MOCPHE | Public Health

MISSOURI CENTER FOR PUBLIC HEALTH EXCELLENCE

Prevent. Promote. Protect.

Public Health Data Disaggregation in Missouri

Missouri Center for Public Health Excellence

Authored By:
Suzanne Alexander, MPH, CPH, City of St. Louis Department of Health
Allie Bodin, MPH, CPH, Missouri Center for Public Health Excellence
Brennan Burk, MPH, St. Charles County Department of Public Health
Dr. Tom Burroughs, St. Louis University
Jordan Coiner, Springfield-Greene County Health Department
Carter Cook, MPH, Clay County Public Health Center
Sarah Crosley, Missouri Center for Public Health Excellence
Cara Erwin, MPS, Springfield-Greene County Health Department
Dr. Alex Francisco, KCMO Health Department
Matt Haslam, MSc, City of St. Louis Department of Health
Jessica Mchugh, MPH, St. Charles County Department of Public Health
Andrew Torgerson, MPH, St. Louis County Health Department
Ashley Wegner, MPH, Clay County Public Health Center

02.01.2013

Executive Summary

Data disaggregation refers to the breaking down of data into smaller groupings, often based on characteristics such as age, ethnicity, gender identity, income, location, race and sex. In public health, this practice allows agencies to better understand the individuals and groups they serve and conditions which may impact those persons disproportionately. Data disaggregation allows public health professionals to create appropriate, efficient and equitable interventions to improve health outcomes. Disaggregated data also enables public health agencies to evaluate programs and policies to ensure effectiveness.

While data disaggregation is not new, the COVID-19 pandemic revealed a greater need for the practice. For example, disaggregated data identified that certain populations, such as elderly individuals living in congregate settings, suffered greater risk of severe illness and death from COVID-19 than did elderly individuals living in single-family homes. Insights derived from disaggregated data enabled states like Missouri to design vaccine roll-out plans that prioritized the most vulnerable. As a result, public health professionals focused initial efforts on distributing COVID-19 vaccines to congregate care facilities in droves, resulting in an almost immediate reduction in the percent of COVID-19 deaths in that setting. Without disaggregated data, public health professionals cannot identify those sub-populations, create policies and interventions to reduce that risk, or evaluate the outcomes of those policies and interventions.

With the benefits of disaggregated data come risks and challenges that public health addresses. Improper disaggregation of data can result in stigmatization of populations, breaches of data security, and disclosure of identifiable or personal health information. Fortunately, current laws regulating public health and additional safeguards ensure safe and proper handling of data.

Overall, the benefits of data disaggregation outweigh the risks. Data disaggregation allows public health agencies to analyze emerging health trends, understand the needs of a particular population, identify potential logistical and operational problems, monitor equity, measure progress and ultimately use information to evaluate and improve programs, policies and systems resulting in better health outcomes for disparate populations.

Introduction

The Centers for Disease Control and Prevention (CDC) proposes a [vision for public health equity](#) which includes:

- One public health community that shares best practices and standards for collecting data on ethnicity, race, and social determinants of health.
- Broader access to integrated data that reflects real-time needs of local communities.
- Timely and complete data to help public health authorities distribute resources equitably, especially for emerging public health threats or environmental disasters.

Accessible, accurate, comprehensive and timely public health data at the local level protects communities by allowing appropriate, real-time resource allocation. Public health data allows for the development and monitoring of targeted interventions, identifies community health threats, and reveals health disparities. Harnessing the full power of public health data requires combining unique data sources (e.g., air quality data, asthma-related prescriptions, ER visits for asthma, neighborhood characteristics) and disaggregating aggregated data sources (e.g., federal health survey data, state-level health indicators) into more specific and smaller data sets to discern health disparities clearly at a group or community level.

Challenges exist which limit the impact and use of disaggregated data in practice. Inaccurate, incomplete or inconsistent data collection and sharing practices delay effective, timely responses to threats. When no data collection for specific groupings like education, ethnicity, race, and sexual orientation occurs, public health professionals may not apply appropriate interventions, impacting outcomes. Without disaggregated data, understanding complex health situations with concurrent risk factors, or examining challenges for individual groups, remains difficult. Few LPHAs consistently utilize disaggregated data currently.

Background

The Common Rule (45 CFR 164.501) legally designates local public health agencies (LPHAs) and the Missouri Department of Health and Senior Services as public health authorities. As partners, both remain responsible for improving and protecting community well-being by preventing disease, illness and injury and impacting economic, environmental and social factors fundamental to health and wellbeing.

Missouri's Foundational Public Health Services (FPHS) model describes the minimum services and capabilities that LPHAs should provide in all communities to ensure equitable opportunities for good health to all Missourians (Missouri Public Health Association, 2021). Capabilities include accountability and performance management, assessment and surveillance, communications, community partnership development, emergency preparedness and response, organizational administrative competencies, and policy development and support. LPHAs utilize these capabilities to offer the following services: chronic disease prevention, communicable disease control, environmental public health, injury prevention, linkage to medical, behavioral and community resources, and maternal, child, and family health. Per the model, all public health work occurs with a health equity lens (Missouri Public Health Association, 2021).

To conduct these responsibilities, the law [45 CFR § 164.512 (b)(1)(i)] authorizes DHSS and LPHAs to collect or receive disaggregated data to control or prevent disability, disease or injury. That authorization includes the reporting of disease, injury and vital events, such as birth or death, along with public health interventions, investigations and surveillance. This data comes

from numerous sources including the American Community Survey, emergency departments, environmental monitoring systems, federal surveys, hospitals, infectious disease surveillance systems, laboratories, local and state data sets, pharmacies, sexually transmitted infection monitoring systems and more.

Data Disaggregation Defined

Data disaggregation refers to the separation of compiled information into smaller segments to discern underlying patterns and trends. Disaggregated data may appear as individual line-level data used by public health practitioners before eventual aggregation into larger geographic levels or tables. The Pan American Health Organization establishes indicators disaggregated by age, disability, education, ethnic origin, geographic location, income/economic status and sex as the minimum set in *Fundamental Principles of Official Statistics* (Pan, 2020).

Examples of data disaggregation include dividing Asian Americans into cultural groups such as Chinese, Korean, Lao, Vietnamese, etc., looking at African American men by age groups, or separating LGBTQIA+ individuals into both gender identity and sexual orientation categories. Disaggregation of data by behavior, culture, or geography also occurs.

Why Disaggregated Data

From community-level statistics to information on individuals, multiple layers of useful disaggregated data exist. Data at the *individual level* allows for appropriate delivery of clinical services such as contract tracing, linkage to care, testing and treatment for sexually transmitted diseases and more. Databases organized by *demographic or geographic characteristics* reveal the needs of certain sub-populations, informing public health programming and evaluations of programs and services. *Community-level data* allows entire communities to assess priority health issues and evaluate the needs of larger geographic areas or jurisdictions. These layers all allow LPHAs to effectively and efficiently deliver services that advance equitable health outcomes to communities. Analysis of that data also measures progress and tracks trends for health outcomes over time.

The CDC, the National Network for Public Health Law, the Robert Wood Johnson Foundation, and World Health Organization all give significant attention and support to the topic of disaggregated data, emphasizing disaggregated data as a critical tool worldwide for designing effective health interventions and eliminating health disparities. According to the *Disaggregation of Public Health Data by Race & Ethnicity* legal handbook (2022), "The development and implementation of public health interventions that effectively reduce health inequities depends on the availability of accurate, representative, and timely health and demographic data" (p. 14).

Aggregated data gives the "big picture," but disaggregated data provides a "full picture" (TolaData, 2021). Not analyzing a health situation accurately may cause "society [to] pay a high opportunity cost when [failing to implement] interventions that yield the highest health return on an investment [ROI]" (Fielding, 2001). Policy makers and public health practitioners require disaggregated data to establish health problems, expose hidden trends, and identify high-risk populations. Such efforts then can inform effective and efficient interventions and policies. Disaggregating accurate, available, and complete ethnic and racial data to gain subpopulation health insights "not only contribute[s] to combating the health impacts of systemic racism and the unequal distribution of public health services [but also] improve[s] the nation's public health system and beyond" (*Disaggregation of Public Health Data by Race & Ethnicity*, 2022, p. 15).

Without granular data to identify high-risk subgroups for targeted vaccination, LPHAs struggle to direct intervention planning and implementation to yield the highest ROI. For example, data collection and reporting of demographics like gender identity and sexual orientation in the 2022 MPOX (formerly known as monkeypox) outbreak demonstrates how aggregation can hide disparities between subgroups. Analyses of the aggregated data may indicate only MSM (men who have sex with men) remain at high risk, but a closer look at disaggregated data reveals more high-risk subgroups such as those within the LBGTQIA+ community and as a subgroup, those who do not identify as MSM.

Disaggregated data informs the design of and helps evaluate interventions. Viewing aggregate data for geographic boundaries (e.g., county or ZIP code) can mask trends in health outcomes, giving an inaccurate community health status when evaluating interventions. While a 2% drop in county obesity rates *appears* to demonstrate improved outcomes and a successful intervention, if the change results from a 6% decrease among white residents and a 4% increase among Black residents (assuming equal populations), a notable disparity in trends exists—hidden without disaggregation. Analyzing data by demographics, location, and socioeconomic variables reveals the intervention as more effective for white community members, and a data-driven decision can help find a more effective intervention for Black residents. Insights gained in this way inform and improve public health programs, allowing local agencies to direct resource allocation optimally.

Finally, using data to understand the burden of disease and level of risk within subpopulations allows equitable distribution of funds and interventions. As such, disaggregated data maximizes the impact of taxpayer dollars to improve health outcomes and reduce health disparities in disproportionately affected populations.

Mitigating Risks and Challenges

Public health professionals and researchers face an ethical dilemma when using disaggregated data given the associated actual and perceived risks. Without proper risk mitigation, data disaggregation could result in stigmatization of populations, breaches of data security, and disclosure of identifiable or personal health information.

First, insights from disaggregating data may reinforce stereotypes or worsen bias (Lee-Ibarra, 2020). Identifying subgroups at risk for a socially undesirable health condition may stigmatize an often already marginalized group.

Additionally, disaggregated data can lead to inadvertent disclosure of confidential information if involving small samples, especially if the reader knows more information (e.g., data released by other agencies or personal knowledge). In a small community with very few Asian residents, if a public health agency reports a diagnosis of hepatitis C in a 52 year-old Asian woman, the potentially small number of people meeting that description could allow residents to infer the identity of the person (Washington, 2018).

While unable to eliminate the risk of inadvertent disclosure, public health professionals take steps to reduce that risk. In fact, bound and required by shared regulations, DHSS and LPHAs must observe all limitations and protections set forth by law that govern data confidentiality, privacy, sharing and use. Agencies working with confidential health information adopt policies for the use of direct and indirect identifiers as well as suppression of small numbers in aggregate data released to the public (Washington, 2018). Professionals handle confidential health information cautiously to prevent theft or inadvertent disclosure, following the Health

Insurance Portability and Accountability Act (HIPAA). Local and state public health agencies also follow policies and procedures to ensure employee compliance with relevant federal and state laws as well as maintain data security. Per Missouri law 19 CSR 20-20.075, LPHAs adopt and abide by confidentiality policies at least as stringent as state policies, and provide annual, comprehensive training on confidentiality and data security to all employees with access to data (Secretary, 2022).

Additionally, modern data infrastructures protect and restrict data access when used in tandem with data use agreements (DUAs). Updated data systems delineate access to authorized parties, giving the minimum data required to accomplish a task. DUAs establish how the use of data can occur and what safeguards each party maintains. DUAs not only institute privacy and publication rights but also denote liabilities for harm from misuse.

Case Studies

Examples of how public health professionals can use disaggregated data to improve equity and health outcomes within communities exist as follows.

Racial Disparities Among COVID-19 Cases in Kansas City

The COVID-19 outbreak in the United States shows the use of disaggregated data as a necessity in public health practice. Disaggregating cases by age, ethnicity, location, and race gives insight into health disparities during the pandemic. For example, the first wave of COVID-19 in Kansas City displays an age-adjusted case rate nearly five times higher for Hispanic residents than white residents. To this day, the age-adjusted death rate remains over twice as high in the Hispanic population relative to the white population.

With this information, the Kansas City Health Department (KCHD) demonstrates an inequity, resulting in the receipt of CDC grants to reduce disparities. The agency then targets vaccination efforts to Hispanic residents and, via more disaggregated data, demonstrates the vaccination rate of Hispanic residents outpaces the citywide average by 4%. By disaggregating data, KCHD identifies a specific population in high need—and then addresses that need via national funding and targeted programming. Without that insight, KCHD may fail to address areas of highest need first, potentially furthering inequities.

Varied Risk of Lung Disease in Springfield-Greene

Analyses of the intersections of housing data, income and smoking rates by census tract help the Springfield-Greene County Health Department determine areas of the community that may benefit most from more smoke-free housing options in multi-family properties. Disaggregating data may reveal a correlation between the cost of rent and secondhand smoke exposure among children, such as that seen in 2018. The lower the rent of the property, the more likely smoking remains permissible. Properties with the highest percent of children as residents appear the most likely to allow smoking. Those insights allow the department to focus interventions on low-income housing properties that can make the most impact by adopting smoke-free policies.

Life Expectancy Estimates in the United States

Disaggregating death data by age and location can estimate life expectancy across the nation. The U.S. Small-Area Life Expectancy Estimates Project (USALEEP) from the National Center for Health Statistics provides this information for nearly the entire United States—partly due to the inability of some jurisdictions to conduct these analyses without disaggregated data. Using

census tract level data results in a nuanced and targeted understanding of how health equity looks in small areas. These types of analyses provide the necessary information for highly targeted interventions to address health equity.

Conclusion

"The better data we have, the smarter the response will be" (Pan American Health Organization, 2020). Missouri ranks 42nd of 50 states and Washington, DC in per capita spending of public health funding (United Health Foundation, 2022) and 44th in per capita CDC grant funding (CDC, 2020). LPHAs in Missouri face difficult decisions when allocating scarce resources, highlighting a greater need for data-driven decision-making. Without disaggregated data, LPHAs feel the compounded impact of budget cuts and constraints: with many persons experiencing health disparities as a result.

To identify populations at high risk and/or in medically underserved communities, the public health workforce capacity must continue data disaggregation practices. To improve outcomes, agencies must direct investments toward vulnerable communities. To advance health equity, resource allocation needs to be data-driven. At a minimum, data disaggregated by demographic factors and social determinants of health remains necessary to design efficient and equitable policies and programs. That data informs LPHAs of health situations and problems accurately, allowing targeted and tailored resource provisions to improve health.

Without data representative of the health of all Missourians, LPHAs struggle to conduct evidence-based public health practice, failing to improve health disparities. The vast diversity of Missouri residents means that the lived experiences of people show greater specificity than information derived from one data metric. Discerning the whole picture to tell a complete story enables public health professionals to more equitably impact the lives of those served.

References

Lee-Ibarra, 2020. "Data Equity: What Is It, and Why Does It Matter?". Hawaii Data Collaborative Blog. https://www.jliconsultinghawaii.com/blog/2020/7/10/data-equity-what-is-it-and-why-does-it-matter

Missouri Public Health Association, "Missouri's Foundational Public Health Services Model." https://www.healthiermo.org/fphs-model

Missouri Secretary of State, 2022. "Rules of Department of Health and Senior Services: Division 20—Division of Community and Public Health. Chapter 20—Communicable Diseases." https://www.sos.mo.gov/cmsimages/adrules/csr/current/19csr/19c20-20.pdf

The Network for Public Health Law, 2022. "Disaggregation of Public Health Data by Race & Ethnicity." https://www.networkforphl.org/wp-content/uploads/2022/12/Network-for-Public-Health-Law-Data-Disaggregation-Handbook_FINAL.pdf

Pan American Health Organization, 2020. "Why Data Disaggregation is Key During a Pandemic." https://iris.paho.org/bitstream/handle/10665.2/52002/Data-Disaggregation-Factsheet-eng.pdf?sequence=19

TolaData, 2021. "Data Disaggregation & its key role in International Development" https://www.toladata.com/blog/data-disaggregation-and-its-key-role-in-international-development/

Washington Department of Health, 2018. "Department of Health Agency Standards for Reporting Data with Small Numbers." https://doh.wa.gov/sites/default/files/legacy/Documents/1500//SmallNumbers.pdf