

# ELECTRONIC MEDICAL RECORDS AS A RESEARCH TOOL:

## THE OPPORTUNITY AND RISK OF BIG DATA

Lawrence I. Bonchek, M.D., F.A.C.C., F.A.C.S.

*Editor in Chief*



### INTRODUCTION

When Electronic Medical Records were first being promoted widely, their full implications were not yet well understood. Back then, the biggest challenge for EMR advocates was simply the need to explain what they were, how they worked, and what they could do, and we published an introductory article to discuss those matters.<sup>1</sup> It took a while for physicians to learn how to use EMRs, and to decide whether they could enhance quality, efficiency, and safety on the clinical front lines enough to justify the training time and expense needed to incorporate them into a practice. After more than a decade it can only be said that many benefits are apparent, but many remain elusive, and inefficiencies persist.

Regardless, however, their adoption is no longer optional. Overriding all other influences, the Accountable Care Act has imposed a variety of carrots and sticks in the form of financial incentives and penalties that have in some ways encouraged, and in other ways mandated, conversion to EMRs. As a result, more than 80% of physicians now use EMRs to an important degree, and the percentage is highest (86%) among Family Physicians.<sup>2</sup> Clearly, there is no turning back.

### THE VALUE OF EMRS AND BIG DATA FOR CLINICAL RESEARCH

With that brief background, I'd like to set aside the well-worn debate about the benefits and costs of EMRs in clinical practice, and instead discuss the use of EMRs as a research tool. Their value for research was perhaps less obvious when they were first introduced, but in our current era of Big Data, EMRs will likely have an enormous impact in the long run.

The term "Big Data" generally describes data sets too large or complex to be processed by conventional (usually desktop) software. Because processing capabilities are also advancing continuously, the definition of

"Big Data" constantly changes, but at last count it was at least several dozen terabytes. Big Data sets of clinical variables are becoming more and more common, in part because data can be gathered so cheaply and conveniently, even on mobile devices. As an example, The Society of Thoracic Surgeons National Database for Cardiac Surgery was the world's first large-scale clinical registry for a specialty, and has been gathering clinical data on cardiac surgery in the U.S. since 1989. It now contains data on more than 5 million procedures, and has more than 3,000 surgeons who participate voluntarily. Over the years it has produced a torrent of studies and articles based on analyses of information in the database.\*

When it comes to health care data, it's worth recalling that even before EMRs came into common use, we had become accustomed to studies based on Big Data from Medicare's database. However, though the CMS database is the granddaddy of medical Big Data, it provides only administrative data, i.e. information relevant to reimbursements. These data include admission and discharge dates, diagnoses, procedures, sources of care, as well as demographic data such as age, date of birth, race, place of residence and date of death. Lacking, however, are the clinical details and correlations that can reveal subtle clinical associations.

EMRs, with their vast numbers of clinical details, fill in those gaps and thus offer the opportunity to detect seemingly countless associations that may have clinical importance and open up new avenues of diagnosis and treatment.

The recent discovery that beta-blockers are associated with a significant increase in survival of patients with ovarian cancer<sup>3</sup> is they type of association that will be more likely in the era of Big Data.<sup>4</sup> In this case, rather than being found unexpectedly, the key association was specifically looked for on the basis of prior evidence that adrenergic activity can influence the immunologic microenvironment and hence tumor

\* The cardiac surgery program at LGH initiated its own data form and database with its first case in 1983. When the STS National Database started in 1989, we transferred all our data to the STS Database, and we have been participants ever since.

growth. A retrospective review of “electronic and paper charts” at four institutions yielded only 1,425 patients who met study criteria and were treated over a 10-year period. An accompanying editorial pointed out that larger, prospective clinical studies will be necessary (and are underway) to validate this finding and determine its applicability to clinical practice.<sup>5</sup>

Consider, if you will, the difference if this analysis could have been carried out using EMRs from not four, but – say – 1,000 institutions. There would have been tens of thousands of eligible patients, more accurate analysis of confounding variables, and conclusions that would likely have been definitive. Thus, for all the excitement surrounding this study’s findings, its failure to provide conclusive findings demonstrates, albeit indirectly, the potential role of EMRs as a clinical research tool that can reveal unexpected associations of clinical variables that are not present in administrative data alone, or remain obscured in the small databases generated by manual chart reviews unless they are specifically looked for.

#### THE CHALLENGES OF BIG DATA

In an era when previously unsuspected associations will be revealed by computerized research, we will face a new challenge as we confront the simple fact that association does not mean causation. As we find new associations by mining Big Data, we will need to distinguish those that are meaningful. Of course, we hope that the statistical advantages of Big Data will make it unnecessary to conduct large randomized clinical trials

to determine the validity of every single association we find. But even if the statisticians can save us from the most egregious errors, we are likely to waste time and resources going down many blind alleys.

Also, there will still be the problem that negative studies are less likely to be published than positive findings that confirm a new association. So, even if the first report of a new “association” is accompanied by appropriate disclaimers, and a subsequent study fails to confirm its findings, the negative study may not be accepted for publication or may not achieve high visibility, whereas the original report will be online forever, often with a title that implies benefit.<sup>5</sup>

#### SPECIAL NOTE ON THIS ISSUE:

As a result of our affiliation with Penn Medicine and with CHOP (The Children’s Hospital of Pennsylvania), we have begun soliciting articles for JLGH from the faculty at the University of Pennsylvania and CHOP. Their pediatric cardiologists are fulfilling an important need for us, and have already established a regular schedule for seeing patients in Lancaster. In this issue we feature an article on congestive heart failure in children by Dr. Matt O’Connor. It’s of some interest that Dr. O’Connor is the brother of Christopher O’Connor, Esq., who was an associate general counsel at LGH until recently when he left for a position in another state. Chris was a regular and highly valued contributor to the Journal on medicolegal affairs. We look forward to a long and productive relationship with our new colleagues.

#### REFERENCES

1. Ripchinski, MR and Eichelberger, DO. What can electronic medical records do for you? How EMR technology can affect the processes of clinical care. *J Lanc Gen Hosp*. 2008; 3: 119-124
2. Office of the National Coordinator for Health Information Technology. Data Brief No. 28; September 2015. [https://www.healthit.gov/sites/default/files/briefs/oncdatabrief28\\_certified\\_vs\\_basic.pdf](https://www.healthit.gov/sites/default/files/briefs/oncdatabrief28_certified_vs_basic.pdf)
3. Watkins JL1, Thaker PH2, Nick AM, et al. Clinical impact of selective and nonselective beta-blockers on survival in patients with ovarian cancer. *Cancer*. 2015 Oct 1;121(19):3444-51. doi: 10.1002/cncr.29392. Epub 2015 Aug 24.
4. Agus DB. *The Lucky Years; how to thrive in the brave new world of health*. Simon and Schuster, New York. 2016
5. Bunch KP, Annunziata CM. Bunch KP1, Annunziata CM1. Are beta-blockers on the therapeutic horizon for ovarian cancer treatment? *Cancer*. 2015 Oct 1;121(19):3380-3. doi: 10.1002/cncr.29394. Epub 2015 Aug 24.