Published in final edited form as:

Vaccine. 2015 August 26; 33(36): 4446–4450. doi:10.1016/j.vaccine.2015.07.037.

Demographic characteristics of members of the Vaccine Safety Datalink (VSD): A comparison with the United States population

Lakshmi Sukumaran, MD, MPH^{1,2}, Natalie L. McCarthy, MPH¹, Rongxia Li, PhD¹, Eric Weintraub, MPH¹, Steven J. Jacobsen, MD, PhD³, Simon J. Hambidge, MD, PhD⁴, Lisa A. Jackson, MD, MPH⁵, Allison L. Naleway, PhD⁶, Berwick Chan, BS⁷, Biwen Tao, MS¹, and Julianne Gee, MPH¹

¹Immunization Safety Office, Centers for Disease Control and Prevention, Atlanta, Georgia

²Department of Pediatrics, Emory University School of Medicine, Atlanta, Georgia

³Department of Research and Evaluation, Kaiser Permanente of Southern California, Pasadena, California

⁴Department of Ambulatory Care Services, Denver Health, Institute for Health Research, Kaiser Permanente Colorado, Department of Pediatrics, University of Colorado, Denver, Colorado

⁵Group Health Research Institute, Seattle, Washington

⁶The Center for Health Research, Kaiser Permanente Northwest, Portland, Oregon

⁷Kaiser Permanente Vaccine Study Center, Oakland, California

Abstract

Background—The Vaccine Safety Datalink (VSD) is a collaboration between CDC and 9 integrated health care systems that serves as a cornerstone of US post-licensure vaccine safety monitoring. Given concerns that potential differences between the insured VSD population and the US population could limit the generalizability of VSD study findings, we performed a comparison of the demographic characteristics between the two populations.

Methods—We collected data from medical records and administrative files at VSD sites in 2010 to compare sex, age, race, ethnicity, income, and educational attainment to the 2010 US Census population. We also compared data on the 2012 VSD Medicaid population to 2012 US Medicaid data.

Results—The VSD population included over 8 million individuals in 2010, which represented 2.6% of the total US population. All major demographic groups were represented in the VSD. We

Corresponding Author: Sukumaran, Lakshmi, xfq3@cdc.gov, 1600 Clifton Rd NE, Building 16, MS D-26, Atlanta, GA 30333.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Contributors: LS, NLM and RL were involved in collecting data and writing the manuscript. All authors contributed to designing the study, interpreting the data, revising the draft, and approved the manuscript for submission.

Conflicts of Interest: All VSD authors are funded by the Centers for Disease Control and prevention. Dr. Sukumaran received research support from the National Institutes of Health. Dr. Naleway received research support from GlaxoSmithKline and Pfizer.

found no major differences in comparing sex, race, ethnicity and educational attainment between the VSD and the US population. Middle income populations were comparable between the VSD and the US. While the percentage of lower income populations was less in the VSD compared to the US, the VSD had over 2 million individuals in this group. Additionally, there were over 600,000 Medicaid members in the VSD in 2012, which represented 1.1% of the US Medicaid population.

Conclusions—We found that the VSD population is representative of the general US population on several key demographic and socioeconomic variables. Despite a few specific groups being underrepresented in the VSD compared to the US, the absolute number of VSD members is large enough to ensure significant representation of these groups in vaccine safety studies that use VSD data.

Keywords

Vaccine Safety Datalink; demographics; census; Medicaid

Introduction

Vaccines are considered one of the most important public health successes of the century, and vaccine safety monitoring is a critical component to any vaccination program [1, 2]. The Vaccine Safety Datalink (VSD) is a collaboration between the Centers for Disease Control and Prevention (CDC) and 9 integrated health care systems (sites) that conducts post-marketing vaccine safety evaluations on approximately 3% of the United States (US) population [3, 4]. The VSD provides essential vaccine safety data to stakeholders (health care providers, public health officials, and the public) and can inform national immunization policy.

The VSD has played a significant role in vaccine safety monitoring in the United States. Important public health investigations conducted by the VSD include studies that have found associations between rotavirus vaccine and intussusception [5–7], and febrile seizures and vaccines [8, 9], as well as a lack of association between vaccines and neurodevelopmental disorders [10, 11]. The VSD has also conducted studies on the safety of influenza and Tdap vaccination in pregnant women [12–14]. A potential limitation of VSD data is a perceived lack of generalizability between the insured VSD population and the more socioeconomically diverse US population. This limitation was commented upon in a recent Institute of Medicine (IOM) report on the safety of the childhood immunization schedule [15]. Differences in education, income, or socioeconomic status may translate to differences in health care seeking behaviors [16, 17], access to healthcare, and differences in adverse events following vaccination [18].

While one study compared the demographic characteristics of one VSD site to its state [19], and other studies have examined VSD vaccination coverage rates and mortality rates relative to those of the US [20, 21], a comprehensive comparison of the demographic characteristics of the entire VSD population with the US population has not been previously conducted. The prior vaccination coverage study showed that the VSD generally had higher vaccination coverage in children compared to the US population [20]. The goal of this study was to

compare the VSD population to the US population in order to examine the generalizability of VSD findings to the broader US population.

Methods

VSD data are provided by nine sites in diverse geographical locations across the US. Of these, six infrastructure sites provide data on a continuous basis: Kaiser Permanente of Southern California (SCK), Kaiser Permanente of Northern California (NCK), Kaiser Permanente Northwest (NWK), Group Health Cooperative (GHC), Kaiser Permanente of Colorado (KPC), and Marshfield Clinic Research Foundation (MFC). Three additional sites can provide data for specific studies: Health Partners Research Foundation (HPM), Harvard Pilgrim Health Care (HAR), and Kaiser Permanente of Georgia (KPG). Each site collaborates to compile data in a standardized manner. The data includes vaccination records, demographic information, enrollment information, birth information, and records on hospitalizations, outpatient visits, emergency department visits, and urgent care visits. Each VSD member has a unique, randomized VSD identification number used to link data on demographics and medical services. VSD data is updated on a weekly basis. In addition, the VSD has ready access to electronic medical records, allowing investigators to conduct chart reviews to validate health outcomes.

For this study, we used data from only the six VSD infrastructure sites. We collected data from 2010 in order to compare the VSD population to 2010 US Census, which is the most recent comprehensive demographic information available on the US population. We also collected information on the VSD Medicaid population from 2005 through 2013. The study was approved by the individual institutional review boards at CDC and all participating VSD sites.

Race and Ethnicity/ Socioeconomic status

We classified race into the following six categories: white, black or African American, American Indian or Alaskan Native, Asian or Pacific Islander, other, and multiple races. Ethnicity was categorized as Hispanic or non-Hispanic. At all sites, race and ethnicity information was obtained from a combination of local administrative data and geocoding data [22], which are neighborhood estimates based on linking health plan member addresses (zip code, block-group, census tract) to US Census block data. While VSD sites used different software for geocoding (e.g., SAS, Geographic Information Coding software), the underlying method of estimating demographic data based on geographic location was similar. At each site, the majority of race and ethnicity information came from administrative data, and a small percentage came from neighborhood estimates. All of the education and income information came from neighborhood estimates.

Medicaid

US Medicaid is a federal and state funded program that provides health insurance for millions of Americans, and includes low-income adults, children, pregnant women, elderly, and disabled populations (<http://www.cms.gov/>). Medicaid information has been collected routinely in the VSD since 1997. Each year, a Medicaid file is created containing Medicaid

coverage status by month for VSD members. Each VSD site creates the Medicaid file using a unique, site-specific algorithm. In addition to Medicaid, each site may also have members participating in local Medicaid related programs that cover specific low income populations. Beginning in 2013 all VSD sites began to include these low income populations from state-subsidized programs in the VSD Medicaid file.

We counted the total number of Medicaid participants each year by taking an average of monthly data. We then looked at the secular trend of the Medicaid population at each site from 2005 through 2013 stratified by the following age groups: <5 years, 5–17 years, 18–49 years, 50–64 years, and 65 + years. We also compared the 2012 VSD Medicaid population with the reference population from the US federal Medicaid program, stratified by sex and age group.

Reference populations

Our reference population was the entire population of the United States, information on which was available from the 2010 US Census. We obtained this information from the US Census Bureau website (<http://www.census.gov/>). Additionally, for the Medicaid population, we used 2012 data from US Centers for Medicare and Medicaid Services available at <http://www.cms.gov/>.

Statistical analysis

For descriptive comparison, we report percentages for the US population and Medicaid data, and the corresponding percentages in the VSD. Formal statistical comparisons were not performed, since both the US and VSD populations are large, and non-meaningful differences would result in statistically significant test results.

Results

Among the VSD infrastructure sites in 2010, there were 8,085,329 people enrolled (Table 1), which represented 2.6% of the total US population. The distribution by sex was similar between the VSD and US populations. The age distributions were also similar between the two groups; however, the 55 to 59 year age group and 60 to 64 year age group were each slightly overrepresented in the VSD (7.2% in the VSD vs. 4.8% in the US and 6.2% in the VSD vs. 3.8% in the US respectively).

When comparing race within the VSD to that of the US population, the most notable difference was in the Asian population (15.4% in the VSD compared to 5.0% in the US). Additionally, the VSD had a smaller proportion of white (64.7% in the VSD compared to 72.4% in the US) and African American members (8.8% in the VSD compared to 12.6% in the US). When comparing ethnicity, the Hispanic population (36.6% in the VSD compared to 16.3% in the US) was overrepresented in the VSD.

A similar percentage of patients in the VSD and the US populations had incomes of \$50,000 to \$74,999. Slightly more than 50% of the US population had annual incomes of less than \$50,000, compared to 34% of the VSD population. Despite these differences, there were

over 2 million VSD members earning less than \$50,000. Conversely, over 43% of the VSD population had incomes over \$75,000, compared to 32% of the US population.

Education attainment comparisons, based on neighborhood estimates, showed that the VSD adult population was similar in educational attainment to the US adult population. The largest differences were between the population with less than 9th grade education, which was overrepresented in the VSD (8.3% in the VSD compared to 4.8% in the US) and the population with high school graduates, which was underrepresented in the VSD (45.1% in the VSD compared to 50.4% in the US).

A total of 629,837 VSD members participated in Medicaid-related programs in 2012, representing 1.1% of the US Medicaid population (Table 2). Despite a large number of total VSD Medicaid participants, the total proportion of the VSD Medicaid population (7.8%) was less than half that of the US Medicaid population (18.8%). The proportion of females was higher in both VSD and the US Medicaid populations; however, this difference was less marked in the VSD population (54.3% female in the VSD vs. 58.1% female in the US). Among different age groups in the VSD, children ages 6–18 years were overrepresented in the VSD (51.5% in VSD compared to 30.3% in the US). Additionally, the 19 to 44 year and 45 to 64 year age groups were underrepresented among VSD members compared to the US (16.4% vs 30.4% and 6.6% vs 12.1% respectively). There has been an increasing trend in the proportion of the VSD Medicaid population from 2005 to 2013 (figure 1) across different age groups, particularly among children 0–17 years old. The increasing trend in children is likely a result of the recent addition of local Medicaid-related programs included in the VSD Medicaid file that offer coverage to children.

Discussion

Post-licensure vaccine safety data monitoring is a crucial part of a successful vaccination program. As the VSD is a network that is used for high impact vaccine safety studies, it is important to know how its population may be different from that of the United States. Overall, we found that the VSD population is representative of the US population for a number of characteristics. Apart from small differences in age distribution, race/ethnicity, income and education levels of the overall population, and minor differences in the age distribution of the Medicaid population, there were no significant demographic differences to suggest a lack of generalizability in population based vaccine safety monitoring studies conducted in the VSD. Additionally, despite specific groups being underrepresented in the VSD compared to the US (i.e., African American, individuals with income less than \$50,000), the absolute number of VSD members is large enough to ensure significant representation of these groups in vaccine safety studies that use VSD data.

All major categories of race and ethnicity are well represented in the VSD. Certain groups, such as Asians and Hispanics, had a larger representation in the VSD than the US as a whole, which may be reflective of the large percentage of the VSD population at the California sites. For groups in the VSD which we found to be underrepresented compared with the US population, specific population sizes nevertheless were considerable: the African-American population consists of over 700,000 individuals, the Medicaid population

consists of over 600,000 members, and the population earning less than \$50,000 consists of over 2 million VSD members. Thus, the large number of individuals within these groups in the VSD facilitate sub-analyses of these populations being conducted in VSD studies. Additionally, separate large scale studies of these and other specific groups can also be done.

The VSD includes data on an insured population, and insurance is commonly provided by employers. Nonetheless, we found substantial coverage of older populations beyond the typical retirement age, which may be a result of the high Medicare population (federal health insurance program for individuals 65 and older) coverage in the VSD sites. With the introduction of the Affordable Care Act (ACA), we expect the VSD population to continue to change [23, 24]. Overall, we expect the population at the VSD sites to increase in diversity. We expect specific population increases to vary by VSD site, due to state differences in health insurance. For example, in some sites, the advent of health care exchanges and more high deductible health plans may attract both younger and lower income enrollees. Conversely, there may be some individuals who currently belong to Medicaid at specific sites who may decide to get insurance outside these VSD sites. Future studies comparing the VSD population to that of the US would be useful in evaluating the impact from the ACA, including the variation of this impact by VSD site.

Our study has a few limitations. In particular, we used data on income and education level derived from neighborhood estimates via the geocoding system. Although this method has been used in other studies [22], census tract/block geocode data may not accurately reflect the characteristics of a specific individual residing in the area. What we can conclude from the geocode data, however, is that the majority of VSD members tend to reside in areas that have distributions of education and income levels that are reflective of the US population. Another limitation is that the most recent US data available for our comparisons was from the 2010 US Census and from 2012 for Medicaid information. Both the VSD population and the US population are dynamic, and some of our data in this study may not be representative of these two populations today. For example, the VSD is actively investigating whether additional health systems may be incorporated to increase its diversity and representativeness with respect to socioeconomically disadvantaged populations, which could impact the future demographic representation in the VSD. Finally, we were only able to compare certain characteristics of the VSD to that of the US population, such as age, sex, race/ethnicity, and markers of socioeconomic status. There may still be differences that we did not capture in this analysis that may impact the generalizability of VSD vaccine safety studies. Future studies can compare other important areas, such as health care utilization, between the VSD and the US populations.

Conclusions

Our study shows that overall, the VSD well represents the US population with regard to sex, race, ethnicity, income, education and Medicare populations. Despite some differences that may affect specific studies, our data suggest findings from studies conducted in the VSD population are largely generalizable to the US population.

Acknowledgements

We gratefully acknowledge James Donahue, DVM, PhD, from the Marshfield Clinic for his review of the manuscript and all of the Vaccine Safety Datalink data managers for assistance in collecting data for this study.

Funding: The project described was supported by the Centers for Disease Control and Prevention (CDC) and Award Number T32AI074492 from the National Institute of Allergy and Infectious Diseases. The content is solely the responsibility of the authors and does not necessarily represent the official policy or position of the Centers for Disease Control and Prevention, the National Institute of Allergy and Infectious Diseases or the National Institutes of Health.

References

1. Roush SW, Murphy TV. Vaccine-Preventable Disease Table Working G. Historical comparisons of morbidity and mortality for vaccine-preventable diseases in the United States. *Jama*. 2007; 298:2155–2163. [PubMed: 18000199]
2. Cooper LZ, Larson HJ, Katz SL. Protecting public trust in immunization. *Pediatrics*. 2008; 122:149–153. [PubMed: 18595998]
3. Baggs J, Gee J, Lewis E, Fowler G, Benson P, Lieu T, et al. The Vaccine Safety Datalink: a model for monitoring immunization safety. *Pediatrics*. 2011; 127(Suppl 1):S45–S53. [PubMed: 21502240]
4. McNeil MM, Gee J, Weintraub ES, Belongia EA, Lee GM, Glanz JM, et al. The Vaccine Safety Datalink: successes and challenges monitoring vaccine safety. *Vaccine*. 2014; 32:5390–5398. [PubMed: 25108215]
5. Weintraub ES, Baggs J, Duffy J, Vellozzi C, Belongia EA, Irving S, et al. Risk of intussusception after monovalent rotavirus vaccination. *The New England journal of medicine*. 2014; 370:513–519. [PubMed: 24422678]
6. Shui IM, Baggs J, Patel M, Parashar UD, Rett M, Belongia EA, et al. Risk of intussusception following administration of a pentavalent rotavirus vaccine in US infants. *Jama*. 2012; 307:598–604. [PubMed: 22318281]
7. Belongia EA, Irving SA, Shui IM, Kulldorff M, Lewis E, Yin R, et al. Real-time surveillance to assess risk of intussusception and other adverse events after pentavalent, bovine-derived rotavirus vaccine. *The Pediatric infectious disease journal*. 2010; 29:1–5. [PubMed: 19907356]
8. Klein NP, Fireman B, Yih WK, Lewis E, Kulldorff M, Ray P, et al. Measles-mumps-rubella-varicella combination vaccine and the risk of febrile seizures. *Pediatrics*. 2010; 126:e1–e8. [PubMed: 20587679]
9. Tse A, Tseng HF, Greene SK, Vellozzi C, Lee GM. Group VSDRCAIW. Signal identification and evaluation for risk of febrile seizures in children following trivalent inactivated influenza vaccine in the Vaccine Safety Datalink Project, 2010–2011. *Vaccine*. 2012; 30:2024–2031. [PubMed: 22361304]
10. Verstraeten T, Davis RL, DeStefano F, Lieu TA, Rhodes PH, Black SB, et al. Safety of thimerosal-containing vaccines: a two-phased study of computerized health maintenance organization databases. *Pediatrics*. 2003; 112:1039–1048. [PubMed: 14595043]
11. Thompson WW, Price C, Goodson B, Shay DK, Benson P, Hinrichsen VL, et al. Early thimerosal exposure and neuropsychological outcomes at 7 to 10 years. *The New England journal of medicine*. 2007; 357:1281–1292. [PubMed: 17898097]
12. Kharbanda EO, Vazquez-Benitez G, Lipkind H, Naleway A, Lee G, Nordin JD, et al. Inactivated influenza vaccine during pregnancy and risks for adverse obstetric events. *Obstetrics and gynecology*. 2013; 122:659–667. [PubMed: 23921876]
13. Kharbanda EO, Vazquez-Benitez G, Lipkind HS, Klein NP, Cheetham TC, Naleway A, et al. Evaluation of the association of maternal pertussis vaccination with obstetric events and birth outcomes. *Jama*. 2014; 312:1897–1904. [PubMed: 25387187]
14. Kharbanda EO, Vazquez-Benitez G, Shi WX, Lipkind H, Naleway A, Molitor B, et al. Assessing the safety of influenza immunization during pregnancy: the Vaccine Safety Datalink. *American journal of obstetrics and gynecology*. 2012; 207:S47–S51. [PubMed: 22920059]

15. Institute of Medicine. The Childhood Immunization Schedule and Safety: Stakeholder Concerns, Scientific Evidence, and Future Studies. Washington (DC): 2013.
16. Wei F, Mullooly JP, Goodman M, McCarty MC, Hanson AM, Crane B, et al. Identification and characteristics of vaccine refusers. *BMC pediatrics*. 2009; 9:18. [PubMed: 19261196]
17. Larson K, Halfon N. Family income gradients in the health and health care access of US children. *Maternal and child health journal*. 2010; 14:332–342. [PubMed: 19499315]
18. Wilson K, Ducharme R, Hawken S. Association between socioeconomic status and adverse events following immunization at 2, 4, 6 and 12 months. *Human vaccines & immunotherapeutics*. 2013; 9:1153–1157. [PubMed: 23328278]
19. Koebnick C, Langer-Gould AM, Gould MK, Chao CR, Iyer RL, Smith N, et al. Sociodemographic characteristics of members of a large, integrated health care system: comparison with US Census Bureau data. *The Permanente journal*. 2012; 16:37–41. [PubMed: 23012597]
20. McCarthy NL, Irving S, Donahue JG, Weintraub E, Gee J, Belongia E, et al. Vaccination coverage levels among children enrolled in the Vaccine Safety Datalink. *Vaccine*. 2013; 31:5822–5826. [PubMed: 24135576]
21. McCarthy NL, Weintraub E, Vellozzi C, Duffy J, Gee J, Donahue JG, et al. Mortality rates and cause-of-death patterns in a vaccinated population. *American journal of preventive medicine*. 2013; 45:91–97. [PubMed: 23790993]
22. Chen W, Petitti DB, Enger S. Limitations and potential uses of census-based data on ethnicity in a diverse community. *Annals of epidemiology*. 2004; 14:339–345. [PubMed: 15177273]
23. Sommers BD, Buchmueller T, Decker SL, Carey C, Kronick R. The Affordable Care Act has led to significant gains in health insurance and access to care for young adults. *Health affairs*. 2013; 32:165–174. [PubMed: 23255048]
24. Blumenthal D, Collins SR. Health care coverage under the Affordable Care Act--a progress report. *The New England journal of medicine*. 2014; 371:275–281. [PubMed: 24988300]

Highlights

- The VSD population included over 8 million members in 2010
- The VSD is representative of the US population in many key demographic categories
- In underrepresented demographic categories, the VSD population remains substantial

Figure 1.

Percentage of VSD enrolled in Medicaid by age: 2005–2013*

*The VSD began collecting Medicaid data on the 18 and older population starting in 2009

Table 1Demographics of the United States and Vaccine Safety Datalink, 2010¹

	VSD	U.S
Total Population (N)	8,085,329	308,745,538
SEX		
Male	47.9	49.2
Female	52.1	50.8
AGE GROUP IN YEARS		
Under 5	5.9	6.8
5 to 9	6.4	7.3
10 to 14	7.1	7.3
15 to 19	7.4	7.2
20 to 24	5.6	6.7
25 to 34	12.3	14.2
35 to 44	13.7	16.0
45 to 54	15.4	13.4
55 to 59	7.2	4.8
60 to 64	6.2	3.8
65 to 74	7.3	6.5
75 to 84	4.0	4.4
85 and over	1.5	1.5
RACE		
White	64.7	72.4
Black or African American	8.8	12.6
American Indian and Alaska Native	0.6	0.9
Asian / Pacific Islander	15.4	5.0
Other races	7.1	6.2
Multiple races	3.3	2.9
ETHNICITY		
Hispanic or Latino (of any race)	36.5	16.3
INCOME IN 2010 ²		
Less than \$10,000	4.6	7.8
\$10,000 to \$14,999	3.1	6.0
\$15,000 to \$34,999	15.9	22.9
\$35,000 to \$49,999	12.4	13.9
\$50,000 to \$74,999	18.6	17.7
\$75,000 to \$99,999	14.6	11.4
\$100,000 to \$149,999	17.1	12.0
\$150,000 to \$199,999	7.2	4.5

	VSD	U.S
\$200,000 or more	6.5	3.9
HIGHEST EDUCATIONAL ATTAINMENT ^{2, 3}		
Less than 9th grade	8.3	4.8
9th to 12th grade, no diploma	8.1	8.9
High school graduate (includes equivalency)	45.1	50.4
Associate degree	8.1	8.6
Bachelor's degree	19.7	18.0
Graduate or professional degree	10.7	9.3

¹ Data represent percentages of total population unless otherwise noted. All percentages are rounded.

² Income and educational information from the VSD population is derived from neighborhood estimates.

³ All educational estimates include only individuals 18 and older, except for 2 sites which include 25 and older.

Table 2

Medicaid data summaries of U.S. and Vaccine Safety Datalink, 2012

	VSD	U.S.
Total Medicaid Population (N)	629,837	57,942,576
% of total population	7.8	18.8
SEX (%)		
Male	45.7	41.7
Female	54.3	58.1
Unknown	0.1	0.2
AGE, years (%)		
< 6	20.0	18.0
6–18	51.5	30.6
19 – 44	16.4	30.4
45 – 64	6.6	12.1
65+	5.5	8.8