	4.05 (3.79, 4.33)	2.29 (2.09, 2.50)	1.00 (0.87, 1.14)	0.86 (0.74, 0.99)	0.47 (0.38, 0.57)	8.66 (8.27, 9.06)

*FOOTNOTE *The traditional reference group was originally defined as "13 or more years of education" using the 1989 revision of the birth certificate. For the current (2003) revision, we substitute "some college credit, but not a degree," or more.*

National PPOR Analysis

For the remaining analyses (nationwide, subpopulations, and county-level), we used the most restrictive group, W3. We selected this group because it produces positive excess mortality in most of our study counties and provides an IH_SUID target rate that has been achieved in other populations (Müller-Nordhorn). Since the W2 period-specific rates are not significantly different from W3 rates, communities may not wish to use an upper age restriction. The six components of excess mortality relative to reference group W3 are shown in Table 10 for the whole US and selected subpopulations. The second row is the reference pool of 42 states plus Washington DC. The remaining rows include all states. Compared to the US, overall mortality is higher for the Southeast and Midwest, for non-Hispanic Black mothers, mothers with no more than a high school diploma or GED, teen mothers, and mothers age 40 and over. Negative excess mortality occurs in the MHP SM component. At the county level, this can occur in any component and, while it precludes a pie chart display, it simply means that the reference group rate is higher than the study group rate for that component. The MHP BW component contributes the most to excess mortality except among Whites, where IH SUID contributes more. MC ranks second among the components, except among Whites, teen mothers, and mothers with low educational attainment.

Table 10 PPOR excess mortality for the US and selected subpopulations

Excess mortality using the most restricted reference group in Table 9 (W3: White, non-Hispanic mothers age 20 -39, with a bachelor's degree or more education).

		MHP BW	MHP SM	MC	NC	IH Other	IH SUID	Overall (sum)
USA (Nationwide)		1.16	0.14	0.71	0.30	0.36	0.57	3.25
USA minus 8 states		1.19	0.17	0.76	0.36	0.40	0.62	3.49
US Census Regions	Northeast	1.04	(0.01)	0.49	0.03	0.16	0.32	2.04
	West	0.62	0.22	0.60	0.21	0.25	0.35	2.24
	Midwest	1.11	0.20	0.71	0.42	0.47	0.60	3.51
	Southeast	1.58	0.13	0.87	0.40	0.45	0.80	4.23
Maternal demographics	Black non-Hispanic	4.15	(0.09)	1.73	0.55	0.81	1.36	8.50
	White non-Hispanic	0.50	0.23	0.49	0.28	0.30	0.53	2.32
	High school or less	1.63	0.20	0.96	0.56	0.66	1.00	5.00
	Age 0 to 19	2.01	0.54	0.90	0.57	0.86	1.29	6.18
	Age 40 or older	2.50	0.36	2.14	1.00	0.61	0.20	6.81

* The eight states excluded due to poor data quality for maternal characteristics were RI, CT, WV, NY, VA, MA, NJ, and CO served as the pool from which reference populations were selected.

County-Level PPOR Analysis

Of the 3,143 US counties or county-equivalents, 294 in our 3-year sample had the minimum 60 PPOR-valid fetal and infant deaths required for the first phase of PPOR analysis. One hundred counties had PPOR-valid denominators of at least 25,000 and were included in this study. The included counties had between 105 and 2,293 PPOR-valid deaths. There were at least seven deaths in any risk period and at least four deaths in any Kitagawa stratum (one county had two Kitagawa strata with n=4, and two others had one stratum with n=4.) The study counties represented 33 states and all four US Census regions. The study counties' overall PPOR mortality averaged 7.30, lower than the US rate (7.51), and ranged from 3.51 to 13.09. Three counties had more than 5% of cases missing the information needed to classify them into a risk period, with the maximum being 7%. A maximum period of five years is recommended for PPOR analysis. With five years of data, we estimate that approximately 500 counties could do PPOR, and 300 would have large enough numbers for at least a partial Phase 2 analyses.

Communities using PPOR typically display their period-specific excess mortality rates in a grid, as shown in Figure 9. The pie chart shows the same county's results after the initial Phase 2 steps that divide the MHP and IH periods. In this county, MHP_BW contributes the most, with 43% of excess mortality due to "too many babies born too small." The MC period contributes 23%, and the community might also prioritize that period for further investigation and community action. The SUID rate was .54, and SUID contributed only 13% of excess mortality. Communities should consider their actual (absolute) mortality in addition to their excess mortality. For example, a second southern county in our sample had a similar percentage of excess mortality due to SUID, but had a SUID rate of 1.63. The second county would have more to gain by prioritizing SUID than the one depicted in Figure 9.

Figure 9 PPOR analysis for a Southern US county using the most restrictive reference group (W3 in Table 4).

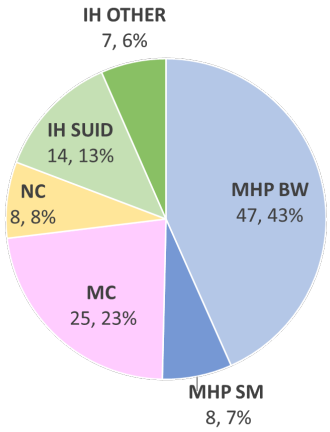
On the left we see the PPOR "Map" depicting the results from Phase 1 analysis. The pie chart on the right shows the six components of excess mortality described above. The IH and MHP periods are subdivided using Phase 2 analytic methods, and multiplying the excess mortality rates by the county's PPOR-valid denominator allows us to express excess mortality as the estimated *number* of preventable deaths. This community might focus on MHP_BW and MC, prioritizing prevention of very low birthweight births and third-trimester stillbirths for further study and community action.

The PPOR "Map" of excess mortality (Phase 1 analysis)

	Fetal Deaths >=24 wks	Neonatal deaths (1-27 days)	Post-neonatal deaths (28-364 days)
500-1499 grams	**Maternal Health / Prematurity (MHP)** 114 fetal and infant deaths 2.95 rate **1.43** excess rate		
1500+ grams	**Maternal Care (MC)** 72 fetal deaths 1.87 rate **0.65** excess rate	**Newborn Care (NC)** 42 infant deaths 1.09 rate **0.22** excess rate	**Infant Health (IH)** 46 infant deaths 1.19 rate **0.54** excess rate

Overall (PPOR-valid) 274 deaths=136 infant + 138 fetal
PPOR denominator (fetal + births)= 38,581
7.10 rate; **2.84 excess rate;** 108 estimated excess deaths

Estimated excess mortality expressed as numbers of preventable deaths, for six components of excess mortality.



Percentile plots of period-specific mortality rates for all 100 counties in our study sample are shown in Figure 10. To use the graph, consider a hypothetical community with an MC mortality rate of 1.60 deaths per thousand. Beginning at 1.60 on the y-axis, move horizontally to the MC line, then drop to the x-axis to see that only about 25% of large US counties have lower MC rates. Thus, even before using a reference group, we can see that our hypothetical county is doing better in the MC period than 75% of counties. Figure 10 also shows that the MHP period has the highest rates and the most variability across the counties, and that the NC period has the lowest rates and the least variation. The SUID component makes up less than half of the IH period for most counties, but more than half for counties in the upper quartile.

Excess mortality percentiles (relative to reference group W3) are plotted for all six components in Figure 11. Except for the MHP period, the shapes are the same. However, the components shift vertically by differing amounts. The MHP BW excess mortality component is highest of the six and shaped very much like overall MHP mortality, because birthweight distribution far exceeds birthweight-specific mortality as a cause of death in the MHP period in most US cities. Notably, after subtracting reference group rates, the IH_SUID component becomes more prominent, reflecting the preventability of many SUID deaths relative to other IH causes. The decreased prominence of the NC component reflects consistency of newborn care and outcomes across the US. MC excess mortality is second highest except for counties in the lowest tenth percentile for that period. We found that MHP BW contributes the largest portion of excess mortality in 81 of the 100 counties. Among the other five components, MC contributes most in 64 counties, and IH contributes most in 25. Negative excess mortality occurs for at least some counties in all components. Although the MHP

SM component is negative for nearly half the counties, it represents more than a quarter of MHP mortality in 10% of counties and can indicate problems with systems of care (Catalano et al., 2017; Kim et al., 2013). Complete tables with counts and excess mortality rates are found in the appendices.

Figure 10 Percentiles of PPOR mortality rates for 100 large US Counties.

PPOR mortality for the four periods of risk, with two additional lines showing the components of the IH period (the two components of MHP are not defined until excess mortality is calculated). The study sample is 100 US counties with PPOR-valid denominators >=25,000 in the 2014-2016 period.

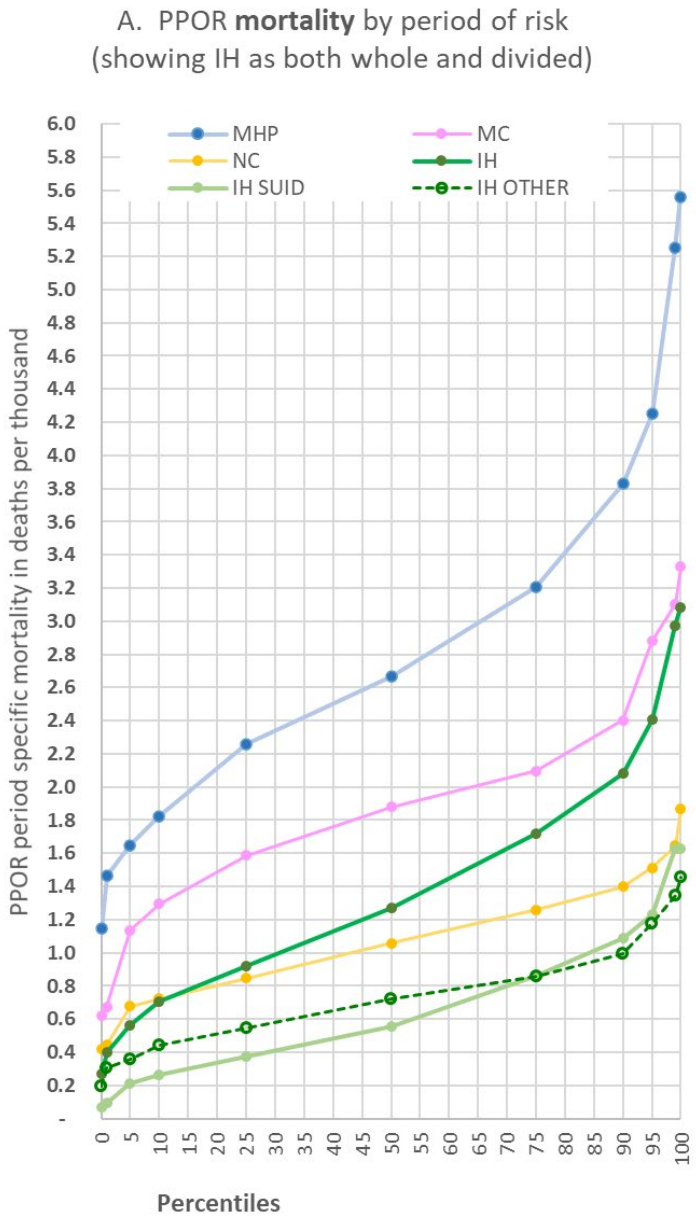
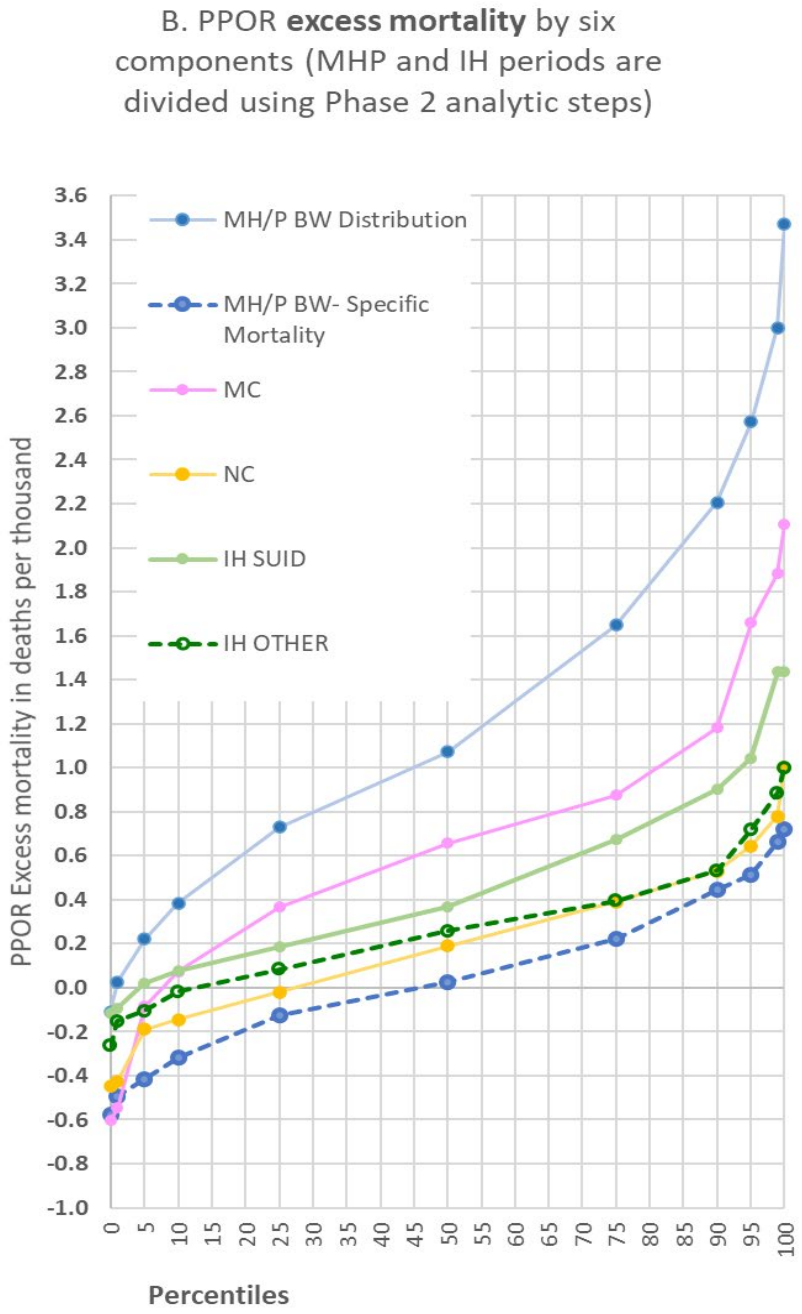


Figure 11 Percentiles of PPOR Excess Mortality for 100 large US Counties

Six components of PPOR ***excess*** **mortality**: the four periods of risk with MHP and IH subdivided. The study sample is 100 US counties with PPOR-valid denominators >=25,000 in the 2014-2016 period.



CHAPTER 3 DISCUSSION

Although local and state reference groups may be desired, many communities cannot use them. In 8 states, more than 30% of fetal deaths are missing the maternal demographic information they would need to assign the case to a reference group. These unclassifiable fetal deaths bias reference group rates downward and artificially increase excess mortality rates in the MHP and MC periods. In many counties, small numbers of educated white mothers limit the county's ability to create a reference group with stable rates. In some counties, high white mortality rates (often high SUID rates) mean that local reference group rates may not be the optimal goals to use. This analysis provides alternatives for these communities and others seeking a national benchmark.

This study shows that PPOR outcomes vary widely across US cities, which is evidence that the approach provides valuable information. The variation can and should inspire action: Implicit beliefs that outcomes cannot change sometimes constrain community leaders and residents alike. Counter-examples provided by the national reference groups and by counties across the nation can be powerful motivators.

A limitation inherent in PPOR analysis is that it is ecological. PPOR is a community process, and its outcomes are measurable only at the community level. Another limitation is that our study uses counties of widely varying numbers of births and deaths, which causes differing statistical precision (Murphy, Xu, Kochanek, Curtin, & Arias)(Murphy, Xu, Kochanek, Curtin, & Arias). Data quality problems also impose limitations, in particular the need to exclude eight states due to high rates of missing maternal education and race/ethnicity. Eliminating those states decreased bias but means that the national reference groups do not truly represent the nation.

Omitting <24-week fetal deaths and all <500-gram deaths and births, i.e., restricting our analysis to "PPOR-valid" cases, leaves out many fetal and infant deaths. The excluded cases are no less important than the included cases. A recent state-level analysis showed that current data quality does not justify lowering the limits. In fact, there is measurable bias even above those limits . Further study is needed to determine whether lower limits are justified *when using state or local reference groups*. This determination would vary by state and depend on the consistency of county, health system, and hospital-level reporting practices.

This study is the first to examine the distribution of excess mortality across six components of excess mortality. This refinement highlights two causes of excess mortality that are often important in communities: VLBW infant & fetal deaths, and SUID deaths. The innovation of a graphical display of PPOR percentiles allows communities to identify unusual findings. Most communities consider only the relative importance of their own components of excess mortality, but comparing to other counties may emphasize different problems and potentially different solutions.

DISCUSSION

INNOVATIONS

The original PPOR exclusion criteria ("cutoffs") were based on graphs of GA and BW distributions, and the expertise of a multi-agency workgroup (Peck et al., 2010). Our study updates this initial assessment, includes analysis by fetal death reporting requirements, and adds adjustment for health differences. We also introduce a simple method for quantifying the bias caused by underreporting, allowing comparisons of different cutoffs. For infant deaths we used the same methods, adding graphical analysis of the birthweight-specific infant mortality rate.

This study is the first to examine the distribution of excess mortality across six components, beginning with the four Periods of Risk defined in PPOR Phase 1 analysis, and subdividing two of them using Phase 2 analytic methods. This refinement highlights two causes of excess mortality that often predominate in communities: VLBW infant & fetal deaths, and SUID deaths. The innovation of a graphical display of PPOR percentiles allows communities to determine whether one of their components is exceptional. Most communities consider only the relative importance of their own components of excess mortality, but their ranking relative to other counties may emphasize different problems and potentially different solutions. For example, a community with high overall mortality might find that their IH SUID component does not contribute as much relative to other components, yet it might be much higher than the SUID component in other communities. Another component that is low in general but high in a few communities is MHP SM. Both these examples might represent "low hanging fruit" opportunities that communities could choose to address before or along with components that contribute larger proportions to their excess mortality.

Another simple innovation this study adds is a quantitative basis for grouping states according to their fetal death reporting requirements. Reporting requirements are specified in terms of birthweight and gestational age, and researchers have grouped them in various ways. Grouping is more complicated than it would seem, in part because the CDC does not have up-to-date information (they rely on states to provide it), and in part because of the variety of requirements that exist. For example, the most common requirement is simply to report fetal deaths if they are “>=20 weeks”. However, some requirements also include a phrase such as “or if they are >=500 grams”, and still others add the phrase “if GA is missing”. Each of these requirements, if followed, would capture different numbers of fetal deaths. To further complicate grouping, state GA limits included 12, 16, and 20 weeks, and BW limits included 350, 400, and 500 grams. Our study is the first we are aware of that group requirements in part by “coverage”, that is, the percentage of fetal death that are theoretically captured by the different reporting specifications. We used the joint BW and GA distribution of the five states that require reporting of all fetal deaths as the benchmark joint distribution. Coverage measures allowed us to quantify differences between the state’s requirements, which varied from 6.1% for New Mexico which required reporting of fetal deaths at 500 grams or more, to 11% for states that required reporting of fetal deaths if they had reached 350 grams OR reached 20 weeks gestation. New Mexico has since instituted the 20-week requirement. The most important conclusion we draw from this is that all other requirements fall far short of the requirement to “report all” fetal deaths. A second conclusion is that adding the 350-gram requirement to the 20-weeks requirement captures more deaths. As a side note it seems also to increase reporting of birthweight slightly.

IMPORTANT FINDINGS

We found strong evidence of underreporting of both fetal and infant deaths. Reporting requirements have a strong relationship with underreporting for fetal deaths, but not for infant deaths.

Our analysis of fetal deaths confirmed previous findings that state reporting requirements affect the distribution of deaths, even in GA and BW ranges for which reporting requirements agree (Goyal et al., 2017; Kirmeyer, 2006). We found that, between 20 and 23 weeks of GA, Gold states as a group have higher fetal death ratios than other reporting requirement groups. Even states that require reporting of fetal deaths at 20 weeks or more are not, in general, reporting the expected FMRs until GA reaches 24 weeks. States that have only a BW requirement do not reach expected FMRs until 25 weeks or later. In fact, most fetal deaths are delivered before 20 weeks GA and are not recorded in the US.

Reporting differences among states are large enough to produce meaningfully biased comparisons, including excess mortality estimates in PPOR analyses. We quantified bias for both fetal and infant deaths by comparing two hypothetical places that report at the 25th and 75th percentiles, respectively. Those places are chosen to represent an average or typical comparison. The bias between two randomly selected places would be larger if they reported at, for example, the 95th and 5th percentiles, or non-existent if they reported at the same percentile. Including all fetal deaths at 20 or more weeks, we found that “average” bias was 15% of the national >20-week fetal mortality rate. The 24-week minimum GA recommended for PPOR analysis reduced bias to about 2%. In PPOR analyses, places with more complete reporting, such as Gold states, would have artificially high excess rates in the Maternal Health and Prematurity Period (<1500 grams) when compared with a reference population with

reporting practices like most states. The reverse is true for places that report fewer early fetal deaths, such as the five states that report based only on BW.

Although fetal death BW distribution was similar to the GA distribution, we found that the fetal death BW cutoff is unnecessary (redundant) if the 24-week GA cutoff is employed. Communities wishing to continue using the BW cutoff for consistency with infant deaths will, on average, exclude few additional cases.

We analyzed infant deaths only by BW because PPOR requirements are specified in terms of BW. PPOR recommendations use BW because it is more precisely and accurately measured than GA for live births. Very few infants born at less than 400 grams are known to survive for a full year, yet the nationwide average birthweight-specific mortality rates indicated that more than 10% of those infants survived. This is strong evidence of underreporting of infant deaths, and it supports excluding deaths of infants delivered at less than 400 grams from PPOR analyses.

We also assessed underreporting of infant deaths using lack of agreement of birthweight distribution among states as evidence, following the method used for fetal deaths. Variation across states increased with decreasing birthweight, beginning with the 700-799 BW category. This would support a cutoff as high as 800. Bias analysis indicated that when all infant deaths are included, variation among states (which may be attributed to reporting differences) is large enough to bias a typical comparison between places by 14% of the overall infant mortality rate (8% with adjustment for health differences). Excluding birthweights below 400 grams reduces the bias to 8% of the overall IMR (4% adjusted) and omitting BW<500 grams reduces it to 5% (3% adjusted). There are both benefits and costs of reducing the PPOR cutoff from 500 grams to 400 grams. It would increase bias by an estimated 2% to 3% but would add

about 9% more deaths to the analysis. While there is a clear reason for the 400-gram cutoff, the evidence is not as strong for omitting infants between 400 and 499 grams.

It is possible that both fetal and infant deaths exclusion criteria could be based on BW instead of GA. However, in practice fetal deaths would still be excluded based on BW, because many fetal deaths with unknown BW would have to be imputed from known GA. Excluding based on GA for both fetal and infant deaths could be considered, though BW is measured more precisely. It is possible that GA measurements using ultrasound are precise and accurate enough for this purpose.

It is likely that several causes combine to produce the apparent underreporting of deaths, i.e. the gap between expected and observed death ratios or rates. Underlying health differences measured at the state level do not “close the gap” for either fetal or infant deaths. For fetal deaths, our study indicates that reporting requirements account for a large proportion of the gap. For infant deaths this does not appear to be true, based on the graphs and on the fact that “fetal death reporting group” was not selected as an explanatory variable in the regression analysis. For infant deaths our analysis showed that misreporting (fetal vs infant), unlinked deaths, and unknown birthweight account for very small portions of the gap. The wide variation in birthweight-specific infant mortality rates among states in the 200 to 399-gram range is evidence that follow-up and verification processes are followed in some states and not others. Adoption of such processes by all states for infant deaths below 800 grams would likely result in more infant deaths being reported in each BW category. Previously published studies lead us to believe that county, hospital, and physician-level reporting practices contribute to differential reporting of both fetal and infant deaths.

LIMITATIONS

A fundamental limitation in this study is that there is no perfect source of information to compare our data to. In the absence of a true "gold standard" (such as clinical records) to judge completeness of reporting, we relied on comparisons within the vital records data.

Fetal death quality posed major limitations. Three strategies were needed to address high rates of unknown fetal BW and GA: (1) presuming that cases with unknown GA in states that report all fetal deaths are <20 weeks, (2) substituting median BW for GA for cases with unknown BW, and (3) excluding fetal deaths at <24 weeks. None of these strategies is perfect. We do not know how many fetal deaths in Gold states had unknown GA that was actually >= 20 weeks. There is wide variation in BW at each week of GA, so substituting median BW is a rough imputation that does not preserve the true variability of the population. Nearly half of GA >=20 fetal deaths are GA 20-23 weeks, and nearly a quarter of infant deaths have BW<500 grams, so these exclusions leave out many deaths. The excluded cases are no less real than the included cases. However, the impact of omitting so many deaths is not as substantial as their numbers would suggest. Including them would primarily increase rates in the MHP period, which is already the predominant period in most counties. If the fetal death GA restriction were lowered some additional deaths would fall into the MC period, encouraging more communities to focus on preventing larger stillbirths.

The restrictions needed to address GA and BW also reduced bias caused by missing maternal educational attainment and race/ethnicity. However, in our Chapter 3 analysis we needed one additional strategy to reduce bias of national reference groups. We omitted eight states with poor data quality from the national pool from which the reference groups were drawn. Omitting states meant that our reference groups were not truly "national". Including these states would have substantively changed the relative importance of the periods of risk,

underestimating the MC period by more than 6% while not underestimating the NC or IH periods. Eliminating the eight states decreased this bias and increased reference group rates because cases that qualify are included rather than inadvertently excluded due to missing maternal information. Our reference rates may slightly overestimate rates from a (hypothetical) unbiased 50-state pool, since the IMR for the included states was 3% higher than the IMR for the whole US. In sum, although our reference groups are not perfect, their bias is not substantive and is balanced across periods of risk.

Our analysis of data quality and underreporting was ecological and done only at the state level. An analysis at both county and state levels could address other sources of variation and would be useful to communities. It is possible that comparisons within a state would be less biased because of common reporting requirements and presumed similar follow-up and verification processes. In states where this is true, lower limits might be justified when using state or local reference groups. However, it is also possible that our state-level analysis obscures even more extreme local variations. Documented county, hospital, and physician sources of variation (Ramsay & Santella, 2011; Williams & Magsumbol, 2010) are not addressed in this study. Further study would be necessary to assess bias caused by reporting variations within states, and conclusions would likely vary by state. It is important to remember that PPOR analysis itself is inherently ecological. PPOR is a community process, and its outcomes are measurable at the community (county, city, and sometimes even sub-city) levels.

Further research could include multilevel studies that have more potential to take into account the multiple sources of variation in reporting practices. More detailed information about state vital records reporting & verification practices, which would have improved this study, is complex and difficult to obtain. Though we corresponded with several state vital records departments, we did not have the resources to collect this information from all states.

Future studies could also investigate the possibility of excluding both fetal and infant deaths based on GA, especially as GA measurement using ultrasound is becoming more prevalent. To improve the analysis of infant deaths it might be possible to identify states with good reporting practices and use those as a benchmark.

CONCLUSIONS

Our investigations show that ignoring underreporting and data quality problems will bias comparisons across states, even states with equivalent reporting requirements. Excluding fetal deaths at <24 weeks gestation and infant deaths at <=500 grams reduces this bias to ignorable levels. Excluding fetal deaths at <500 grams excludes few additional cases and provides consistency with infant death BW exclusion criteria. Because known sources of variation operate at county, hospital, and physician levels, it is likely that differential reporting also exists within states.

The set of national reference groups provided in Chapter 3 will allow more communities to use PPOR, and will facilitate comparisons across communities. Because the reference groups are large and have stable rates, they may be useful for several years, as the original reference groups were.

Findings detailed in Chapter 3 show that PPOR outcomes vary widely across US cities, which is evidence that the approach provides valuable information. While MHP BW is predominant in most counties, exceptions occurred in 20%. In addition, there is substantial variation in absolute excess mortality rates and relative importance of the five other components. The variation can and should inspire action: Implicit beliefs that outcomes cannot change sometimes constrain community leaders and residents alike. Counter-examples

provided by reference populations shown in Table 4 and the national distributions shown in Figure 4 can be powerful motivators.

REFERENCES

Allen, M. C., Donohue Pk Fau - Dusman, A. E., & Dusman, A. E. The limit of viability--neonatal outcome of infants born at 22 to 25 weeks' gestation. (0028-4793 (Print)).

Assari, S. (2020). Protective Effects of Maternal Education against Low Birth Weight Deliveries: Blacks' Diminished Returns. *Res Health Sci, 5*(4), 1-17. doi:10.22158/rhs.v5n4p1

Bell, E. F., & Zumbach, D. K. (2011). The Tiniest Babies: A Registry of Survivors With Birth Weight Less Than 400 Grams. *Pediatrics, 127*(1), 58-61. doi:10.1542/peds.2010-1855

Besculides, M., & Laraque, F. (2005). Racial and ethnic disparities in perinatal mortality: applying the perinatal periods of risk model to identify areas for intervention. *Journal of the National Medical Association, 97*(8), 1128-1132.

Braveman, P. G., Laura. (2014). The Social Determinants of Health: It's Time to Consider the Causes of the Causes. *Public health reports (Washington, D.C.: 1974), 129*.

Brisendine, A. E., Rice, W. S., Goldfarb, S. S., & Wingate, M. S. (2020). The weathering hypothesis and stillbirth: racial disparities across the life span. *Ethnicity and Health, 25*(3), 354-366. doi:10.1080/13557858.2017.1420145

Brumberg, H. L., Dozor, D., & Golombek, S. G. History of the birth certificate: from inception to the future of electronic data.

Catalano, A., Bennett, A., Busacker, A., Carr, A., Goodman, D., Kroelinger, C., . . . Barfield, W. (2017). Implementing CDC's Level of Care Assessment Tool (LOCATe): A National Collaboration to Improve Maternal and Child Health. *J Womens Health (Larchmt), 26*(12), 1265-1269. doi:10.1089/jwh.2017.6771

Christiansen-Lindquist, L., Silver, R. M., Parker, C. B., Dudley, D. J., Koch, M. A., Reddy, U. M., . . . Hogue, C. J. R. (2017). Fetal death certificate data quality: a tale of two U.S. counties. *Annals of Epidemiology, 27*(8), 466-471.e462. doi:10.1016/j.annepidem.2017.07.001

CityMatCH. (2012). Celebrating the 15th Anniversary of PPOR's Conception. *CityLights.* Retrieved from https://www.citymatch.org/wp-content/uploads/2012/01/CityLights201212.pdf

Ekeke, P., Mendez, D. D., Yanowitz, T. D., & Catov, J. M. (2020). Racial Differences in the Biochemical Effects of Stress in Pregnancy. *International Journal of Environmental Research and Public Health, 17*(19). doi:10.3390/ijerph17196941

El-Metwally, D., Vohr B Fau - Tucker, R., & Tucker, R. Survival and neonatal morbidity at the limits of viability in the mid 1990s: 22 to 25 weeks. (0022-3476 (Print)).

Fishman, S. H., Hummer, R. A., Sierra, G., Hargrove, T., Powers, D. A., & Rogers, R. G. (2020). Race/ethnicity, maternal educational attainment, and infant mortality in the United States. *Biodemography and Social Biology, 66*(1), 1-26. doi:10.1080/19485565.2020.1793659

Gould, J. B. (1999). Vital records for quality improvement. *Pediatrics, 103*(1 Suppl E), 278-290.

Gould, J. B., Chavez, G., Marks, A. R., & Liu, H. (2002). Incomplete birth certificates: a risk marker for infant mortality. *American Journal of Public Health, 92*(1), 79-81. doi:10.2105/ajph.92.1.79

Goyal, N. K., DeFranco, E., Kamath-Rayne, B. D., Beck, A. F., & Hall, E. S. (2017). County-level Variation in Infant Mortality Reporting at Early Previable Gestational Ages. *Paediatric and Perinatal Epidemiology, 31*(5), 385-391. doi:10.1111/ppe.12376

Gregory, E. C., Driscoll, Anne. (2014). *User Guide to the 2014 Fetal Death Public Use File*. National Center for Health Statistics, Centers for Disease Control and Prevention Retrieved from

https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/fetaldeath/2014FetalUserGuide.pdf

Gregory, E. C. W., Martin, J. A., Argov, E. L., & Osterman, M. J. K. (2019). *Assessing the Quality of Medical and Health Data From the 2003 Birth Certificate Revision: Results From New York City*. Retrieved from Hyattsville, MD:

Heck, K. E., Schoendorf, K. C., & Parker, J. (1999). Are very low birthweight births among American Indians and Alaska Natives underregistered? *International Journal of Epidemiology, 28*(6), 1096.

Hodgkinson, S., Beers, L., Southammakosane, C., & Lewin, A. (2014). Addressing the mental health needs of pregnant and parenting adolescents. *Pediatrics, 133*(1), 114-122. doi:10.1542/peds.2013-0927

Hoyert Dl Fau - Gregory, E. C. W., & Gregory, E. C. W. Cause-of-death Data From the Fetal Death File, 2015-2017. (1551-8930 (Electronic)).

Iyasu, S., Becerra, J. E., Rowley, D. L., & Hogue, C. J. R. (1992). Impact of very low birthweight on the black-white infant mortality gap. *American Journal of Preventive Medicine, 8*(5).

Khoshnood, B., Wall, S., & Lee, K. S. (2005). Risk of low birth weight associated with advanced maternal age among four ethnic groups in the United States. *Matern Child Health J, 9*(1), 3-9. doi:10.1007/s10995-005-2446-4

Kim, E. W., Teague-Ross, T. J., Greenfield, W. W., Keith Williams, D., Kuo, D., & Hall, R. W. (2013). Telemedicine collaboration improves perinatal regionalization and lowers statewide infant mortality. *Journal of Perinatology, 33*(9), 725-730. doi:10.1038/jp.2013.37

Kirmeyer, S., MacDorman, M. (2006). *User Guide to the 2006 Fetal Death Public Use File*. Hyattsville, MD: National Center for Health Statistics Retrieved from https://ftp.cdc.gov/pub/Health_Statistics/NCHS/Dataset_Documentation/DVS/fetaldeath/2006FetalUserGuide.pdf

Kitagawa, E. M. (1955). Components of a difference between two rates. *Journal of the American Statistical Association, 50*(272), 1168-1194.

Kleinman, J. C. (1986). Underreporting of infant deaths: then and now. *American Journal of Public Health, 76*(4), 365-366. doi:10.2105/ajph.76.4.365

Kowaleski, J. L. (1997). *State definitions and reporting requirements for live births, fetal deaths, and induced termination of pregnancy. 1997 edition* (ISBN 0-8406-0534-X). Retrieved from Hyattsville, Maryland:

Lumley, J. (2003). Defining the problem: the epidemiology of preterm birth. *BJOG: An International Journal of Obstetrics and Gynaecology, 110 Suppl 20*, 3-7.

Martin, J. (2014). The Birth Certificate (Finally) Goes National. Retrieved from https://blogs.cdc.gov/inside-nchs/2014/07/30/the-birth-certificate-finally-goes-national/

McCarthy, B. J., Terry, J., Rochat, R. W., Quave, S., & Tyler, C. W., Jr. (1980). The underregistration of neonatal deaths: Georgia 1974--77. *American Journal of Public Health, 70*(9), 977. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1619518/pdf/amjph00682-0051.pdf

Melnik, T. A., Guldal, C. G., Schoen, L. D., Alicandro, J., & Henfield, P. (2015). Barriers in Accurate and Complete Birth Registration in New York State. *Maternal and child health journal, 19*(9), 1943. doi:10.1007/s10995-015-1711-4 [doi]

Metz, T. D., ; Snipe Berry, Rana; Fretts, Ruth C.; Reddy, Uma M.; Turrentine, Mark A. (2020). Management of Stillbirth: Obstetric Care Consensus No, 10. *Obstetrics and Gynecology, 135*(3), e110-e132. doi:10.1097/aog.0000000000003719

Murphy, S. L., Xu, J., Kochanek, K. D., Curtin, S. C., & Arias, E. Deaths : final data for 2015.

NAPHSIS, T. N. A. f. P. H. S. a. I. S. (2021). Jurisdictions. Retrieved from https://www.naphsis.org/systems

NCHS. (2016). *User Guide to the 2016 Period Linked Birth/Infant Death Public Use File*. Hyattsville, Maryland

Nijkamp, J. W., Sebire, N. J., Bouman, K., Korteweg, F. J., Erwich, J. J. H. M., & Gordijn, S. J. (2017). Perinatal death investigations: What is current practice? *Seminars in Fetal & Neonatal Medicine, 22*, 167-175. doi:10.1016/j.siny.2017.02.005.

Peck, M. G., Sappenfield, W. M., & Skala, J. (2010). Perinatal periods of risk: a community approach for using data to improve women and infants' health. *Maternal and child health journal, 14*(6), 864-874. doi:10.1007/s10995-010-0626-3

Prince, C. B., Young, M. B., Sappenfield, W., & Parrish, J. W. (2016). Investigating the Decline of Fetal and Infant Mortality Rates in Alaska During 2010 and 2011. *Maternal and child health journal*. doi:10.1007/s10995-015-1906-8 [doi]

Ramsay, S. M., & Santella, R. M. (2011). The definition of life: a survey of obstetricians and neonatologists in New York City hospitals regarding extremely premature births. *Maternal and child health journal, 15*(4), 446-452. doi:10.1007/s10995-010-0613-8 [doi]

Rodriguez, R. N., & Yao, Y. (2017). *Five things you should know about quantile regression.* Paper presented at the Proceedings of the SAS global forum 2017 conference, Orlando.

Roth, J., Hendrickson, J., Schilling, M., & Stowell, D. W. (1998). The risk of teen mothers having low birth weight babies: implications of recent medical research for school health personnel. *Journal of School Health, 68*(7), 271-275. doi:10.1111/j.1746-1561.1998.tb00581.x

Sappenfield, W. M., Peck, M. G., Gilbert, C. S., Haynatzka, V. R., & Bryant, T., 3rd. (2010a). Perinatal periods of risk: analytic preparation and phase 1 analytic methods for investigating feto-infant mortality. *Matern Child Health J, 14*(6), 838-850. doi:10.1007/s10995-010-0625-4

Sappenfield, W. M., Peck, M. G., Gilbert, C. S., Haynatzka, V. R., & Bryant, T., 3rd. (2010b). Perinatal periods of risk: phase 2 analytic methods for further investigating feto-infant mortality. *Matern Child Health J, 14*(6), 851-863. doi:10.1007/s10995-010-0624-5

Schummers, L., Hacker, M. R., Williams, P. L., Hutcheon, J. A., Vanderweele, T. J., McElrath, T. F., & Hernandez-Diaz, S. (2019). Variation in relationships between maternal age at first birth and pregnancy outcomes by maternal race: a population-based cohort study in the United States. *BMJ Open, 9*(12), e033697. doi:10.1136/bmjopen-2019-033697

Shapiro-Mendoza, C. K., Camperlengo, L., Ludvigsen, R., Cottengim, C., Anderson, R. N., Andrew, T., . . . MacDorman, M. (2014). Classification system for the Sudden Unexpected Infant Death Case Registry and its application. *Pediatrics, 134*(1), e210. doi:10.1542/peds.2014-0180 [doi]

Shapiro, S. (1950). Estimating Birth Registration Completeness. *Journal of the American Statistical Association, 45*(250), 261-264. doi:10.1080/01621459.1950.10483357

Sims, J., & Coley, R. L. (2019). Variations in links between educational success and health: Implications for enduring health disparities. *Cultural Diversity & Ethnic Minority Psychology, 25*(1), 32-43. doi:10.1037/cdp0000239

Singh, G. K., & Yu, S. M. (2019). Infant Mortality in the United States, 1915-2017: Large Social Inequalities have Persisted for Over a Century. *Int J MCH AIDS, 8*(1), 19-31. doi:10.21106/ijma.271

Smith, L. K., Hindori-Mohangoo, A. D., Delnord, M., Durox, M., Szamotulska, K., Macfarlane, A., . . . Zeitlin, J. (2018). Quantifying the burden of stillbirths before 28 weeks of completed

gestational age in high-income countries: a population-based study of 19 European countries. *Lancet, 392*(10158), 1639-1646. doi:10.1016/s0140-6736(18)31651-9

Tyler, C. P., Grady, S. C., Grigorescu, V., Luke, B., Todem, D., & Paneth, N. (2012). Impact of fetal death reporting requirements on early neonatal and fetal mortality rates and racial disparities. *Public health reports (Washington, D.C.: 1974), 127*(5), 507-515. doi:10.1177/003335491212700506 [doi]

Vilda, D., Hardeman, R., Dyer, L., Theall, K. P., & Wallace, M. (2021). Structural racism, racial inequities and urban-rural differences in infant mortality in the US. *Journal of Epidemiology and Community Health*. doi:10.1136/jech-2020-214260

Wallace, M., Crear-Perry, J., Richardson, L., Tarver, M., & Theall, K. (2017). Separate and unequal: Structural racism and infant mortality in the US. *Health Place, 45*, 140-144. doi:10.1016/j.healthplace.2017.03.012

Weinberg, D., Stevens, G., Finkenauer, C., Brunekreef, B., Smit, H. A., & Wijga, A. H. (2019). The pathways from parental and neighbourhood socioeconomic status to adolescent educational attainment: An examination of the role of cognitive ability, teacher assessment, and educational expectations. *PloS One, 14*(5), e0216803. doi:10.1371/journal.pone.0216803

Westat. (2016). *Electronic Death Reporting System Online Reference Manual: A Resource Guide for Jurisdictions*. Retrieved from https://www.cdc.gov/nchs/data/dvs/EDRS-Online-Reference-Manual.pdf

Williams, B. L., & Magsumbol, M. S. (2010). Inclusion of non-viable neonates in the birth record and its impact on infant mortality rates in Shelby County, Tennessee, USA. *Pediatric Reports, 2*(1), e1. doi:10.4081/pr.2010.e1

Wingate, M. S., & Alexander, G. R. (2006). Racial and ethnic differences in perinatal mortality: the role of fetal death. *Annals of Epidemiology, 16*(6), 485. doi:S1047-2797(05)00091-8 [pii]

Woodall AM, D. A. (2020). Racial and Ethnic Differences in Mortality rate of Infants Born to Teen Mothers: United States, 2017-2018. *NCHS Data Brief, 2020 Jul;(371):1-8*.

Xaverius, P., Salas, J., Kiel, D., & Woolfolk, C. (2014). Very low birth weight and Perinatal Periods of Risk: Disparities in St. Louis. *Biomed Res Int, 2014*, 547234. doi:10.1155/2014/547234

APPENDICES

Appendix A: PPOR counts and rates for national and state reference groups

Group	USA	RefPool	W1	W2	W3	B1	B2	B3
Description	USA, 50 states plus Washing-ton DC	USA with 8 states excluded	White, age 20+ , some college credit or more.	White, age 20+ , bachelor's degree or more	White, age 20-39, bachelor's degree or more	Black, age 20+, some college credit or more.	Black, age 20+, bach-elor's degree or more	Black, age 20-39, bach-elor's degree or more
All births	11,912,448	9,970,390	3,684,551	2,028,353	1,950,520	725,283	227,845	214,642
All fetal deaths	155,476	87,974	17,851	9,169	8,488	7,645	2,351	2,099
All linked infant	69,410	60,096	14,382	6,369	5,972	7,127	1,930	1,804
PPOR valid denom-	11,924,983	9,981,618	3,687,827	2,029,748	1,951,812	724,745	227,631	214,415
Excluded deaths	132,144	68,378	12,229	6,422	5,909	7,116	2,208	1,994
Period not known	3,131	2,257	478	244	226	201	55	53
MH/P counts	33,771	28,801	6,811	3,175	2,984	3,343	930	868
MC counts	23,028	19,806	5,205	2,542	2,383	1,674	543	490
NC counts	13,930	12,212	3,705	1,819	1,694	903	239	214
IH counts	18,882	16,616	3,805	1,336	1,264	1,535	306	284
IH Other counts	9,805	8,568	2,170	945	898	732	198	184
IH SUID counts	9,077	8,048	1,635	391	366	803	108	100
MHP Rates	2.83	2.89	1.85	1.56	1.53	4.61	4.09	4.05
MC Rates	1.93	1.98	1.41	1.25	1.22	2.31	2.39	2.29
NC Rates	1.17	1.22	1	0.9	0.87	1.25	1.05	1
IH Rates	1.58	1.66	1.03	0.66	0.65	2.12	1.34	1.32
IH Other Rates	0.82	0.86	0.59	0.47	0.46	1.01	0.87	0.86
IH SUID Rates	0.76	0.81	0.44	0.19	0.19	1.11	0.47	0.47
Overall Rates	7.51	7.76	5.29	4.37	4.27	10.29	8.87	8.66

Appendix B: Kitagawa counts and rates for national and state reference groups

	KitagawaCategory	Births	Infant	Fetal	Inf+Fet	Birth+Fet	Denom	C	P
USA	1-KITA1500-749g	29,312	11,217	4,460	15,677	33,772	11,924,983	0.4642	0.0028
	2-KITA2750-999g	32,979	4,039	3,900	7,939	36,879	11,924,983	0.2153	0.0031
	3-KITA31000-1249g	39,240	2,335	3,074	5,409	42,314	11,924,983	0.1278	0.0036
	4-KITA41250-1499g	48,384	1,953	2,793	4,746	51,177	11,924,983	0.0927	0.0043
	5-KITA51500-1999g	187,955	4,560	5,401	9,961	193,356	11,924,983	0.0515	0.0162
	6-KITA62000-2499g	607,525	5,950	5,275	11,225	612,800	11,924,983	0.0183	0.0514
	7-KITA72500-8165g	10,942,333	22,302	12,352	34,654	10,954,685	11,924,983	0.0032	0.9186

	KitagawaCategory	Births	Infant	Fetal	Inf+Fet	Birth+Fet	Denom	C	P
Reference Pool	1-KITA1500-749g	24,680	9,384	3,813	13,197	28,493	9,981,618	0.4632	0.0029
	2-KITA2750-999g	27,781	3,489	3,310	6,799	31,091	9,981,618	0.2187	0.0031
	3-KITA31000-1249g	32,991	2,039	2,673	4,712	35,664	9,981,618	0.1321	0.0036
	4-KITA41250-1499g	40,410	1,687	2,406	4,093	42,816	9,981,618	0.0956	0.0043
	5-KITA51500-1999g	157,348	4,019	4,708	8,727	162,056	9,981,618	0.0539	0.0162
	6-KITA62000-2499g	509,018	5,268	4,504	9,772	513,522	9,981,618	0.019	0.0515
	7-KITA72500-8165g	9,157,382	19,541	10,594	30,135	9,167,976	9,981,618	0.0033	0.9185

	KitagawaCategory	Births	Infant	Fetal	Inf+Fet	Birth+Fet	Denom	C	P
W1	1-KITA1500-749g	5,444	2,161	863	3,024	6,307	3,687,827	0.4795	0.0017
	2-KITA2750-999g	7,019	846	754	1,600	7,773	3,687,827	0.2058	0.0021
	3-KITA31000-1249g	9,134	566	605	1,171	9,739	3,687,827	0.1202	0.0026
	4-KITA41250-1499g	11,762	438	578	1,016	12,340	3,687,827	0.0823	0.0034
	5-KITA51500-1999g	47,325	1,075	1,111	2,186	48,436	3,687,827	0.0451	0.0131
	6-KITA62000-2499g	147,197	1,287	1,099	2,386	148,296	3,687,827	0.0161	0.0402
	7-KITA72500-8165g	3,451,941	5,148	2,995	8,143	3,454,936	3,687,827	0.0024	0.9369

	KitagawaCategory	Births	Infant	Fetal	Inf+Fet	Birth+Fet	Denom	C	P
W2	1-KITA1500-749g	2,596	1,045	413	1,458	3,009	2,029,748	0.4846	0.0015
	2-KITA2750-999g	3,440	367	341	708	3,781	2,029,748	0.1873	0.0019
	3-KITA31000-1249g	4,579	281	274	555	4,853	2,029,748	0.1144	0.0024
	4-KITA41250-1499g	6,059	184	270	454	6,329	2,029,748	0.0717	0.0031
	5-KITA51500-1999g	24,735	518	505	1,023	25,240	2,029,748	0.0405	0.0124
	6-KITA62000-2499g	75,155	574	527	1,101	75,682	2,029,748	0.0146	0.0373
	7-KITA72500-8165g	1,909,344	2,063	1,510	3,573	1,910,854	2,029,748	0.0019	0.9414

	KitagawaCategory	Births	Infant	Fetal	Inf+Fet	Birth+Fet	Denom	C	P
W3	1-KITA1500-749g	2,451	1,000	389	1,389	2,840	1,951,812	0.4891	0.0015
	2-KITA2750-999g	3,247	343	319	662	3,566	1,951,812	0.1856	0.0018
	3-KITA31000-1249g	4,287	257	252	509	4,539	1,951,812	0.1121	0.0023
	4-KITA41250-1499g	5,639	168	256	424	5,895	1,951,812	0.0719	0.003
	5-KITA51500-1999g	22,934	469	464	933	23,398	1,951,812	0.0399	0.012
	6-KITA62000-2499g	70,466	531	495	1,026	70,961	1,951,812	0.0145	0.0364
	7-KITA72500-8165g	1,839,189	1,958	1,424	3,382	1,840,613	1,951,812	0.0018	0.943

	KitagawaCategory	Births	Infant	Fetal	Inf+Fet	Birth+Fet	Denom	C	P
B1	1-KITA1500-749g	4,282	1,365	410	1,775	4,692	724,745	0.3783	0.0065
	2-KITA2750-999g	4,337	415	321	736	4,658	724,745	0.158	0.0064
	3-KITA31000-1249g	4,412	193	255	448	4,667	724,745	0.096	0.0064
	4-KITA41250-1499g	5,014	161	223	384	5,237	724,745	0.0733	0.0072
	5-KITA51500-1999g	17,816	372	460	832	18,276	724,745	0.0455	0.0252
	6-KITA62000-2499g	51,835	461	400	861	52,235	724,745	0.0165	0.0721
	7-KITA72500-8165g	634,166	1,605	814	2,419	634,980	724,745	0.0038	0.8761

	KitagawaCategory	Births	Infant	Fetal	Inf+Fet	Birth+Fet	Denom	C	P
B2	1-KITA1500-749g	1,241	404	96	500	1,337	227,631	0.374	0.0059
	2-KITA2750-999g	1,301	117	74	191	1,375	227,631	0.1389	0.006
	3-KITA31000-1249g	1,353	52	79	131	1,432	227,631	0.0915	0.0063
	4-KITA41250-1499g	1,538	48	60	108	1,598	227,631	0.0676	0.007
	5-KITA51500-1999g	5,115	95	140	235	5,255	227,631	0.0447	0.0231
	6-KITA62000-2499g	14,311	105	123	228	14,434	227,631	0.0158	0.0634
	7-KITA72500-8165g	201,920	345	280	625	202,200	227,631	0.0031	0.8883

	KitagawaCategory	Births	Infant	Fetal	Inf+Fet	Birth+Fet	Denom	C	P
B3	1-KITA1500-749g	1,159	384	90	474	1,249	214,415	0.3795	0.0058
	2-KITA2750-999g	1,203	109	67	176	1,270	214,415	0.1386	0.0059
	3-KITA31000-1249g	1,233	49	76	125	1,309	214,415	0.0955	0.0061
	4-KITA41250-1499g	1,386	41	52	93	1,438	214,415	0.0647	0.0067
	5-KITA51500-1999g	4,613	82	127	209	4,740	214,415	0.0441	0.0221
	6-KITA62000-2499g	13,109	89	116	205	13,225	214,415	0.0155	0.0617
	7-KITA72500-8165g	190,937	327	247	574	191,184	214,415	0.003	0.8917