

An evaluation of electronic laboratory data quality and a health information exchange

Brian E. Dixon^{1,2*} and Shaun J. Grannis^{2,3}

¹Indiana University School of Informatics, Indianapolis, IN, USA; ²Regenstrief Institute, Indianapolis, IN, USA; ³Indiana University School of Medicine, Indianapolis, IN, USA

Objective

To examine the completeness of data submitted from clinical information systems to public health agencies as notifiable disease reports.

Introduction

Electronic laboratory reporting (ELR) was demonstrated just over a decade ago to be an effective method to improve the timeliness of reporting as well as the number of reports submitted to public health agencies (1). The quality of data (including completeness) in information systems across all industries and organizations is often poor (2), and anecdotal reports in the surveillance literature suggest that ELR may not improve the completeness of the data in the submitted reports (3).

Methods

The scope of our research included the following aims: (1) the development of a method for evaluating the completeness of laboratory data in the context of public health reporting; (2) measuring the completeness of laboratory data received from clinical information systems and an HIE using the method; and (3) comparing the completeness of the 'raw' data from clinical information systems (e.g., unaltered and unedited ELR messages) with the completeness of 'enhanced' data from the HIE (e.g., ELR messages having syntax corrected and concepts mapped to standard vocabularies).

Results

A comparison of 7,592,039 raw messages and 16,365 enhanced messages revealed a number of differences with respect to data completeness. Data field completeness within the ELR messages varied from 0.01% to 84.6% across the two samples. Completeness was generally higher in the enhanced message sample as shown in Table 1.

Conclusions

To effectively perform surveillance, public health agencies require access to 'timely, accurate, and complete data' (4). Unfortunately, data quality is an issue for many clinical information systems that capture data utilized in public health surveillance processes. This study assessed the completeness of real-world ELR data from

multiple provider organizations using a variety of laboratory information systems, documenting evidence that ELR data are heterogeneous in their completeness across and within information systems. In many cases, data important to public health surveillance processes are missing, indicating suboptimal ELR data quality. The study further documented evidence that a statewide or regional HIE can employ methods to mitigate ELR data deficiencies, leading to improvements in the completeness of ELR data prior to transmission to public health agencies.

Keywords

Electronic laboratory reporting; health information exchange; data quality; completeness

Acknowledgments

This study was supported, in part, by the Indiana Center of Excellence in Public Health Informatics through a grant award (1P01HK000077-01) from the U.S. Centers for Disease Control and Prevention.

References

1. Effler P, Ching-Lee M, Bogard A, Jeong MC, Nekomoto T, Jernigan D. Statewide system of electronic notifiable disease reporting from clinical laboratories: comparing automated reporting with conventional methods. *JAMA*. 1999;282:1845-50.
2. Wang RY, Strong DM. Beyond accuracy: what data quality means to data consumers. *J Manage Inform Syst*. 1996;12:5-34.
3. Overhage JM, Grannis S, McDonald CJ. A comparison of the completeness and timeliness of automated electronic laboratory reporting and spontaneous reporting of notifiable conditions. *Am J Public Health*. 2008;98:344-50.
4. Lombardo JS, Buckeridge DL, editors. *Disease surveillance: a public health informatics approach*. Hoboken: John Wiley & Sons; 2007.

*Brian Dixon

E-mail: bdixon@regenstrief.org

Table 1. Completeness of data fields in ELR messages before and after enhancement by a health information exchange.

Key data element	Percent complete raw	Percent complete enhanced	Difference
Patient's identifier	99.9%	100%	+0.01%
Patient's name	99.4%	100%	+0.06%
Patient's date of birth	97.8%	99.8%	+2.0%
Sex (gender)	95.8%	99.9%	+4.1%
Race	38.4%	60.3%	+21.9%
Patient's address	41.5%	63.3%	+21.8%
Patient's home phone number	38.5%	72.8%	+34.3%
Ethnicity	3.5%	18.3%	+14.8%
Name of attending physician or hospital or clinic or submitter	57.4%	66.5%	+8.9%
Telephone number of attending physician or hospital or clinic or submitter	0.15%	73.3%	+73.2%
Address of attending physician or hospital or clinic or submitter	NA	84.6%	+84.6%