Identifying and Quantifying Medication Errors: Evaluation of Rapidly Discontinued Medication Orders Submitted to a Computerized Physician Order Entry System

ROSS KOPPEL, PHD, CHARLES E. LEONARD, PHARM D, A. RUSSELL LOCALIO, JD, PHD, ABIGAIL COHEN, PHD, RUTHANN AUTEN, BA, BRIAN L. STROM, MD, MPH

Abstract
All methods of identifying medication prescribing errors are fraught with inaccuracies and systematic bias. A systematic, efficient, and inexpensive way of measuring and quantifying prescribing errors would be a useful step for reducing them.

We ask if rapid discontinuations of prescription-orders—where physicians stop their orders within 2 hours—would be an expedient proxy for prescribing errors?

To study this we analyzed CPOE-system medication orders entered and then discontinued within 2 hours. We investigated these phenomena in real time via interviews with corresponding ordering physicians. Each order was also independently reviewed by a clinical pharmacist or physicians. We found that of 114 rapidly discontinued orders by 75 physicians, two-thirds (35 of 53, PPV = 66; 95% CI = 53–77) of medication orders discontinued within 45 minutes were deemed inappropriate (overdose, underdose, etc.). Overall, 55% (63 of 114; 95% CI = 46–64%) of medication orders discontinued within 2 hours were deemed inappropriate.

This measure offers a rapid, constant, inexpensive, and objective method to identify medication orders with a high probability of error. It may also serve as a screening and teaching mechanism for physicians-in-training.

Introduction
Prescribing errors are one of the most frequent types of medical errors1-5 and largest proportion of medication errors, causing ill effects in 1% of inpatients.2-8 Prescribing errors, however, are among the most preventable and are therefore a focus of patient safety interventions.9-11 Identifying prescribing errors, unfortunately, is itself fraught with inaccuracy.12,13 Each method of detection and reporting is subject to systematic bias:

1. Medical record analysis catches errors that may not reach the patient and misses errors linked to undocumented diagnoses and due to improper or delayed diagnoses.13,14 It is also time-consuming and expensive.
2. Many manifestations of medication errors go unrecognized because symptoms are often complex, patients have multiple systems problems, and polypharmacy may obscure causes and symptoms.
3. Self-reports of prescribing errors are rare because they are seldom known (if physicians realize they are making prescribing errors, they correct them). Self-reports also are limited by concerns of litigation and status loss.13,15
4. Errors intercepted by colleagues are often corrected informally and are infrequently reported.13,16,17
5. Observational methods require knowledgeable personnel, are more time- and money-consuming than chart

Affiliations of the authors: Department of Sociology, University of Pennsylvania, (RK), Philadelphia, PA; Department of Biostatistics and Epidemiology, Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania School of Medicine, (RK, CEL, ARL, AC, RA, BLS), Philadelphia, PA; Center for Education and Research on Therapeutics, University of Pennsylvania School of Medicine (CEL, ARL, AC, RA, BLS), Philadelphia, PA; Department of Medicine (General Medicine Division), Department of Pharmacology, University of Pennsylvania School of Medicine (BLS), Philadelphia, PA.

The authors thank the following researchers (Melissa Iwanowski) and clinicians (Manisha C. Shanbhag, MD and Ann G. Till, MD) for their contributions. The authors thank David Wordell, RPh for data access and helpful insights; Amy Praestgaard, MS, for statistical analysis; and Edmund Weisberg, MS, for help with editing.

No member of the research team has a financial relationship with the Eclipsys Corporation or with the Pharmacy computer system, Megasource. There are no known conflicts of interests with any vendor or other entity.

This research was approved by the Institutional Review Board of the University of Pennsylvania’s Office of Regulatory Affairs, as is so noted in the text.

The data were under the control and supervision of Dr. Koppel and the co-authors. Dr. Leonard supervised the data collection. Drs. Koppel, Cohen, Localio, and Strom designed the study and analyzed the findings. Dr. Koppel and Dr. Leonard wrote the early drafts; others contributed to subsequent editions.

This research was funded by The Agency for Healthcare Research and Quality P01 HS11530.

Correspondence: Ross Koppel, Ph.D. 813 Pardee Lane, Wynnewood, PA 19095; e-mail: <rkoppel@sas.upenn.edu>.

Received for review: 07/06/07; accepted for publication: 03/19/08

analysis, and are far more likely to capture medication administration and dispensing errors rather than misdiagnoses, missing information, and poor prescribing.\textsuperscript{18,19}
6. Even pharmacy-intercepted errors are generally handled with deference and only occasionally logged.\textsuperscript{6,10,11,15}
7. Sentinel or trigger signals (e.g., certain abnormal lab results, leukopenia, and use of certain drugs) used to detect or prevent adverse drug events (ADEs) show great promise for both patient safety and cost effectiveness. Unfortunately, they quickly generate alert fatigue. Also, few have been implemented and those that have are implemented in non-standardized ways and therefore difficult to compare.\textsuperscript{20–23}
8. Use of combined methods would produce a more comprehensive analysis, thus mitigating the deficiencies of each approach. Such an undertaking, however, would be resource intensive and would neither eliminate undercounts nor avoid systematic biases.

Thus, development of a systematic, valid, and efficient way of measuring and quantifying prescribing errors would be a useful step toward reducing medication errors. The idea is not entirely new: 15 years ago, Classen et al.\textsuperscript{22} included “abrupt stops” as a possible trigger to adverse drug events. However, with the technology available then, such a measure was cumbersome, expensive, and, required post hoc chart analysis. Others have examined this type of metric,\textsuperscript{20,21,23} but it has not received systematic analysis with modern computer systems. With the anticipated widespread implementation of computerized physician order entry (CPOE) systems, a computer-assisted and validated method of quickly identifying medication errors would be of even greater utility.

We hypothesized that a rapid discontinuation of a prescription order, where the physician stops the order within 1 minute to 2 hours, might be a valuable and expedient proxy for prescribing errors. Whether the error was detected by a physician, pharmacist, nurse, or patient, we suspected that a change that soon may suggest the original order was probably suboptimal. We were able to examine this hypothesis via real-time access to orders placed in a CPOE system and almost-immediate interviews of physicians entering those orders.

**Methods**

**Setting and Design**

The setting was a tertiary-care, urban, academic hospital with 750 beds and a CPOE system (SCM 3.5, Eclipsys Corporation, Boca Raton, FL). At the time of the study, the clinical decision support system was operational for drug-drug interactions and for known drug allergies.

All medication orders are entered into the hospital’s CPOE system and then transmitted to the pharmacy for approval and dispensing. The hospital pharmacy department provided the investigators with access to real-time orders, allowing immediate identification of medication orders and discontinuations. Orders were evaluated during a non-random sample of 24 days between March 14, 2005 and May 10, 2005, including nights and weekends. Times-of-day during the 24 days were intentionally scheduled to ensure a full and representative sample of orders, i.e., researchers’ observation times paralleled the order flow.

We interviewed housestaff who discontinued a medication order within 2 hours of writing that order, focusing on the rationale for quickly discontinuing that order. We also segmented and analyzed the discontinued orders within 15-minute time blocks.

**Identification of Subjects and Questionnaire Administration**

Subjects were physicians ordering medications during the study period. We telephoned those discontinuing a medication order within 2 hours of the initial order and for whom we had sufficient identifying information. Those giving verbal consent were interviewed. Physicians completing the interview were offered a $5 discount coupon to an in-hospital coffee shop. Interviews were conducted within 30 minutes to 6 hours of the discontinuation time. All interviews were conducted by a clinical research pharmacist or specially-trained nursing student under his direct supervision. The interview schedule had 4 elements:

1. A reminder of the details of the order.
2. Inquiry about their reason(s) for discontinuing the initial medication and for any information used in this decision (e.g., advice from pharmacist, nurse, colleague, or new data).
3. Inquiry about any subsequent medication order made in light of the recent discontinuation.
4. Further elaboration.

**Categorization and Validation of Outcome**

Interview data were retrospectively reviewed by a clinical pharmacist to determine the proportion of rapidly discontinued medication orders where the initial order was deemed inappropriate, i.e., a medication error or suboptimal order (e.g., overdose, underdose, drug allergy ignored, wrong eye specified, suboptimal antimicrobial choice, inappropriate intravenous fluid). To facilitate this judgment, the clinical pharmacist also used any information in the electronic medical record—data also available to the ordering physician at the time of the order. When the clinical pharmacist was unable to determine the appropriateness of the initial medication order, one of two resident physicians reviewed the interview results and medical records to make the judgment. These resident physicians were not members of the study cohort.

**Statistical Analysis**

For each 15-minute time segment, we calculated the proportion of medication orders discontinued within 2 hours that were identified by the expert reviewers as inappropriate (along with positive predietive values (PPVs) and confidence intervals). Equally important, we sought to find the optimal time segments for identifying rapidly discontinued medication orders, i.e., the highest ratios of possible errors within time span groupings—the first 15 minutes, 30 minutes, 45 minutes, 1 hour, 2 hours. Of note, to avoid including immediately self-corrected mistakes, typographic, and mouse errors, we excluded all orders discontinued within one minute of the initial order.

This study was approved by the University of Pennsylvania’s Institutional Review Board.
Results
In the specified 24 days, we identified 398 medication orders that were discontinued within 2 hours (see Figure 1). We attempted to contact all physicians with known phone numbers who had not previously rejected our calls. Reasons for non-contact included: physician ID in computer not linked to known physician, “new” refusals, physician known but contact information faulty, and medication order not relevant (e.g., order for Accu-Check glucose testing strip). This process resulted in 228 orders for possible analysis. We were able to interview the physicians (N = 75) who wrote 114 separate rapidly-discontinued medication orders (see Figure 1 for reasons for refusal). The study’s pharmacist requested that 65 of the 114 orders (57%) also be reviewed by the study’s physicians to determine the orders’ appropriateness.

Two-thirds (35 of 53) of the medication orders discontinued within 45 minutes were deemed inappropriate (PPV = 66%; 95% CI = 53–77%). By comparison, 55% (63 of 114; 95% CI = 46–64%) of all medication orders discontinued within 2 hours were deemed inappropriate.

Table 1 illustrates the time segments examined to determine the most efficacious time parameter to maximize ratios of inappropriate-to-appropriate medication orders. In general, the proportion of inappropriate orders appears highest in the first 45 minutes, and then diminishes somewhat for the next 45 minutes, but increases again in the 90-to-120-minute period. Orders stopped within 30 minutes were most likely to be deemed inappropriate. However, the percentage, 67%, is essentially the same as that found inappropriate at 31–45 minutes.

Trend Analyses
When we considered time intervals up to 120 minutes, we detected a marginally significant trend (p = 0.06) with the Cochran-Armitage trend test. That same test also shows that the percentage of inappropriate orders decreases across time categories up to 90 minutes (p = 0.02).

Reasons and Stimulus for Discontinuations
In our interviews, physicians noted the reasons orders were discontinued were drug-disease reconsiderations, drug-drug interactions, and patient preferences. The impetus for the change was from: other housestaff, nurses, and pharmacists (who often call the physicians with questions), and attendings. Physicians frequently reported they caught their own mistakes. Most of the changes, not surprisingly, were drug and dosage changes.

The classes of drugs most likely to be quickly discontinued were: low therapeutic index drugs, insulin, antiretrovirals, antineoplastics, and immunosuppressive drugs.

Discussion
These data reveal that rapidly discontinued medication orders are indicators of risks for inappropriate prescribing, or near misses. Two-thirds of orders discontinued within 45 minutes were judged to be inappropriate. Over half (55%) of medication orders discontinued within 2 hours were likewise deemed inappropriate.

Most research on medication prescribing errors suggests error rates of 1–3%,6–24 a large figure in light of the many millions of hospital medication orders written annually. The proxy measure proposed here, while preliminary and im-
Table 1 • Proportion of Rapidly Discontinued Medication Orders Found to be Inappropriate: Analyzed by 15-Minute, 30-Minute, and 45-Minute Time Segments

<table>
<thead>
<tr>
<th>15-minute slices</th>
<th>Time to discontinuation</th>
<th>1–15 m</th>
<th>16–30 m</th>
<th>31–45 m</th>
<th>46–60 m</th>
<th>61–75 m</th>
<th>76–90 m</th>
<th>91-105 m</th>
<th>106–20 m</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inappropriate</td>
<td>10 (62.5%)</td>
<td>16 (69.6%)</td>
<td>9 (64.3%)</td>
<td>8 (44.4%)</td>
<td>7 (50.0%)</td>
<td>2 (15.4%)</td>
<td>4 (66.7%)</td>
<td>7 (70.0%)</td>
<td>63 (55.3%)</td>
<td></td>
</tr>
<tr>
<td>“Not” Inappropriate or undetermined</td>
<td>6 (37.5%)</td>
<td>7 (30.4%)</td>
<td>5 (35.7%)</td>
<td>10 (55.6%)</td>
<td>7 (50.0%)</td>
<td>11 (64.6%)</td>
<td>2 (33.3%)</td>
<td>3 (30.0%)</td>
<td>51 (44.7%)</td>
<td></td>
</tr>
<tr>
<td>PPV (95% CI)</td>
<td>63 (38–81)</td>
<td>70 (49–84)</td>
<td>62 (38–84)</td>
<td>44 (24–66)</td>
<td>50 (27–73)</td>
<td>15 (5–43)</td>
<td>67 (29–90)</td>
<td>70 (39–89)</td>
<td>55 (46–64)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>30-minute slices</th>
<th>Time to discontinuation</th>
<th>1–30 m</th>
<th>31–60 m</th>
<th>61–90 m</th>
<th>91–120 m</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inappropriate orders</td>
<td>26 (67%)</td>
<td>17 (53%)</td>
<td>8 (35%)</td>
<td>12 (60%)</td>
<td>63 (55%)</td>
<td></td>
</tr>
<tr>
<td>“Not” Inappropriate or undetermined orders</td>
<td>13 (33%)</td>
<td>15 (47%)</td>
<td>15 (65%)</td>
<td>8 (40%)</td>
<td>51 (45%)</td>
<td></td>
</tr>
<tr>
<td>PPV (95% CI)</td>
<td>67 (51–79)</td>
<td>53 (36–69)</td>
<td>35 (19–55)</td>
<td>60 (38–78)</td>
<td>55 (46–64)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>45-minute slices</th>
<th>Time to discontinuation</th>
<th>1–45 m</th>
<th>46–90 m</th>
<th>91–120 m</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inappropriate orders</td>
<td>35 (66%)</td>
<td>17 (38%)</td>
<td>11 (69%)</td>
<td>63 (55%)</td>
<td></td>
</tr>
<tr>
<td>“Not” Inappropriate or undetermined orders</td>
<td>18 (34%)</td>
<td>28 (62%)</td>
<td>5 (31%)</td>
<td>51 (45%)</td>
<td></td>
</tr>
<tr>
<td>PPV (95% CI)</td>
<td>66 (53–77)</td>
<td>38 (25–52)</td>
<td>69 (44–86)</td>
<td>55 (46–64)</td>
<td></td>
</tr>
</tbody>
</table>

Trend Analyses: When we considered time intervals up to 120 minutes, we detected a marginally significant trend (p = 0.06) with the Cochran-Armitage trend test. That same test also shows that the percentage of inappropriate orders decreases across time categories up to 90 minutes (p = 0.02).”

perfect, indicates that up to 67% of prescriptions discontinued within 45 minutes after their origination are inappropriate. Even beyond the ratio comparisons, the value of this proxy measure is several-fold: When linked to a CPOE system, it is rapid, constant (24/7), and does not depend on possibly biased evaluators, self-report, or others’ reports. Data collection is also cost-free as part of a CPOE system.

These data reveal that any order discontinued within a short time has a substantial probability of having been inappropriate. Moreover, although the remaining proportion of orders could not be substantiated as inappropriate, it could be argued that almost all of these orders were perceived as problematic by their authors—that is why they discontinued them. Even in the best of circumstances, when the initial order appeared temporarily correct but was shown to be unwise by quickly emerging information, it may be that the physician should have awaited the additional information before ordering. Difficulties identifying and measuring medication errors are a constant theme of the hospital patient safety literature. Many scholars indicate such difficulties are critical barriers to addressing medication errors. The relative advantages and disadvantages of each identification method are frequently contrasted—comprising observation, voluntary reporting, clinician intervention, computer monitoring, and chart review.

The identification method proposed here circumvents several of the obstacles found in the other methods but lacks the specific focus of each. It does not replace the other methods, but may add a technique that appears both efficient and objective.

With our technique, a third of all discontinued orders within 45 minutes were not found to be inappropriate. If we had used a less conservative measure of inappropriate medication orders, that third would have undoubtedly declined, but concerns for false positives would increase. The alternative explanation, that any order quickly discontinued by its author is by definition a probable error, requires additional examination.

**Limitations**

This study introduces the use of this method to screen for medication errors. There are, accordingly, several limitations: 1) The small sample, which did not allow exploration of the reason for the time slice patterns; 2) Our inability to contact a substantial portion of the physicians (Figure 1); 3) Lack of detailed information on the reasons for discontinuations; 4) A single study site, a teaching hospital where most orders are entered by house officers; and 5) Because of significant non-response, the sample may not be fully representative.

Some physicians we contacted were unaware their orders had been discontinued—a theoretical impossibility given the electronic security system, but nonetheless reported. We believe that data provided to us from the Pharmacy Department were incomplete and did not enable us to identify some of the physicians.

A few physicians could not remember why they discontinued the order, undoubtedly a reflection of the hectic pace when entering many orders for many patients while stopping or modifying other orders. In any event, such information gaps add unwanted noise to our analysis.

Future research should ascertain and enumerate the types of medication orders stopped within 45 minutes; prospectively validate this measure; involve larger samples and analysis of non-respondents; measure inter-rater reliability; incorporate pharmacy logs; systematically determine the reasons why the orders were discontinued; and involve a range of hospitals, including non-teaching hospitals. Also, if this measure is widely adopted, administrators should determine if it is “reactive,” i.e., residents try to avoid rapid order discontinuations to evade scrutiny by senior physicians.
Conclusion
Medication prescribing errors are a prime concern for all in healthcare. Identifying medication prescribing errors is especially difficult because both the errors and their effects are often obscured by the messy reality of illness, computer ordering systems, multifaceted treatments, and the rapid pace and complexity of an acute care hospital. This proposed measure offers a fast, constant, inexpensive, and objective method to identify medication orders that have a high probability of reflecting possible errors. It might be employed as a screening mechanism to find orders for closer analysis or to find physicians-in-training who could benefit from additional mentoring. It would also serve as an extraordinarily focused teaching tool, precisely targeting issues that have emerged in recent cases. The proposed measure can also identify specific rotations, shift configurations, or types of medications that require extra attention.

References