Commentary

The perils of meta-regression to identify clinical decision support system success factors

Christopher L. Fillmore *, Casey A. Rommel, Brandon M. Welch [1], Mingyuan Zhang, Kensaku Kawamoto

Department of Biomedical Informatics, University of Utah, 421 Wakara Way, Ste 140, Salt Lake City, UT 84108, United States

A B S T R A C T

Clinical decision support interventions are typically heterogeneous in nature, making it difficult to identify why some interventions succeed while others do not. One approach to identify factors important to the success of health information systems is the use of meta-regression techniques, in which potential explanatory factors are correlated with the outcome of interest. This approach, however, can result in misleading conclusions due to several issues. In this manuscript, we present a cautionary case study in the context of clinical decision support systems to illustrate the limitations of this type of analysis. We then discuss implications and recommendations for future work aimed at identifying success factors of medical informatics interventions. In particular, we identify the need for head-to-head trials in which the importance of system features is directly evaluated in a prospective manner.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Clinical decision support (CDS) interventions are typically heterogeneous in nature, making it difficult to identify why some interventions succeed while others do not [1]. In recent years, investigators have sought to address this difficulty through the use of meta-regression techniques [1–6]. In this approach, investigators apply regression techniques to identify features of information systems (e.g., workflow integration or patient engagement) that are significantly associated with desired outcomes. However, meta-regression analysis has recognized limitations [7–9]. A particularly important limitation is the observational nature of this type of analysis. As such, causal relationships cannot be illuminated by this approach alone. To illustrate why care should be taken when performing or interpreting this type of analysis, this manuscript provides a cautionary case study in the context of CDS evaluation.

In 2013, the *British Medical Journal* published a meta-regression analysis by Roshanov et al. [6] of randomized controlled trials (RCTs) of CDS systems. The study sought to identify features associated with effective systems and resulted in conclusions that differed significantly from similar studies, including a prior meta-regression analysis in the *British Medical Journal* by

Kawamoto et al. [1] Notably, Roshanov et al. [6] found that advice given automatically in workflow was not significantly associated with system success in their initial model. As a result, this feature was removed from their final model.

This particular finding was unexpected, as it differed significantly from findings suggested by previous reviews addressing clinical decision support [1,2,4,10–12]. Moreover, RCTs that directly evaluated the importance of this feature have found automatic provision to be important [13,14]. Specifically, within the clinical context of hyperlipidemia management, van Wyk et al. [14] compared alerts provided automatically to physicians within an EHR versus on-demand CDS which had to be proactively accessed by physicians within the same EHR. In this cluster RCT involving 38 Dutch general practices and 87,886 patients, 65% of the patients requiring screening were screened in the automatic CDS group, as compared to 35% in the on-demand CDS group (adjust relative risk 1.40; 95% confidence interval [CI], 1.08 to 1.81) [14]. In another RCT directly evaluating the importance of providing CDS automatically, Scheepers-Hoeks et al. [13] compared alerts provided automatically to physicians within an EHR versus the same information provided on-demand in the EHR. In this RCT, which was conducted in an intensive care unit (ICU) regarding 13 locally developed clinical rules, compliance with the CDS recommendations was 41% in the automatic alerting group, versus 19% in the on-demand EHR group ($p < 0.0001$) [13]. Such findings from head-to-head RCTs must be considered very seriously, as they directly evaluate causative relationships between a CDS feature and system impact, rather than merely correlation as in the case of a meta-regression analysis.

---

 * Corresponding author. Present address: Homer Warner Center for Informatics Research, Intermountain Healthcare, 5171 S Cottonwood St, Ste 220, Salt Lake City, UT 84107, United States.
    E-mail address: chris.fillmore@imail.org (C.L. Fillmore).
 [1] Biomedical Informatics Center, Medical University of South Carolina, 55 Bee St, Charleston, SC 29425, United States.

Given the stark discrepancy in findings between the meta-regression analysis by Roshanov et al. [6] and these prior studies on the topic, we sought to discover an explanation for these findings. Here, we describe our findings and discuss implications for future work aimed at identifying success factors of medical informatics interventions.

2. Methods

2.1. Initial investigation

Based on results from other studies identifying the importance of automatic CDS provision, as well as our initial review of the source data set provided in Appendix of the Roshanov et al. [6] article, we suspected that the differences between the two meta-regression analyses were more likely due to discrepancies in the source data rather than differences in statistical analysis methods. In particular, in the meta-regression analysis by Kawamoto et al. [1], CDS systems that automatically provided their advice had a success rate of 75%, versus a success rate of 0% for systems that did not (difference = 75%). In contrast, in the systematic review by Roshanov et al. [6], the difference in success rate was only 60% versus 54% (difference = 6%).

In examining potential reasons for the differences in the source data set, we found that the differences appeared to stem from discrepancies in the determination of whether specific explanatory features were present or absent in a given CDS system. In particular, Roshanov et al. [6] appeared to consider many CDS systems whose use was required by a study protocol to not be automatically provided as a part of clinician workflow, whereas the end result would be the same: clinicians would always be exposed to the intervention.

2.2. Hypothesis

In our opinion, mandated, protocol-driven use of a CDS system is functionally equivalent to automatic provision of CDS. Therefore, we hypothesized that the results would be more consistent with prior studies if we considered protocol-driven provision of CDS to constitute automatic provision. To test this hypothesis, we repeated the same statistical analysis conducted by Roshanov et al. using this updated definition of automatic CDS provision.

2.3. Dataset

We used the dataset of 162 randomized controlled trials of CDS systems identified by Roshanov et al. [6]

2.4. Outcome and explanatory variables

As in the original systematic review, "effective" systems were defined as those systems that improved primary (or 50% of secondary) reported outcomes of process of care or patient health. For the outcome measure, we maintained the determinations made by Roshanov et al. [6] As in the original study, we focused on six potential explanatory variables for CDS outcomes: automatic provision of CDS, development by authors, feedback at the time of care, advice presented in electronic charting or order entry, advice for patient, and requires reason for override.

2.5. Feature reanalysis

We re-assessed each source study for the presence of "automatic provision of CDS" using an updated definition that included cases where use of the CDS was mandated by protocol. If both the

intervention and control groups involved use of a clinical information management tool (e.g., EHR or computerized provider order entry [CPOE] system), the CDS was considered to be automatically provided if exposure to the CDS did not require any end-user initiative beyond simply using the base information management tool. If it was unclear whether the CDS was provided automatically, we deferred to the determination made by Roshanov et al. [6].

In addition, we sought to improve the quality of the data set for the reanalysis by completing incomplete explanatory variable classifications. Specifically, we reanalyzed the "author developed" and "feedback at time of care" features for trials where Roshanov et al. [6] were unable to determine the presence or absence of those features. We considered a CDS intervention to have been developed by the authors if the study stated or implied that one or more of the study authors were involved in the development of the system. If the authors developed the underlying clinical algorithm or knowledge base used in a CDS system, we also considered the CDS intervention to have been developed by the authors. We considered a CDS intervention to have delivered feedback at the time of care if the CDS was provided to the clinician while the clinician was with the patient in question for the purposes of clinical care. If we were unable to make a determination based on the manuscript, we attempted to make direct contact with the authors to make these determinations. If it was unclear whether the CDS system was author developed or provided feedback at the time of care, and if we were unable to make contact with the authors, we maintained the unknown status.

The presence or absence of features was determined through the consensus of four of the authors (CLF, MZ, BMW, and KK). These reviewers consisted of two physician informaticists (KK and CLF), one nurse informaticist (MZ), and one additional health informaticist (BMW).

2.6. Exclusions

We excluded studies where the intervention arm with the CDS system was disadvantaged in a significant way compared to the control arm. Example disadvantages included care delivery by practitioners with significantly less training (e.g., nurse versus physician), or the removal of a key diagnostic resource.

2.7. Statistical analysis

The original data set and statistical analysis file used in the manuscript by Roshanov et al. [6] was kindly provided by the authors. Using this data set and analysis file, the primary analysis from the original study was replicated. Following this replication of findings, the data set was modified as described above, and the primary analysis was repeated with the updated data. In brief, the primary analysis as developed by Roshanov et al. [6], and replicated here without modification, consisted of the following: (i) the selection of explanatory variables for inclusion in the final primary model, with variables selected if they were associated with the study outcomes with a $p < 0.1$; and (ii) the development of a final model to explain the study outcomes using these selected explanatory variables. For these analyses, Firth's profile penalized likelihood method was used as in the original study. All statistical analyses were performed using Stata 13.1 [15].

3. Results

3.1. Feature reanalysis

In the reanalysis, we revised the "automatic provisioning" feature for 46 (28%) of the original 162 studies in the data set. Of these