Difference or Disparity? Will Big Data Improve Our Understanding of Sex and Cardiovascular Disease?

Karen E. Joynt, MD, MPH; Jessica L. Mega, MD, MPH; Michelle L. O’Donoghue, MD, MPH

Variations in clinical outcomes, such as mortality rates and complications, between men and women have been reported in cardiovascular disease for many years. However, whether these are the result of divergent underlying biology, which we often refer to as a difference, or whether they are the consequence of inadequate or inappropriate treatment, which we refer to as a disparity, remains unclear. Even the definitions of these 2 words are not universally agreed upon, and they are often used interchangeably. However, the distinction is important to inform the solutions clinical leaders and policymakers might put into place.

The major barrier to understand the proportion of variation that is because of different biology versus variation that is because of disparate treatment or differential response to treatment, particularly in complex real-world settings, is that we rarely have the opportunity to concurrently examine data that might elucidate each. However, our growing ability to leverage big data through novel analytics may begin to change that in meaningful ways. The objective of this piece is to present information on areas in which existing data suggest that there are differences, areas in which existing data suggest that there are disparities, and ways in which new big data sources and analytic techniques may allow us to disentangle the 2.

Potential Mechanisms

Women have long been reported to have ≤2-fold higher rates of in-hospital mortality in the setting of acute myocardial infarction compared with their male counterparts; this relationship is particularly powerful in women <50 years of age and when assessed in community settings. However, women are typically older when they present with cardiovascular disease, and studies suggest that much, though not all, of the increased mortality risk for women with acute myocardial infarction may be explained by clinical variables, such as age and a higher burden of comorbidities. It is thus possible that both difference and disparity contribute to the variation in outcomes we observe.

One piece of evidence that supports biological difference as a mechanism is the finding that women with stable angina and nonobstructive coronary artery disease have rates of recurrent hospitalization and cardiovascular events that are more than twice as high as men with similar coronary findings. This may suggest that chest pain in women, while less likely to be indicative of obstructive epicardial coronary artery disease, may instead be indicative of microvascular disease or alternate disease states that carry their own risk of poor outcomes. Women also have a higher risk of stroke in the setting of atrial fibrillation and a higher lifetime risk of any stroke. Taken together, these findings suggest that there may be differences in the underlying biology in women’s vascular beds than in those of men.

On the other hand, women might have worse outcomes than men in part because they receive lower quality care than their male counterparts; this would be referred to as a disparity. We know that women are less likely to receive cardiac catheterization in the setting of acute coronary syndrome, for example, and have longer reperfusion delays. These delays are likely explained by a failure to recognize and diagnose acute coronary syndrome in women rather than a conscious choice to deny treatment to women who warrant it but nonetheless constitute an important and potential remediable disparity in care. Women with cardiovascular disease are also less likely to receive a host of important outpatient therapies, including statins, renin–angiotensin blockade, and β-blockers, although these differences have been declining over time.

Disparity in cardiovascular outcomes in women may also be driven by a host of risk factors that are not typically evaluated in either biological or clinical research, namely socioeconomic and psychosocial factors. Women are more likely to live in poverty than men and tend to have higher rates of depression; there is some evidence to suggest that socioeconomic and psychosocial risk factors may have more powerful negative associations with cardiovascular outcomes in women. Few previous studies examining sex-based differences in care have accounted for these factors.
Finally, women may derive differential benefit from established therapies because they have been historically under-represented in phase III clinical trials. Although hypothesis-generating only, subgroup analyses of many cardiovascular trials, as well as meta-analyses combining these subgroups, have raised the question of whether women have a differential response than men to medications, such as glycoprotein IIb/IIIa inhibitors or device therapies, such as implantable cardiac defibrillators. As well, data are lacking to suggest that an invasive strategy is beneficial in low-risk women in the setting of acute coronary syndrome. Women have been demonstrated to have higher rates of bleeding in the setting of antiplatelet therapy and coronary intervention, which may also change risk-benefit calculations in the setting of both acute and stable coronary artery disease.

This may also make judging the quality of care delivered to women more difficult because reasonable people could differ in treatment patterns when there are inadequate data to inform decision-making. However, if a discrepant response to therapeutics indeed exists between men and women, the underpinnings remain elusive, as they could be explained by differences in drug pharmacology, underlying biology, or contextual factors, such as access to healthcare or psychosocial concerns.

Although many thought that true personalized medicine was on us when the first human genome was sequenced, it has proven to be much more difficult than we had hoped to predict someone’s medical future based on her genes alone. Similarly, although many thought that electronic health records would make real-time risk assessment and prediction of outcomes ubiquitous, our ability to determine who will or will not have a particular adverse event remains poor. One need look no farther than the popular number needed to treat website to see that our hopes of delivering highly effective medications to highly targeted populations is not yet realized. However, big data techniques may help change this.

**Promise of Big Data**

What is big data? Big data is a term used to describe collections of large datasets that were previously extremely challenging, if not impossible, to work with. Examples include the immense amount of data generated by electronic health records in the course of caring for a patient or repositories of millions of individuals generating billions of insurance claims, or genetic and genomic information not just on individuals but on populations. Many of the insights that could theoretically be gleaned from big data lie in waiting for our computational power and methodologies to catch up, but the process is beginning. The power lies in combining biometric, phenotypic, health service, and contextual data all at once. For example, the National Patient-Centered Clinical Research Network is a national effort to bring together detailed patient-level data from around the country in real time. The rapidly growing use of biorepositories that are linked to electronic health records and other rich sources of phenotypic data can be seen in such efforts as the UK biobank, PatientsLikeMe and other social networking medical sites, Health eHeart, and the American Heart Association’s Genome-Phenome Study.

So, how might big data help us understand the relationship between sex and cardiovascular disease? It may, for the first time, allow us to understand how biology, quality, and context interact with one another to produce variations in cardiovascular health and outcomes. Only in bringing huge data sources, such as biometric data, clinical trial data, and health service data, together, we can begin to solve this puzzle, and in doing so, we can begin to sort out why men and women differ in their response to medications, adverse event rates, and clinical outcomes.

For example, take the finding that women, particularly older women, have higher rates of stroke in the setting of atrial fibrillation than men. If this is because of differences in underlying vascular biology, we could potentially reduce strokes in women with new targeted therapeutics. If this is because of different genetically based response to anticoagulants, we could reduce strokes with different dosing strategies in women for these medications. If this is because of differential quality of maintaining therapeutic anticoagulation, we could reduce strokes via better systems to improve monitoring. And if this is because of higher rates of poverty and less ability to afford medications, we could reduce strokes by improving access to affordable care for this population. In reality, it is probably because of all of these things and others. But any study that only looks at one factor in isolation will likely be confounded, may overestimate the effect of that factor, and could misdirect our efforts to reduce the incidence of strokes in women. Only with a broader approach to data, we can gain precision in our findings.

It is, of course, also worth noting the limitations to the use of big data; one of which is the existence of appropriate and novel analytical methods. Because of the enormous sample sizes generally at play, statistical significance must be handled differently from when dealing with smaller datasets, and techniques, such as data mining, machine learning, and cluster analysis, will play larger roles. Big data and their attendant methods introduce their own limitations, however, most importantly ensuring data quality and consistency, sorting signal from noise, and inferring causal inference only where appropriate. Another important limitation to the use of big data is maintaining privacy and security while linking patient-level data across multiple platforms (such as genetic and health service data or socioeconomic and biomarker data).

If these limitations could be overcome, it is feasible that the insights generated by big data techniques could be translated into meaningful changes in clinical practice. Ultimately, we should hope to move toward a paradigm in which a personalized risk assessment, incorporating behavior and evidence and biology, can help us make ideal clinical decisions at the point of care. For an 89-year-old woman with rheumatoid arthritis and renal insufficiency admitted with an inferior MI, using genes alone or clinical trial results alone, or health services-driven guidelines alone is not enough. Big data could allow us to go beyond personalized medicine in the usual sense, targeting treatment based on our patient’s genetic ability to respond to medications, and also taking into consideration high-quality clinical trial data, suggesting that elderly women are better-served with a particular combination of antiplatelet
therapies, or prescribing a medication developed specifically to address small-vessel coronary artery disease in women, or using tools developed by health services researchers to more accurately convey risk and inform patient preferences and integrating these tools in real time. And where evidence does not yet exist to inform patient care, we can embed clinical tri- als within these large data structures and leverage this wealth of data to generate new insights.

Conclusions
Variations in clinical outcomes between men and women likely represent both biological differences as well as disparities. Insights garnered from the merging of health services, clinical trials, and genetic or genomic data have the potential to help us sort out which is which, and in doing so, they help us begin to target our therapies in the most efficient and equitable way possible. For our female patients, this could lead to better, more appropriate care. For clinicians, it may mean increased decision support at the point of care, novel therapeutics, and better evidence on which to rely when making complex decisions for populations for whom evidence may have been lacking in the past. For policymakers, better ability to identify when a true disparity exists will allow the development of policies that effectively target these inequities while improving overall quality. Overall, big data, if harnessed appropriately, may have the ability to change the care we deliver and improve our patients’ clinical outcomes.

Disclosures
Dr Mega has received research support from Janssen/Bayer, Daiichi Sankyo, and AstraZeneca and serves as a consultant to Janssen, Bayer, Portola, and AstraZeneca. Dr O’Donoghue has received research support from Merck, GlaxoSmithKline, Eisai, and AstraZeneca. Dr Joynt reports no conflicts.

References


35. UK Biobank. 2015.


37. Health eHeart. 2015.


**Key Words:** healthcare disparities ■ quality of healthcare
Difference or Disparity: Will Big Data Improve Our Understanding of Sex and Cardiovascular Disease?
Karen E. Joynt, Jessica L. Mega and Michelle L. O'Donoghue

Circ Cardiovasc Qual Outcomes. 2015;8:S52-S55; originally published online February 24, 2015;
doi: 10.1161/CIRCOU115.001701
Circulation: Cardiovascular Quality and Outcomes is published by the American Heart Association, 7272 Greenville Avenue, Dallas, TX 75231
Copyright © 2015 American Heart Association, Inc. All rights reserved.
Print ISSN: 1941-7705. Online ISSN: 1941-7713

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://circoutcomes.ahajournals.org/content/8/2_suppl_1/S52

Permissions: Requests for permissions to reproduce figures, tables, or portions of articles originally published in Circulation: Cardiovascular Quality and Outcomes can be obtained via RightsLink, a service of the Copyright Clearance Center, not the Editorial Office. Once the online version of the published article for which permission is being requested is located, click Request Permissions in the middle column of the Web page under Services. Further information about this process is available in the Permissions and Rights Question and Answer document.

Reprints: Information about reprints can be found online at:
http://www.lww.com/reprints

Subscriptions: Information about subscribing to Circulation: Cardiovascular Quality and Outcomes is online at:
http://circoutcomes.ahajournals.org//subscriptions/