

# Biases in electronic health record data due to processes within the healthcare system: retrospective observational study

*BMJ* 2018; 361 doi: <https://doi.org/10.1136/bmj.k1479> (Published 30 April 2018)

## Abstract

**Objective** To evaluate on a large scale, across 272 common types of laboratory tests, the impact of healthcare processes on the predictive value of electronic health record (EHR) data.

**Design** Retrospective observational study.

**Setting** Two large hospitals in Boston, Massachusetts, with inpatient, emergency, and ambulatory care.

**Participants** All 669 452 patients treated at the two hospitals over one year between 2005 and 2006.

**Main outcome measures** The relative predictive accuracy of each laboratory test for three year survival, using the time of the day, day of the week, and ordering frequency of the test, compared to the value of the test result.

**Results** The presence of a laboratory test order, regardless of any other information about the test result, has a significant association ( $P < 0.001$ ) with the odds of survival in 233 of 272 (86%) tests. Data about the timing of when laboratory tests were ordered were more accurate than the test results in predicting survival in 118 of 174 tests (68%).

**Conclusions** Healthcare processes must be addressed and accounted for in analysis of observational health data. Without careful consideration to context, EHR data are unsuitable for many research questions. However, if explicitly modeled, the same processes that make EHR data complex can be leveraged to gain insight into patients' state of health.

## Reviewer: 4 - Patient and Public Reviewer

Comments:

This study is important and--as far as I know--groundbreaking, at least in scope. From the perspective of this patient-reviewer, it holds particular promise for improving patient outcomes and both public and private education about healthcare.

The following points were of particular interest: 1) Using PP and especially HSD data can tell us what clinical trials do not and can not. (In fact, a bit more could have been made about this very important point.) 2) Though all the data came from only two in many ways quite similar hospitals in essentially the same location (an acknowledged limitation), the applicability of this approach to hospitals (of all kinds and sizes) elsewhere and even to non-hospital settings (rehab centres, outpatient clinics, primary care practices, etc.) is clear and tantalizing. 3) The instances where PP data are better predictors, where HSD data are better predictors, and where the combination of the two is a better predictor are well documented and fairly well explained. 4) The case for treating HSD data as much more than noise is well made without being repeated too much.

The following are some suggestions and concerns:

1) I understand no patients were included in the design of the study and I have a sense of why that would have been difficult if not impossible, but I can see a huge potential value in involving them in further such studies, especially in choosing outcomes that matter most to them and in identifying ways the findings can help in public and patient education.

2) Perhaps data related to patients' SES, their co-morbidities, etc. could also be captured and used. This would make the findings of even greater value to clinicians, health economists, and policy makers.

3) While I understand why PP and HSD data "can be used as proxies for each other in some instances" (p.14), I think some acknowledgement (preferably with an example or two) that they can't always be used as proxies for each other would be a useful caution.

4) The statement about "more accurately aligned incentives for both patients and providers that result in better health" (p. 14-15) could use a good example. While reading the article, I was struck at least a few times of its relevance to the "Choosing Wisely" campaign here in Canada. I suspect it will ring different but related bells for readers in other countries.

Additional Questions:

Please enter your name: Frank Gavin

Job Title: 1) Public Member, Canadian Drug Expert Committee and 2) Board Member, Ontario SPOR Support Unit

Institution: 1) Canadian Agency for Drugs and Technologies in Health and 2) Ontario Strategy for Patient-Oriented Research Support Unit

Reimbursement for attending a symposium?: No

A fee for speaking?: No

A fee for organising education?: No

Funds for research?: No

Funds for a member of staff?: No

Fees for consulting?: No

Have you in the past five years been employed by an organisation that may

in any way gain or lose financially from the publication of this paper?: No  
Do you hold any stocks or shares in an organisation that may in any way  
gain or lose financially from the publication of this paper?: No  
If you have any competing interests (please see BMJ policy) please declare them  
here:

Note: Accompanying reviews for this paper can be found at:  
<https://www.bmj.com/content/361/bmj.k1479/peer-review>