

Why Count and Measure?

Justin Glasgow, MD

Introduction

The first recorded sign of counting in early humans is attributed to a collection of tally marks found on a baboon fibula in the Congo region of Africa, which dates to approximately 40,000 years ago. While it is unknown what was recorded (or if counting was the true intention), the recovered segment of bone has 29 tally marks, leading to a hypothesis that people were tracking the lunar phase. A health related hypothesis is that women were tracking menstrual cycles. The process of counting advances with time, and clearly is in use in 4,000 BC, with development of urban cities in Sumeria. Cities bring together a collection of people and resources in close proximity, necessitating a process to track and monitor people and resources. The Egyptians in 3,000 BC expanded on counting and developed measuring, which supports the ability to build pyramids and temples. The next major development, to the bane of many teenagers and perhaps their parents, was the Greeks development of more advanced mathematics, building the base for Algebra and Trigonometry.

In healthcare, perhaps the most famous early story of counting and measuring is that of John Snow, commonly considered a founding father of Epidemiology. His story is notable for the thoroughness with which he counts and tracks cases of cholera in an outbreak, and is able to show how they congregate around specific water pumps in London. Not only did this introduce the concept of epidemiologic case tracing, but perhaps also established the basis for studying how social determinants created health disparities.

Counting in Modern Healthcare

Modern healthcare seems to have an endless array of counts and measures. From the standard Complete Blood Count (CBC) to complex National Surgical Quality Improvement Program (NSQIP), healthcare providers and healthcare systems collect innumerable measures of health and healthcare. While the field strives to understand outcomes, it often has a limited or superficial understanding of the patients receiving care. With the Centers of Medicare and Medicaid Services (CMS) promoting electronic health records through the meaningful use (MU) incentives program, there has been some standardization of information collection. The 2014 Edition of the Stage 1 MU program incentivized the systematic collection of a patient's preferred language, gender, race, ethnicity and date of birth. This program has explicit definitions for race and ethnicity categories, but does not define expectations on gender collection. In fact, a summary document from CMS uses gender and sex interchangeably.

The development of the MU program precedes current recognition about risk of privacy invasion with data collection of large internet and social media companies, but within that context it is critical to ask whether it is necessary for healthcare systems and providers to collect this data. The stated purpose within MU is that the systematic collection of data elements will serve as a platform for understanding health disparities and driving efforts to improve quality, safety, and efficiency. This noble pursuit to understand and reduce health disparities is indeed necessary if we are going to continue to effectively improve the quality of health in our own neighborhoods and across the nation.

Health Disparities – Why Does Counting Matter?

Health disparities are any measure of higher burdens of illness, injury, disability or mortality experienced by one group compared to another. Health disparities exist in most everywhere that they have been explored. Amongst those who identify as a gender or sexual minority (GSM), disparities have been identified in a number of arenas. As a population, there are higher rates of tobacco, alcohol, and illicit substance use. Studies have also identified higher rates of many chronic conditions such as asthma, ischemic heart disease, hypertension, obesity, depression, and anxiety. And perhaps most unfortunate (but not uncommon): in settings where health disparities exist, GSM populations also have lower rates of seeking healthcare services.

Talking about health disparities in GSM populations as a single monolithic population ignores disparities that exist between various groups within the populations. Individuals identifying as transgender have starkly worse health behaviors and health outcomes than those who identify as a sexual minority. Due to the recent growth in recognition of individuals who identify as non-binary, there is little understanding of what specific health disparities they face, or how those differ from disparities previously identified among individuals identifying as transgender. Among those who identify as a sexual minority, when analyzed separately, individuals identifying as bisexual have greater health disparity than those who identify as the same gender but only identify as having same sex attraction (i.e. gay men have elevated rates of heavy drinking, but bisexual men have even higher rates of drinking).

Given good recognition regarding the existence of health disparities, is there truly a need to further collect data to assess health disparities? Answering this question requires some consideration of the scientific field that serves as the basis for studies on health disparities: epidemiology. Epidemiology is classically defined as the study of the diseases that affect a population. As the field has evolved, it more broadly examines how different factors impact the health of populations. Those trained in epidemiology build a strong foundation in study design and statistical analysis, which includes a particular emphasis on understanding the strengths and weaknesses of different study designs and analytic techniques. The popular media can make jokes about how one study says coffee will help you live longer while the next says coffee will kill you, but to an epidemiologist who has reviewed the research, they can often identify how different decisions in a study design can lead to these disparate findings. This is, of course, assuming that the splashy headline also doesn't just reflect an over simplification of the research study conclusions promoted to garner attention and drive views.

Epidemiology and Counting

There are many considerations epidemiologists use when evaluating research, here we will discuss two that can help us understand the importance of collecting structured data to help better analyze and understand health disparities. The concepts of internal validity and external validity are critical to understanding why GSM individuals (and quite frankly, anyone) should want to be accurately counted and should want to ensure that their health care organizations are consistently and systematically capturing appropriate demographic information.

Internal validity addresses how well a study measures the relationship of interest and whether it appropriately accounts for how other variables (called confounders) impact that relationship. For example, a study may wish to examine the relationship between GSM identity and utilization of emergency departments (ED) for care that can more effectively and efficiently be provided in a

primary care office (also called ambulatory care sensitive conditions). If a study only measured an individual's GSM identity and their rates of ED and primary care office utilization over a specific time frame, it would miss other critical confounders that impact why individuals choose to use one health care setting over another. A common confounder here is insurance status: GSM populations generally have higher rates of being uninsured, and being uninsured increases the likelihood of utilizing an ED for care over a primary care setting (as uninsured individuals generally do not have a longitudinal primary care relationship). So if a study finds an association between GSM identity and excess ED utilization, but does not account for differential rates of insurance, we cannot safely conclude that study found a true relationship: it would be considered to have weak internal validity. The role of unmeasured confounders is a common cause for why nutrition studies (i.e. what is the impact of coffee or chocolate) will sometimes show health benefits and other times health hazards.

In contrast, **external validity** addresses how well a sample of patients within a study represents the greater population at large. Issues with external validity are commonly found in many clinical drug trials. Given the high expense involved in clinical drug trials, these studies are designed to have high internal validity and promote the ability to find a benefit of the drug in question if it exists. However, this leads to excluding patients from the study population who may receive the drug once it becomes available widely in clinical practice. An easy example is that clinical trials rarely include pregnant women, so little is known about the effectiveness of many drugs in pregnant women. In many instances, the only information about risks of fetal teratogenicity is based on animal studies or several post-market birth defect registries. While the story of thalidomide occurred before (and is a primary driver for) the current paradigm of drug testing and approval, it demonstrates how external validity impacts study findings. Initial studies on thalidomide focused on the drug as a sedative, and found that it was essentially impossible to overdose on the medication. This led to its approval, and in some countries the medication was even sold without a prescription. However, those studies did not include many populations, particularly pregnant women. As its clinical use expanded from a sedative to being used to treat nausea (specifically morning sickness in pregnancy), this meant the early studies did not have sufficient external validity to address the safety of the medication. It was recognized that thalidomide was a fetal teratogen, and resulted in unknown numbers of miscarriages, as well as numerous birth defects, limb defects being the most famously linked.

In general, studies examining health disparities in GSM populations often suffer from major risks to both their internal and external validity. Internal validity is often weak because studies try to draw a simple line between GSM identity and either a health behavior or health outcome. For example, a study documenting increased rates of alcohol use among gay and bisexual identified men compared to heterosexual identified men also found higher reported rates of severe psychological distress among gay and bisexual identified men. This easily raises the question of how does severe psychological distress and alcohol consumption interact, and might that relationship confound the independent relationship with sexual identity (internal validity). It is certainly reasonable to hypothesize that psychological distress can lead to alcohol use as an (ineffective) attempt at coping, but similarly, alcohol is a depressant and could impact how individuals perceive their psychological distress. It is likely that alcohol and psychological distress confound the relationship of the other with sexual identity, however without a sufficient population size (external validity), the study could not statistically account for this possibility, leaving an open question for further studies.

The previous study data came from the National Health Interview Survey (NHIS), which suggests the study sample is representative of the large population (supporting generalizability), but that sort of study sample is rare in the literature evaluating health disparities in GSM populations. Frequently, studies utilize a sample of convenience, which helps develop preliminary findings, but limits how broadly findings should be applied. Common settings for identifying a large enough cohort of GSM identified individuals to participate in a study frequently means working in an urban setting and recruiting from settings that cater to the population, such as a gay bar or a community free clinic for sexual health services. It doesn't require much creative thinking to recognize that, when a study recruits from a bar, rates of alcohol and tobacco consumption (in the times when you could still smoke indoors) in that group may not reflect the behaviors of the greater GSM population. Similarly, those receiving care in an urban free clinic likely have a different healthcare experience than their urban compatriots with health insurance, whose healthcare experience may also be dramatically different from GSM individuals residing in a rural environment which may not have easy access to the same level of public health resources.

The Importance of Counting and Measuring GSM Populations

In the context of healthcare, and more importantly public health, why should we want to count and measure GSM populations accurately? In 2017, healthcare spending accounted for 17.9% of the Gross Domestic Product (GDP) or about \$3.5 trillion annually. Disparities in health will either directly or indirectly cost the country more over time than if we were providing everyone with the best high value care. The challenge is that if we have not accurately measured and assessed the health disparities, then our attempts to develop and implement interventions to reduce disparities are likely to fall woefully short of their aims. If we look back to the example of the relationship between GSM identity, psychological distress, and alcohol consumption, a poor understanding of this relationship is likely to lead to an ineffective intervention. An intervention focused only on achieving alcohol sobriety is unlikely to be successful if it does not address the underlying psychological distress that may be driving alcohol consumption. The intervention may achieve a brief period of sobriety in individuals, but since a key driver of alcohol consumption might not have been addressed, later events that increase stress will place an individual that lacks the adequate coping skills to address this driver at a higher risk of relapse.

Unfortunately, in many concepts of public health, the causal relationship between an individual's demographics and health outcomes are too complex for a simple survey to assess with sufficient internal or external validity. Simply put, a survey of 700 gay men (with an appropriate heterosexual comparison group) cannot sufficiently help us understand the complex relationship between sexual identity, psychological distress and substance use. Without large datasets that allow for careful and systematic study of complex relationships, there will be a persistent high risk of drawing the wrong conclusions about health disparities faced by minority populations.

How Can You Be Counted?

There are opportunities beyond collecting data during routine medical care that can also contribute to our efforts to understand and alleviate health disparities. For GSM populations, the PRIDE study (pridestudy.org) is the first long-term national health study of LGBTQ people with periodic surveys to assess physical, mental and social health. Additionally, the National Institutes of Health is supporting the All of Us study (allofus.nih.gov, Joinallofus.org) which represents an

effort to gather data from one million or more people, with the goal of collecting data from a broad representative sample of all individuals living in the United States. These programs represent opportunities to contribute to the health and wellbeing of all.

So why, as a GSM individual, should you want to share your identity with a healthcare provider or system? In a world of explicit discrimination and unconscious bias, there are countless barriers to wanting to share information with a provider. Long term, helping a medical provider understand the whole of your physical, mental and social health will support a strong patient provider relationship and drive effective diagnosis and management. In many instances, your identity (whether related to race, ethnicity, gender identity, or sexual identify) is unlikely to change the medical care you receive. Yet, as I think about common diagnoses, I have a hard time identifying a single diagnosis where a sexual or gender identity wouldn't ever potentially contribute to helping a provider develop an appropriate treatment plan. Common causes of symptoms may not be related to GSM identities, but if those are ruled out, knowledge about GSM identity may help formulate questions to better understand potential for more rare causes of symptoms.

Conclusion

To summarize, counting and measuring were established early in civilization as a way to help distill complex systems into an understandable process. The expansion of electronic health records provides an opportunity for well-structured interventions to address quality in healthcare and, more importantly, disparities in health. However, our current understanding of health disparities may be superficial, as it is often built on studies with significant limitations from an epidemiologic perspective when it comes to internal and external validity. In order to support improvements in public health (and consequently, help address unsustainable healthcare spending) there should be broad support for collecting data to help understand populations, such as gender and sexual minority patients, whether as part of routine healthcare or as participants in national studies.

Copyright (c) 2019 Delaware Academy of Medicine / Delaware Public Health Association.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc-nd/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.