

# Use of Clinical Alerting to Improve the Collection of Clinical Research Data

Krystl Haerian, MD,<sup>1</sup> Jon McKeeby, DSc,<sup>2</sup> Gary DiPatrizio, PharmD,<sup>2</sup> and James J Cimino, MD<sup>1</sup>

<sup>1</sup> Laboratory for Informatics Development

<sup>2</sup> Department of Clinical Research Informatics

National Institutes of Health Clinical Center, Bethesda, Maryland, USA

## Abstract

*Data errors in electronic health records have been shown to have the potential to adversely impact the conclusions drawn from clinical research. We prospectively studied the efficacy of a new alert to infer errors in previously stored data and to decrease the frequency of data entry errors, in an attempt to improve the quality of data for clinical trials. For the purpose of this study, we monitored data entry errors in height or weight measurements. We pre-determined the criteria for probable error as a ten percent variance from a patient's reference value. The care provider entering a value satisfying our error criteria received a disruptive pop-up alert message. The study revealed a significant decrease in the frequency of data errors stored in the EHR, from 2.4% before the alert to 0.9% after the alert. These findings have implications for the development of clinical research trial data collection support tools.*

## Introduction

Goethe wrote, "Nothing is more harmful to a new truth than an old error."<sup>1</sup> Clinical research studies search for new truths, but inherent data errors stored in the clinical research databases have the potential to impact those findings negatively. Studies have shown that even seemingly trivial data errors can substantially affect the accuracy of research findings, especially for rarely occurring adverse events.<sup>2</sup>

Numerous studies have examined the rates of data errors in research databases. Using a "double entry" method, Goldberg and colleagues detected data error rates that ranged from 2.3 to 26.9%.<sup>3</sup> A literature review of 42 articles that used the source to database method, which compares case report forms to database entries, found the average error rate was 9.76%.<sup>4</sup> Double data entry may allow the detection of some errors. However, this method is typically applied during transcription of original data into research databases; it does not prevent or identify errors in the data as they are initially recorded.

Data errors not only impact the statistical conclusion of research studies, they also generate a financial burden. A large multi-institutional clinical trial can cost as much as \$100 million dollars. An average 12

month clinical trial with 2,000 patients may generate three million data points, and at least 15 percent of expenses are related to monitoring to assure the validity and accuracy of the data.<sup>5</sup>

When data errors are detected after collection, they may affect the ability to analyze study results or may force the disqualification of research subjects, which is likely to increase the cost of a study and may even render it invalid. When data errors are not detected, they may lead to incorrect conclusions

We conducted this study to determine if the alerting system in a commercial electronic health record (EHR) could be applied effectively to improve data collection in clinical trials and to add to the limited literature available on the types of data errors that exist in clinical trial databases.

## Background

The Clinical Center at the National Institutes of Health (NIH) in Bethesda, Maryland, is the nation's largest hospital devoted exclusively to clinical research. It houses 234 inpatient beds and 82 outpatient day hospital stations. There are currently approximately 1,600 clinical research trials being conducted at the Clinical Center.

As the support center for clinical trials, the Clinical Center has a unique patient population. Every patient seen at the Clinical Center is actually a subject participating in a research study. All patients are treated according to standardized research protocols. These protocols address many aspects of the patient's care, including the dosage and timing for administering medications, when to perform diagnostic procedures and which laboratory tests and other values need to be monitored.

Since 2004, clinicians at the NIH Clinical Center have used the Clinical Research Information System (CRIS), which is an implementation of Sunrise Clinical Manager (Eclipsys Corporation, Boca Raton, FL). Research protocols are carried out through the use of CRIS's order entry functions to schedule all data collection and interventions. One type of data collection that is common to virtually all protocols is the routine recording, by nurses, of the patients' vital

signs, including heart rate, blood pressure, respiratory rate, temperature, height and weight.

The CRIS system includes, as part of the standard commercial product, an alerting-and-reminder function that is typically used to check physician orders for patient allergies, drug interactions, and proper dosing of medications. The logic of the alerting system uses previously acquired patient data, such as allergies and current medications, to determine whether or not to send an alert.

Dosing alerts often depend on the availability of accurate patient vital sign data, such as heights and weights. The process by which vital signs are recorded in CRIS is subject to several types of human error, such as incorrect recording, incorrect transcription due to illegibility, and typographical errors. Attempting to improve the recording of vital signs through, for example, dual recording or dual entry is impractical and expensive. However, the presence of errors can be inferred when variations in sequentially recorded data appear physiologically implausible. We therefore sought to use such variations as a means to detect errors in previously recorded data and to support the logic of an alert that would attempt to catch and prevent the recording of new erroneous data.

## Methods

We conducted a prospective study of a new alert, with a before-and-after comparison of the frequency of data that appeared erroneous. All patients seen at the Clinical Center from October 2007 to January 2009 who had two or more documented height or weight measurements were included in this study.

The criterion for a probable error in recorded values was pre-determined to be the finding of a greater than 10% variance from a patient's prior last recorded value. Using standard alert authoring functions, we constructed an alert in the CRIS system that compares a height or weight, upon entry by a care provider, with the most recent comparable value. A value that varies by 10% or more will trigger a disruptive pop-up alert message. The care provider is given the opportunity to delete the value,

correct the value, or proceed to save the value despite the alert.

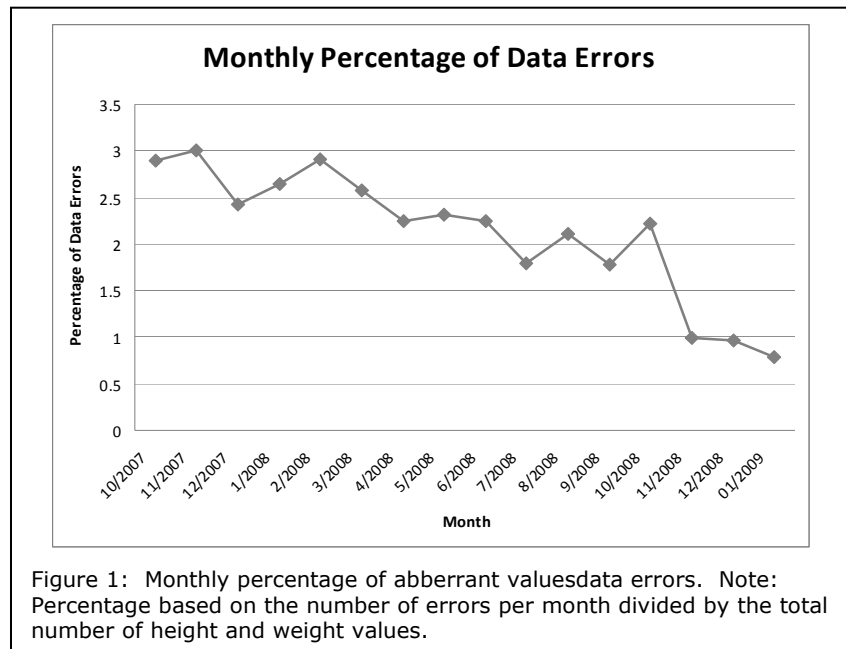
We used the same 10% change principle for identification of aberrant values in height and weight data recorded in the database. The number of aberrant values per month were tallied for comparison before and after introduction of the alert.

## Results

A functional alert was introduced into CRIS on 10/22/08. It fired an average of 158 times per month. There were 47 alerts in October, 159 in November, 157 in December, and 171 in January.

A total of 110,426 data entries were included in our analysis. The average time that elapsed between measurements when a >10% change occurred was 21 days (SD 36 days), with a minimum of a few minutes to a maximum of 294 days. Weights are measured more frequently than heights, with an approximate ratio of 3.2 weight data points per one height measurement. Our findings are the result of a higher representation of weight measurements.

Of the 90,684 values recorded in the twelve months before introduction of the alert, 2156 were classified as aberrant, while 176 values were classified as aberrant among the 19,742 values recorded in the three months subsequent to the implementation of the alert. The monthly frequencies of aberrant values over the year prior to the introduction of the alert through the end of the study period are shown in Figure 1. Thus, the introduction of the alert was associated with a change in the frequency of aberrant



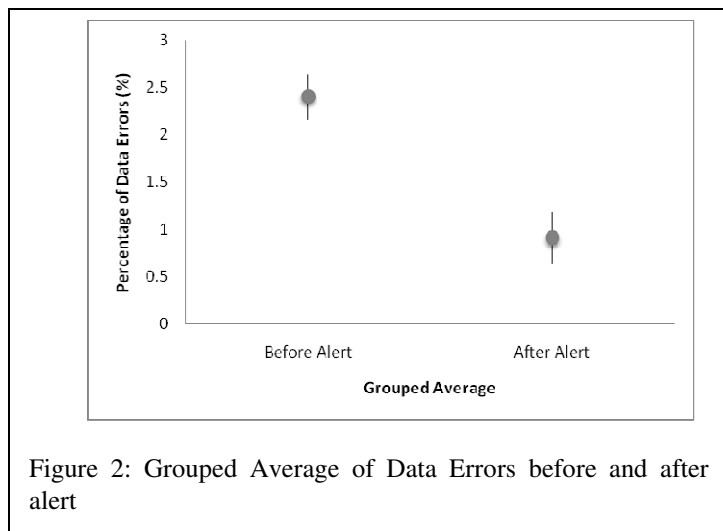


Figure 2: Grouped Average of Data Errors before and after alert values from 2.4% (95% CI: 2.16-2.64) to 0.9% (95% CI: 0.64-1.18). (Figure 2).

### Discussion

Alerts have been an effective clinical decision support tool for clinical care.<sup>6</sup> Some of the ways in which alerts have been demonstrated to be effective in clinical decision support include: improving immunization administration, improving safe and effective medication use, decreasing ordering of drugs contraindicated due to renal insufficiency, and preventing patient allergic reactions.<sup>7,8,9</sup> The intention of such clinical alerts is to improve patient care.

This study was conducted at the NIH Clinical Center, a facility that specializes in the care of patients enrolled in clinical research protocols. In a clinical research environment, medical providers have dual responsibilities, they must provide clinically appropriate care and they must collect research data in accordance with a clinical trial protocol.

Health care providers undergo intensive training in patient care, which facilitates an innate understanding of the importance of the medical logic behind well formulated clinical alerts, yet even alerts that provide meaningful clinical decision support can be ignored or overridden 49-96% of the time.<sup>10</sup> Our alert was overridden about 38% of the time.

We were not certain, at the start of this study, that the significance of our alert would be clear to the clinicians. Since alerts whose worth is not understood tend to be the ones that are ignored, we wanted to investigate if the height and weight alerts would be a valuable tool in clinical trial data collection. Entering factual data into a database is a different, perhaps more passive thought process, but

through this study we showed that alerts appear to be an effective support tool for data collection. Whereas this study was limited to an alert and analysis of height and weight data, these results certainly raise the possibility that other types of alerts may facilitate accurate data collection in clinical research trials.

Our choice of a 10% change as the trigger for our alert was multi-factorial. Some of the reasons behind our choice included: 1) A desire to achieve accurate data entry. An ideal alert would predict the numerical data value that should be collected and only fire when this value is entered incorrectly. In the case of height and weight measurements, no definitive value is available, so an acceptable range of plausible values was chosen. 2) The need to ensure the accurate recording of the values used to calculate a patient's Body Surface Area, as these values are particularly important for subsequent critical decisions, such as medication dosing. 3) The ethical responsibility to maximize patient safety balanced with the reality of alert fatigue and the need to minimize false alerts.

The choice of a 10% variation in values as our threshold appeared sufficiently small that it would identify the majority of data entry errors (and virtually all of the clinically significant errors), while being large enough to limit the number of false positives and possible resultant "alert fatigue".

The purpose of the alert was to prevent data entry errors, but it is possible that over time a person may gain or lose 10% of their body weight, triggering a false alert. A false alert occurs when a popup alert occurs informing the practitioner that a possible data entry error may have occurred when in fact the data entered was accurate. As the popup message displayed includes both the current weight value and the value for the last recorded weight, this false alert may provide a clinical benefit. Weight changes in adults often have clinical implications. A significant weight loss could signal a loss of appetite seen in depression or as a medication side effect. Obesity is a risk factor for a number of health conditions, such as diabetes and heart disease, and a patient who is gaining weight may require dietary counseling as a preventive health measure. Therefore, while the alert is classified as a false positive as far as our clinical research data collection purposes are concerned, a benefit may exist from a clinical care perspective in raising awareness that a weight change has occurred and may need to be addressed.

Besides the implications for clinical trial support, this study also has implications for patient safety. Some potentially toxic medications, such as those used in chemotherapy, are dose calculated based on a patient's height and weight, therefore getting the correct data into the EHR is highly valuable. Latent errors have been defined as often unrecognized errors that have the ability to result in multiple types of active errors.<sup>5</sup> We have identified a latent source of error at Clinical Center: clinical decision support automatically calculates medication dosage based on the height and weight stored in the database. An error can be passed from the storage in the database to an electronic prescription, as automation has given clinicians less tangible contact with the ordering process. The loss of this cognitive checkpoint makes it imperative that the correct data are available in the system. Our alert proved a simple and effective way of addressing this problem. The alert resulted in a significant decrease in the percentage of data errors from 2.4% (95% CI: 2.16-2.64) to 0.9% (95% CI: 0.64-1.18).

Our study had a number of limitations. By choosing a 10% change as our criteria for an alert, a number of false positive alerts were generated. Different parameters could have been chosen to allow for greater fluctuation in a patient's weight over time to decrease false alerts, but this would have resulted in fewer patients benefiting from the safety benefits of the alert. Patients are at increased risk of receiving the incorrect dose of a medication based on Body Surface Area (BSA) if an incorrect weight is present in their EHR. Antineoplastic drug doses are frequently normalized based on BSA to achieve a constant therapeutic effect and for avoidance of toxic side effects. We felt it was advantageous to improve the data entry used in dosage calculations for clinical trial participants, regardless of the time between height and weight measurements. We decided that the ethical importance of improved patient safety outweighed the possible benefit of decreasing erroneous alerts.

Our intention was to identify and correct as many errors as possible. To avoid alert fatigue, different criteria should be used for younger pediatric patients, who may be rapidly changing in height and weight. From birth to five years, rapid changes in height and weight occur in pediatric patients. Stepwise alerts based on growth charts would be an appropriate standard to consider in younger patients.

While all this speaks to a limitation in the alert design and frequency of false positive alerts, our overarching goal was not to design the perfect alert for height and

weight data collection- but to test the hypothesis that an alert would be effective in improving the accuracy of data collection.

The alert is not 100% sensitive and specific to all data entry errors, but merely a tool to identify all substantive errors with a minimum of false alarms. We presume that when clinicians are presented with our alerts, that they will know whether their values are accurate or suspect and, when in doubt, return to the patient to obtain a repeat measurement. However, the study of what clinicians *actually* do is an important subject of further study, since simply overriding the alert or deciding to enter no value at all are potential alternative actions.

Clinical research studies provide an entirely new venue for alert analysis. These trials might require more frequent monitoring, or procedures such as blood draws, when compared to an average inpatient admitted to the hospital solely for clinical care. In addition, patients in such trials might be taking a novel therapeutic medication that is not present in the drugs and pharmaceutical database. In this latter situation some drug-allergy information cannot be ascertained. In future studies we plan to explore how alerts that relate to other aspects of patient care can further support clinical research studies.

Now that we can conclude that an alert is effective at improving data collection, the next step would be to study alerts that facilitate data collection at multiple points in a research protocol. This could include error range checking and reminders for time point data collection. Ultimately, an automated system that would develop alerts based on the research protocol submitted to an Internal Review Board may be designed.

## Conclusions

Our ultimate goal is to transform the processes through which data entry and collection occur in clinical research studies. This study supports the concept that prospective alerts may significantly increase the quality of data entry in clinical research. Alerts have previously been shown to decrease certain types of clinical errors; this study demonstrated their usefulness in preventing data collection errors. Identification of research data errors is of particular importance in clinical studies, where certain patient parameters are measured and used for evaluation of study metrics. Additionally, discovering a mistake is an opportunity to improve the quality of research by offering a real time opportunity to collect the right value.

## Acknowledgments

This work is supported intramural research funds from the NIH Clinical Center and the National Library of Medicine.

We would like to thank Patricia Sengstack and Lincoln Farnum, members of DCRI, for assistance implementing the alert.

## References

1. Wood, James. Dictionary of Quotations from Ancient and Modern English and Foreign Sources. London: Frederick Warne and Co., 1899. ("Einer neuen Wahrheit nichts ist schadlicher als ein alter Irrtum" -Johann Wolfgang von Goethe)
2. Gallivan S, Stark J, Pagel C, Williams G, Williams WG. Dead reckoning: can we trust estimates of mortality rates in clinical databases? *Eur J of Cardiothoracic Surg* 2008;33: 334-340.
3. Goldberg S, Niemierko A, Turchin T. Analysis of Data Errors in Clinical Research Databases. *AMIA Annu Symp Proc*. 2008; 242-246.
4. Nahm ML, Pieper CF, and Cunningham MM. Quantifying data quality for clinical trials using electronic data capture. *PLoS ONE* 2008; 3(8): e3049.
5. Davis JR, Nolan VP, Woodcock J, Estabrook RW. *Assuring Data Quality and Validity in Clinical Trials for Regulatory Decision Making*: Institute of Medicine. Washington, DC: National Academy Press, 1999.
6. Bates DW, Kuperman GJ, Wang S, Gandhi T, Kittler A, Volk L, Spurr C, Khorasani R, Tanasijevuc M, Middleton B. Ten commandments for effective clinical decision support: making the practice of evidence based medicine a reality. *J Am Med Inform Assoc*. 2003; 10(6):523-530.
7. Fiks AG, Grundmeier RW, Biggs LM, Localio AR, Alessandrini EA. Impact of clinical alerts within an electronic health record on routine childhood immunization in an urban pediatric population. *Pediatrics*. 2007 Oct;120(4):707-14.
8. Payne TH. Computer decision support systems. *Chest* 2000; 118:47S-52S.
9. Galanter WL, Didomenico RJ, Polikaitis. A Trial of Automated Decision Support Alerts for Contraindicated Medications Using Computerized Physician Order Entry. *J Am Med Inform Assoc*. 2005;12:269-274.
10. 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.