

Measuring the Impact of Health Information Technology

D. Hanauer¹; K. Zheng^{2,3}

¹Department of Pediatrics, University of Michigan Medical School, Ann Arbor, MI;

²School of Public Health Department of Health Management and Policy, University of Michigan, Ann Arbor, MI;

³School of Information, University of Michigan, Ann Arbor, MI

Correspondence to

David A. Hanauer
5312 CC, SPC 5940
1500 E Medical Center Dr.
Ann Arbor, MI 48109
Phone: 734-615-0599
Fax: 206-338-4213
E-mail: hanauer@umich.edu

Appl Clin Inf 2012; 3: 334-336

doi:10.4338/ACI-2012-06-LE-0025

received: June 21, 2012

accepted: August 30, 2012

published: September 12, 2012

Citation: Hanauer D, Zheng K. Measuring the impact of health information technology. *Appl Clin Inf* 2012; 3: 334-336

<http://dx.doi.org/10.4338/ACI-2012-06-LE-0025>

To the Editor

We read with great interest the article by Litwin, Avgar, and Pronovost discussing the multiple ways in which current studies of health information technology (HIT) may inadequately capture relevant measures. Lacking appropriate details could result in biased conclusions and may explain why some studies have reported negative perceptions or suboptimal outcomes of the technology being implemented [1].

With respect to the concept of 'time horizons' (i.e., the time elapsed from HIT implementation to measurement), we would like to propose an additional perspective to consider: while organizations should be able to optimize a system over time to achieve desired effectiveness, it may not always be the case that "studies with shorter time horizons are likely to yield weaker results." Alert fatigue and other factors can set in and begin to undermine the effectiveness of a system [2], resulting in waning usage and satisfaction.

Alert fatigue occurs when too many (often inconsequential) warnings or messages are presented to a user, resulting in interrupted workflow and an increased mental burden, ultimately driving the user to ignore the alerts, even those that are clinically important [3]. The current literature suggests that alert fatigue may develop over time. For example, Embi and Leonard recently reported on the gradual onset of alert fatigue for patient recruitment alerts over a 36-week period [4]. Another study of allergy alerts showed a steady decline in adherence to these warnings over a five-year period [5]. Therefore, single snapshots in time may not show the temporal trends of system usage that may only become apparent with longitudinal measures. We have observed system usage fade substantially in as little as 3 months as users lost interest in interacting with a system [6]. This dynamic was also evident in a study of a clinical reminder system in which usage over time increased for some users and decreased for others [7].

Litwin et. al also noted that "studies must allow enough time for the organization to make necessary changes and to undergo necessary learning around new technologies before taking a 'post-' health IT performance measure." This is often referred to as the 'optimization period,' a time in which multiple changes occur, often to both the system itself and external workflows, to ensure that every-

thing functions as intended. This line of reasoning can be found in multiple studies of health IT. For example, a study of drug interaction and allergy alerts measured the incidence of overridden alerts 9 months after a computerized provider order entry implementation to “allow clinicians ... to become accustomed to the upgrade” [8]. Another study of specialty care providers using a newly implemented electronic health record waited to conduct their observations 6 to 9 months post-implementation in order to “allow practice habits to stabilize” [9].

We certainly agree that collecting measurements is important only after a system is considered ‘stable,’ but we wonder if the management literature provides any insight regarding the selection of an ideal time period to wait between implementation and measurement. A rule of thumb we have often seen among HIT evaluation studies is to wait at least 3 months, but we know of no such consensus, nor do we claim that a consensus is even practical, possible, or desired among the research community. A recent analysis we conducted of HIT evaluations that used observational time and motion approaches uncovered tremendous variability in the post-implementation time frames for measurements [10]. First, nearly 20% of the studies that could have reported the ‘time horizon’ (also described as the ‘intervention maturity’) did not actually report when they began to carry out their observations with respect to the HIT implementation. Second, among those that did report the measure, the actual timing ranged from as little as 2 months [11] to as much as 3 years later [12].

This lack of standard reporting and comparable time frames may hinder the generation of useful insights. To that end, we would like to point out recent work of ours in which we outlined a minimum set of data elements related to measuring the impact of HIT implementation on clinical workflow, known as the ‘Suggested Time And Motion Procedures’ (STAMP) [10]. STAMP describes nine overall areas with 33 distinct elements that we believe should be reported in time and motion studies so that salient study design characteristics can be compared effectively. While some of the areas we suggested are specific to time and motion studies, such as training of observers and categorization of clinical activities to be observed, other areas are more broadly applicable and could be generalized to a larger set of study methodologies. These include details about the intervention itself such as the type of system studied (e.g., computerized prescriber order entry or electronic health record), the system genre (e.g., commercial or homegrown), and intervention maturity. Other details include descriptive characteristics of the empirical setting including the institution type (e.g., academic or non-academic center) and care areas (e.g., inpatient or outpatient). We found that studies in the current literature often reported these details in an inconsistent manner, if at all.

The reporting guidelines in STAMP are in alignment with other initiatives to standardize research conduct and reporting in health informatics, including the Statement on Reporting of Evaluation Studies in Health Informatics (STARE-HI) [13]. At the very least, further development of and adherence to such reporting standards would help researchers produce more generalizable results and readers interpret the results in the true context in which the study took place. Our STAMP recommendations were to simply report the measures; we did not attempt to make any determination about what a ‘right’ study design should be, including how to determine system maturity. We welcome further thoughts on the issues and additional guidance from all disciplines, including management science, on how best to design studies for evaluating HIT implementation and how best to record and report study characteristics and contexts.

Conflict of Interest Statement

The authors declare that they have no conflicts of interest in the research.

Human Subjects Research

The work described in this letter did not involve a study of human subjects.

References

1. Litwin AS, Avgar AC, Pronovost PJ. Measurement error in performance studies of health information technology: Lessons from the management literature. *Applied Clinical Informatics* 2012; 3: 210–220.
2. Ash JS, Sittig DF, Campbell EM, Guappone KP, Dykstra RH. Some unintended consequences of clinical decision support systems. *AMIA Annu Symp Proc* 2007: 26–30.
3. van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. *J Am Med Inform Assoc* 2006; 13: 138–147.
4. Embi PJ, Leonard AC. Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study. *J Am Med Inform Assoc* 2012; 19(e1): e145–e148.
5. Abookire SA, Teich JM, Sandige H, Paterno MD, Martin MT, Kuperman GJ, Bates DW. Improving allergy alerting in a computerized physician order entry system. *Proc AMIA Symp* 2000: 2–6.
6. Hanauer DA, Wentzell K, Laffel N, Laffel LM. Computerized Automated Reminder Diabetes System (CARDS): e-mail and SMS cell phone text messaging reminders to support diabetes management. *Diabetes Technol Ther* 2009; 11: 99–106.
7. Zheng K, Padman R, Johnson MP, Engberg J, Diamond HH. An adoption study of a clinical reminder system in ambulatory care using a developmental trajectory approach. *Stud Health Technol Inform* 2004; 107(Pt 2): 1115–1119.
8. Weingart SN, Toth M, Sands DZ, Aronson MD, Davis RB, Phillips RS. Physicians' decisions to override computerized drug alerts in primary care. *Arch Intern Med* 2003; 163: 2625–2631.
9. Lo HG, Newmark LP, Yoon C, Volk LA, Carlson VL, Kittler AF, Lippincott M, Wang T, Bates DW. Electronic health records in specialty care: a time-motion study. *J Am Med Inform Assoc* 2007; 14: 609–615.
10. Zheng K, Guo MH, Hanauer DA. Using the time and motion method to study clinical work processes and workflow: methodological inconsistencies and a call for standardized research. *J Am Med Inform Assoc* 2011; 18: 704–710.
11. Rotich JK, Hannan TJ, Smith FE, Bii J, Odero WW, Vu N, Mamlin BW, Mamlin JJ, Einterz RM, Tierney WM. Installing and implementing a computer-based patient record system in sub-Saharan Africa: the Mosoriot Medical Record System. *J Am Med Inform Assoc* 2003; 10: 295–303.
12. Hollingworth W, Devine EB, Hansen RN, Lawless NM, Comstock BA, Wilson-Norton JL, Tharp KL, Sullivan SD. The impact of e-prescribing on prescriber and staff time in ambulatory care clinics: a time motion study. *J Am Med Inform Assoc* 2007; 14: 722–730.
13. Talmon J, Ammenwerth E, Brender J, de Keizer N, Nykanen P, Rigby M. STARE-HI – Statement on reporting of evaluation studies in Health Informatics. *Int J Med Inform* 2009; 78: 1–9.