Opinion

Big Data?
When a Mere Datum Won't Do

The term Big Data always seems to be capitalized, as if it were Bigger than can be adequately expressed in lowercase. Just as Big Steel and Big Oil imply their domination through the use of capital letters, Big Data sounds impressive and demanding of attention. It's a new term for many of us in medicine, but astrophysicists have been wallowing in terabytes for much longer. It refers to the aggregation of large amounts of data and the analysis of them to find new relationships and new answers to complex questions. Now the term is popping up in the biological sciences, too, because we have the computational power to handle massive amounts of information and electronic data generators to supply the volume. And clinical registries have begun to aggregate patient data, by diseases, to improve care in clinics and accountable care organizations. Medical societies are getting on board the data express, too.

But Big Data has taken on some of the negative connotations of Big Government. It is viewed as a resource hog, stealing critical funding from carefully controlled investigations that use small data sets and increasing overall expenditures. It is also viewed by some as a statistical quagmire. The number of subjects is so large in these data sets that statistical significance can be conferred on trivial chance relationships. The data rarely arrive from their sources in a "clean" state—that is, there may be systematic errors related to one site of origin versus another; other problems may be as simple as coding mismatches. And there is likely to be bias, whether introduced by selection of the individuals on whom the data are collected or by the questions asked by the investigators.

Why, then, is the Academy launching in 2014 its new Big Data project, the IRIS Registry (Intelligent Research in Sight)? The answer is simple: It brings the potential for huge improvements in quality of patient care. This will be accomplished in an environment of complete privacy, both for the patients and for the participating ophthalmologists. The IRIS Registry will automatically harvest data from the electronic health records of each practice; little time will be required from the Academy member or office staff. Participation will be voluntary, but a critical mass of participants will ensure its success. Certainly, there will be questions about data analysis validity and errors, but because of the environment of privacy, these questions will be addressed by data validation and comparison to other studies.

Think of how you might use your data (which, by the way, you own) within the IRIS Registry. You could compare your rate of capsular rupture with the national statistics. If you are above average, as most of us are, you can take some pride in your skills. If you are not, it could be that you have a high-risk group of patients or other extenuating circumstances. But it could mean that a refresher course is in order. It's your choice.

If you're going to the Annual Meeting, you'll have ample opportunity to learn more about the IRIS Registry (see page 61). This Opinion hasn't even mentioned several other important benefits of participation—explore them in person in New Orleans or online at www.aao.org/irisregistry.

For those who say that the data from a registry are likely to be flawed, I'd like to point out our current clinical decision making is based largely on anecdotes. Surely we can do better than a datum at a time.

RICHARD P. MILLS, MD, MPH
CHIEF MEDICAL EDITOR, EYENET