Research Paper ■

Improving Ambulatory Prescribing Safety with a Handheld Decision Support System: A Randomized Controlled Trial

Eta S. Berner, EdD, Thomas K. Houston, MD, MPH, Midge N. Ray, RN, MSN,
Jeroan J. Allison, MD, MPH, Gustavo R. Heudebert, MD, W. Winn Chatham, MD,
John I. Kennedy, Jr., MD, Gerald L. Glandon, PhD, Patricia A. Norton, EdD,
Myra A. Crawford, PhD, Richard S. Maisiak, PhD, MSPH

A b s t r a c t   **Objective:** To evaluate the effectiveness of a personal digital assistant (PDA)–based clinical decision support system (CDSS) on nonsteroidal anti-inflammatory drug (NSAID) prescribing safety in the outpatient setting.

Design: The design was a randomized, controlled trial conducted in a university-based resident clinic. Internal medicine residents received a PDA-based CDSS suite. For intervention residents, the CDSS included a prediction rule for NSAID-related gastrointestinal risk assessment and treatment recommendations. Unannounced standardized patients (SPs) trained to portray musculoskeletal symptoms presented to study physicians. Safety outcomes were assessed from the prescriptions given to the SPs. Each prescription was reviewed by a committee of clinicians blinded to participant, intervention group assignment, and baseline or follow-up status.

Measurements: Prescriptions were judged as safe or unsafe. The main outcome measure was the differential change in unsafe prescribing of NSAIDs for the intervention versus the control group.

Results: At baseline, the mean proportion of cases per physician with unsafe prescriptions for the two groups was similar (0.27 vs. 0.29, $p > 0.05$). Controlling for baseline performance, intervention participants prescribed more safely than controls after receiving the CDSS (0.23 vs. 0.45 [$F = 4.24$, $p < 0.05$]). With the CDSS, intervention participants documented more complete assessment of patient gastrointestinal risk from NSAIDs.

Conclusion: Participants provided with a PDA-based CDSS for NSAID prescribing made fewer unsafe treatment decisions than participants without the CDSS.

■ **J Am Med Inform Assoc.** 2006;13:171–179. DOI 10.1197/jamia.M1961.

There are a variety of methods of providing support for clinical decision making, but computer-based approaches, usually referred to as clinical decision support systems (CDSS), can be broadly defined as computer programs designed to

provide support for clinician decision making. Implementation trials have demonstrated the positive influence of CDSS on physician decisions for medication ordering in the inpatient setting and an expert panel report endorsed by major informatics professional associations has recommended use of CDSS for electronic prescribing.[1–5] These and other data have prompted the Leapfrog Group, a coalition of *Fortune 500* companies advocating major changes to improve patient safety in health care (www.leapfroggroup.org), to include computerized provider order entry (CPOE), with its potential to include clinical decision support, as part of their major "leaps" to improve patient safety in inpatient settings. Despite this evidence of CDSS improving physician prescribing in the inpatient setting, there is a dearth of randomized, controlled trials and, in particular, a lack of information about the impact of CDSS in ambulatory settings.[1]

A common medication error in ambulatory patients that could potentially be reduced by appropriate use of CDSS is the inappropriate use of nonsteroidal anti-inflammatory drugs (NSAIDs).[6,7] Although recent data have raised concerns about the risks of cardiovascular events with selective cyclooxygenase 2 (COX-2) inhibitors,[8–10] these drugs are still

considered safer than traditional NSAIDs in terms of gastrointestinal (GI) risk factors.[11,12] However, many patients with GI risk factors are still prescribed nonselective NSAIDs without gastroprotection (i.e., concomitant treatment with a proton pump inhibitor or misoprostol).[7,13] A decision support system that assists in risk assessment and provides guidance for prescribing could help mitigate these errors.

One approach for addressing clinician information needs in the ambulatory setting is to provide the decision support tools on handheld computers or personal digital assistants (PDAs). Clinicians have begun to use PDAs to provide decision support in the outpatient arena, especially in regard to appropriate medication use.[14–19] A study by Rothschild et al.[17] that surveyed physician users of Epocrates™, a free drug database for handheld computers, found that 63% of the users claimed that the database reduced potential adverse drug events (ADEs), with half estimating it reduced at least one ADE per week. According to data from Epocrates, there are currently almost 200,000 physicians who are actively using their product (M. Snyder, personal communication, March 31, 2005). Although self-report data are encouraging, experimental studies of the impact of handheld decision support systems in the outpatient setting are needed.

Methods

Study Design and Hypotheses

We conducted a randomized, controlled trial to examine the impact of a handheld computer-based clinical decision support program on medication prescribing in a primary care residency education setting. The focus of the decision support was a clinical prediction rule to assess NSAID-related GI risk

and provide real-time treatment recommendations based on the patient's risk.[20] Our primary hypothesis was that clinicians provided with a CDSS that provides recommendations for risk assessment and treatment will prescribe NSAIDs more safely than clinicians without that support. In addition to the obvious effect of the CDSS, providing accurate risk assessment based on clinical factors, using the CDSS could also provide a cue to thorough collection of risk factors in the patient history. Thus, we also examined the impact of the CDSS on participant's gathering key risk factor data.

Participants

Participants were recruited from a pool of 105 internal medicine residents assigned to an urban university-based, resident-staffed clinic, and 68 provided consent and enrolled in the study. Participating residents were randomly assigned to either the intervention or control group by means of a computerized random number generator that ensured that an equal number were in each group (Fig. 1). Participants were informed that the study was to examine their use of PDA-based decision support programs and their impact on care but were blinded to the specific outcome of interest. Participants were told that they would all have slightly different sets of decision support rules but were not aware which rule was the specific focus of the study. The study protocol and consent procedure were approved by the University of Alabama at Birmingham Institutional Review Board.

Details of Study Design

The study design was a randomized trial using preintervention and postintervention assessments of safe and unsafe prescribing. Residents were randomized to intervention and control groups prior to baseline data collection. Outcomes
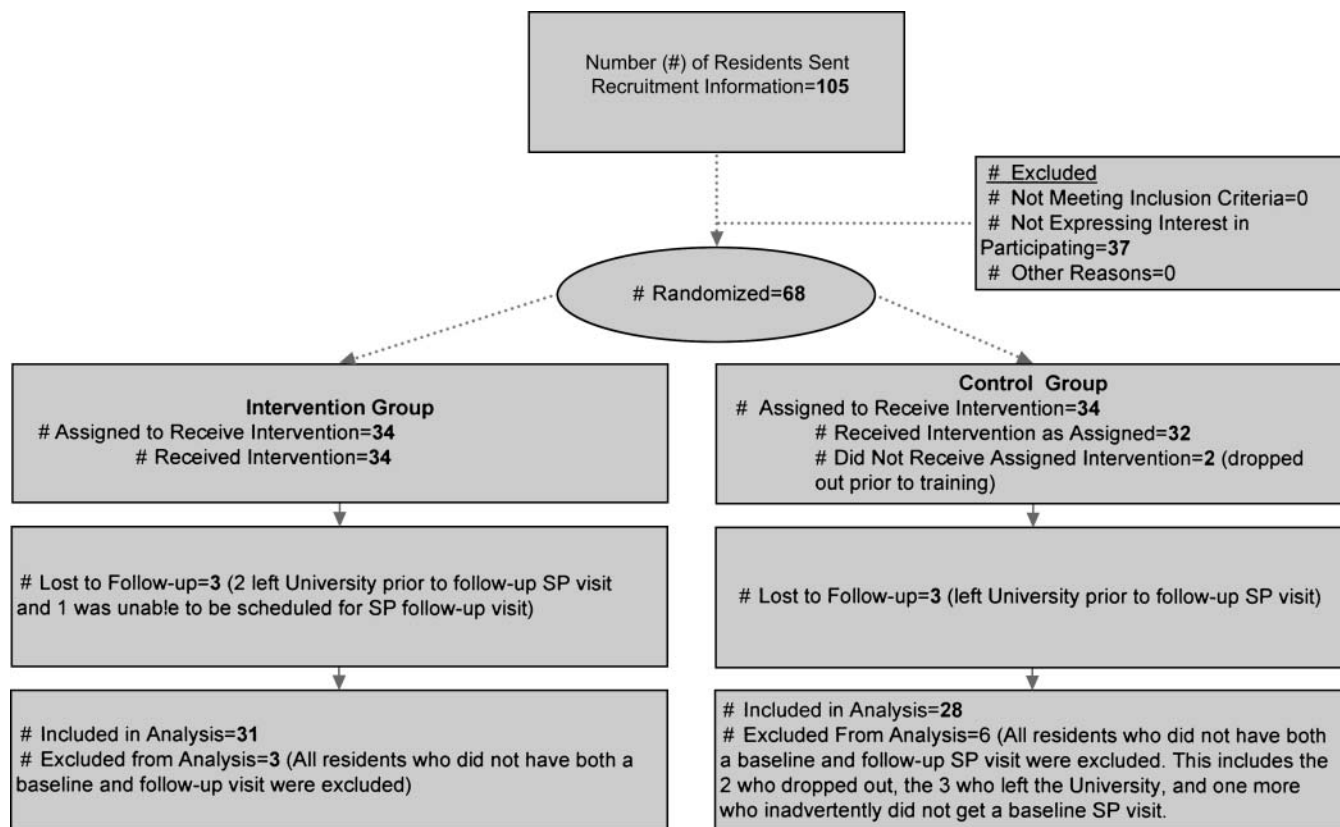

F i g u r e 1. Flow diagram of movement through study.

were assessed by examining the treatment provided for unannounced standardized patients (SPs). Standardized patients are lay individuals trained to portray particular clinical conditions.[21] Participants agreed to see up to four unannounced SPs over a period of time during their outpatient clinic day. Each SP was trained to portray one of four cases that posed a risk for GI hemorrhage if nonselective NSAIDs were prescribed for an extended period of time without gastroprotection. Each participant saw up to two of the four cases at baseline and two at follow-up.

At baseline, all participants were provided with a PDA that ran the Palm Operating System (OS) 4.01. Participants were allowed to load their own programs on the PDA. Approximately six months later, after baseline prescribing performance data were collected, we added the decision support programs to each PDA and trained participants in their use. The primary CDSS for this study, which we entitled MedDecide, was developed by the investigators using Satellite Forms™ Version 4 for Palm OS. MedDecide contained a suite of clinical prediction rules, based on published evidence-based literature.[22–38] There were 19 rules representing diagnostic, risk assessment, and treatment recommendations. Participants were randomized to receive one of six different sets of these rules for their PDA, with each set containing 14 rules. All participants received some rules that could potentially apply to the SP's complaints. Only the participants who were randomized to the intervention arm received, within their set of 14 rules, the rule for GI risk assessment when prescribing NSAIDs (defined as the Intervention Rule). The rest of the rules were distributed randomly in blocks. Participants also received Epocrates, Medcalc, Medmath, and a breast cancer risk calculator, all of which were freely available for use on Palm OS PDAs (www.epocrates.com; www.pdacortex.com/software_palm.htm). A prototype of a PDA diagnostic decision support system that used the knowledge base of the QMR™ diagnostic decision support system was also provided by First Databank. The additional programs, as well as the multiple sets of rules, assisted in masking the specific NSAID-risk rule, which was the focus of the study.

For intervention and control groups, one training session was held for small groups of participants at a time, grouped by intervention status. The training session lasted approximately 30 minutes and included demonstration and instruction on the use of all programs and illustration of some, but not all, of the clinical prediction rules. Demonstration of any given rule took less than a minute. In the intervention group, the demonstrations included a demonstration of the Intervention Rule.

Data collection for the follow-up phase took place over an eight-month period after participants received the CDSS.

Sources of data for both the baseline and follow-up phase were prescriptions and other materials given to the SPs, the participants' dictations and chart notes, and attending physician documentation of participants' initial diagnosis and treatment plan. An outcomes committee, composed of five academic clinicians, reviewed the participants' performance according to explicit criteria. The primary outcome was whether the participant prescribed safely or unsafely. The Intervention Rule, SP case descriptions, procedural details, outcome assessment, and data analysis methods are described below.

The Intervention Rule (Nonsteroidal Anti-inflammatory Drug Gastrointestinal RISK)

The Intervention Rule assessed six established risk factors for GI complications from NSAIDs: age, self-assessed health status, diagnosis of rheumatoid arthritis, steroid use, a history of GI hemorrhage or hospitalization for ulcer, and symptoms with NSAIDs.[20,39–41] This rule has been validated to predict GI hemorrhage and is used by the Veterans Administration as a guideline for NSAID prescribing.[20,42–44] License for the rule, (known as SCORE[©]), which was incorporated into MedDecide, is available from Stanford University. According to the rule, GI risk scores of 15 and below are considered safe for nonselective NSAIDs, while scores 16 and above have been shown to pose at least moderate risk with extended use of nonselective NSAIDs.[20,42] In our adaptation of the rule for the PDA, we allowed ulcer history the same weight as previous hospitalization for ulcer. Users enter all six elements into the PDA via pull-down menus and tap a submit button on the PDA screen to receive the score and recommendation.

Standarized Patient Case Descriptions

We developed three cases for the SPs and adapted one case previously used by Tamblyn et al.[6] in their study using SPs for assessment of NSAID prescribing practices of primary care physicians in Canada. The cases were developed using iterative review as well as pilot testing by chief medical residents who were not part of the study. Each case had at least moderate risk of a GI adverse event according to the NSAID risk assessment rule (Table 1).[43]

Each of our 13 SPs was trained on one case. Training consisted of three sessions, each lasting approximately one hour and included evaluation by a clinician for clinical accuracy prior to portraying the case.

Integration of Standardized Patients into Clinic Workflow

Appointments for SPs were fit into participants' regular schedule of patients during their assigned half-day weekly clinics. The SPs were scheduled as new patients so as not to arouse suspicion about missing medical information.

Table 1 ■ Standardized Patient Case Characteristics

Case	Presenting Symptom	Age, yr	Overall Health Status	Diagnosis	Key Risk* Factor Other Than Age and Health Status	NSAID GI Risk Score
1	Foot pain	66	Fair	Plantar fasciitis	Steroid use	16
2	Knee pain	44	Fair	Osteoarthritis	GI hemorrhage history	16
3	Hip pain	66	Good	Osteoarthritis	Ulcer history	22
4	Shoulder pain	46	Good	Rotator cuff tendinitis	GI hemorrhage history	17

NSAID = nonsteroidal anti-inflammatory drug; GI = gastrointestinal.
*This risk factor was primarily responsible for the higher score, indicating increased risk of NSAID-induced complications.

Residents followed their routine clinic procedures when examining the SPs. This procedure includes examining their patients prior to discussing them privately with their supervising attending physician. The residents then return to their patient and give the patient any handwritten prescriptions, instructions, or order slips they deem appropriate. Documentation of the encounter is done after the patient leaves via dictation, which is subsequently typed and inserted into the paper chart. There is no electronic prescribing, but residents could elect to use the PDA as an information/decision support source at any point in this process.

Immediately after leaving the clinic, the SPs returned to study personnel all prescriptions, medication samples, lab slips, and any other materials they were given. The typed encounter notes were collected when they became available. Any other handwritten resident or nurse's notes, as well as an attending checklist describing participant's initial diagnosis and treatment plan, were also collected after the visit. All collected materials were assembled into a "patient chart" and constituted the documentation for the visit. To assess blinding to the SPs, the participating residents were queried via e-mail every six weeks as to whether they thought they had seen any SPs.

Outcomes Assessment

Patient safety outcomes were determined by a review of documentation from SP encounters. Five physicians with experience in health services research constituted the outcomes committee. Each chart was independently reviewed by two clinicians who were blinded to participant, timing (baseline or follow-up), and group (intervention or control).

Because all the SPs were at risk of NSAID-induced GI complications, specific definitions of duration and dose of NSAIDs were used to define unsafe prescribing. We developed an evidence-based algorithm to classify the treatment regimens. Each prescription was coded as safe or potentially unsafe regarding GI complications. (See Appendix for details of the definition of safe and unsafe prescribing.) Prescriptions or samples handed to the SP were the main source of data, but, in the absence of a prescription or sample, the encounter note or the attending documentation was used. Any disagreements in determination of unsafe prescribing were adjudicated by group review by the committee. Of 189 patient encounters, 16 required group review. Presence or absence of the key risk factor (Table 1) in the notes and/or dictation was also noted.

Statistical Analysis

Fisher exact tests of proportions and t-tests were used to test for significant differences in participant characteristics between groups. At baseline and follow-up, the primary outcome measure for each participant was defined as the proportion of cases per physician with unsafe NSAID prescriptions. For each SP case seen by the participant, the outcomes committee defined safety as a binary variable (unsafe prescription = 1, safe prescription = 0; see Appendix for criteria for safe and unsafe). The case data were aggregated at the participant level. The proportion of cases with unsafe prescriptions for each physician at the baseline or follow-up stage was calculated by dividing the number of unsafe prescriptions by the total number of cases seen at that stage. Consistent with an intention-to-treat analysis design, the data from participants were included regardless of whether they used the clinical decision support software.[45]

For each SP case, we also defined whether the key risk factor driving the GI risk score (Table 1) was obtained (obtained = 1, not obtained = 0). These data were also aggregated at the participant level to calculate the proportion of cases per physician with a key risk factor recorded at baseline or follow-up.

The primary hypothesis of a difference between the two groups in follow-up unsafe prescribing was tested with an analysis of covariance (ANCOVA) using the baseline proportion of cases per physician with unsafe prescriptions as a covariate and participant as the unit of analysis. Correlations were examined between the proportion of cases per physician with unsafe prescriptions and several potential confounding variables such as participant demographics, primary care interest, and whether the participant had obtained the key risk factor for the particular case. A two-tailed α level of 0.05 was used for all statistical tests. Effect sizes were computed using Cohen's method.[46] SPSS was used for statistical analysis.[47]

Results

Participant Characteristics

Fifty-nine of the 68 participants saw at least one SP during both baseline and follow-up periods and were considered to have completed the study. Participant characteristics were similar in the intervention and control groups (Table 2). Intervention participants were slightly younger (1.2 years), and more frequently first-year residents (PGY-1), compared with control participants (p < 0.05 for both). Overall there was no evidence that the 59 study completers differed significantly from the nine noncompleters.

Unsafe Prescribing at Baseline

At baseline, the mean proportions (0.27 vs. 0.29) of cases per physician with unsafe prescriptions were similar (p > 0.05) for the intervention and control groups. Although the two groups differed in distribution of participants in different residency years, the correlation of residency year with proportion of cases per physician with unsafe prescriptions at both baseline (r = 0.01) and follow-up (r = 0.01) was low and non-significant (p > 0.05) as was the correlation with age (baseline, r = 0.07; follow-up, r = 0.21).

Changes in Unsafe Prescribing

The mean number of SP visits per participant at baseline was 1.34 visits (SD = 0.48) and at follow-up it was 1.56 visits (SD = 0.50). There was no statistically significant difference between intervention or control groups in number of visits

Table 2 ■ Participant Characteristics

Characteristic	Intervention Group	Control Group
Male, no./total no. (%)	23/31 (74%)	20/28 (71%)
Ethnicity, no./total no. (%)		
White, non-Hispanic	24/31 (77%)	21/28 (75%)
Other	7/31 (23%)	7/28 (25%)
Plan to practice primary care, no./total no. (%)	4/31 (13%)	5/28 (18%)
Mean age, yr (standard deviation)	27.35 (2.18)	28.57 (2.28)
Postgraduate year, no./total no. (%)		
1	14/31 (45%)	4/28 (14%)
2	9/31 (29%)	13/28 (47%)
3	8/31 (26%)	11/28 (39%)

at either time period. Figure 2 shows the mean proportion of cases per physician with unsafe prescriptions for each group at baseline and follow-up. The mean proportions of the intervention and control groups were similar at baseline, but the mean proportion increased greatly for the control group at follow-up. Using participant as the unit of analysis, the ANCOVA showed a significantly lower mean proportion of cases per physician with unsafe prescriptions for the intervention group compared to the control group after adjustment for baseline rates (F = 4.24, p < 0.05, effect size = 0.54). There were no statistically significant associations with other potential confounders of the primary hypothesis such as any participant characteristics, SP gender, or recognition of the SP by the participant. The mean proportions were also not significantly different among the different presenting complaints, and there was no group-chief complaint interaction effect.

Key Risk Factor Assessment

The intervention group more frequently obtained the key risk factor during follow-up (0.58 vs. 0.45), even though they tended to be worse at baseline (0.69 vs. 0.74), but these differences were not statistically significant. We did note that the follow-up mean proportion of cases per physician with unsafe prescriptions tended to be lower among those participants who collected the key risk factor (r = −0.23, p < 0.10), compared with those who did not get that information. This association was stronger and statistically significant for the intervention group (r = −0.43, p < 0.05).

Because of the statistically significant association between unsafe prescribing and failing to obtain the key risk factor in the intervention group, we examined whether the statistically significant differences in proportion of cases per physician with unsafe prescriptions between the intervention and control group participants were still present if we only included cases where the subjects had obtained the key risk factor. Twenty-nine (12 in the control group and 17 in the intervention group) participants recorded the key risk factor for at
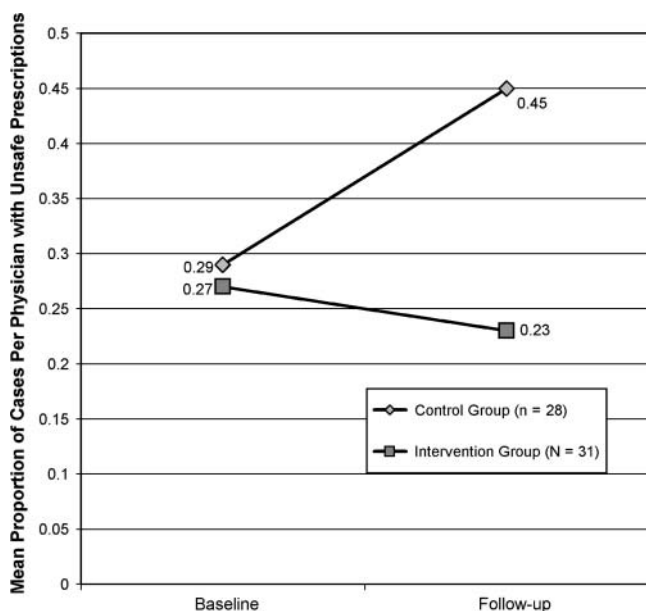
least one of their cases at both baseline and follow-up. The mean proportions of cases per physician with unsafe prescriptions by assessment time and group for this subset of participants who recorded the key risk factor are presented in Figure 3.

At baseline, the intervention and control groups' performance appeared similar and only slightly better than that of the total group. At follow-up, the intervention group's mean proportion of unsafe prescriptions on cases where the key risk factor was recorded was zero, while the control group's mean proportion of unsafe prescriptions on such recorded cases was 0.50. Even with the smaller number of subjects, the ANCOVA showed that the mean proportion of cases per physician with unsafe prescriptions of the intervention group at follow-up was significantly (p < 0.001, effect size = 1.12) lower than that of the control group.

Discussion

We conducted a randomized, controlled trial of the impact of a CDSS designed to assist with GI risk assessment and treatment recommendations when prescribing NSAIDs. We demonstrated improved clinical safety in the intervention group. Specifically, the mean proportion of cases per physician with unsafe prescriptions in the intervention group was approximately one-half that of physicians without access to the CDSS. Our analysis of key risk factors suggests that, at least in part, the CDSS acted as a cue to collect appropriate patient history related to risk. In fact, among those residents who had the Intervention Rule and obtained the key risk factor, there were no unsafe prescriptions for the SP visits. Thus, our results suggest that the combination of collection of data on risk factors and accurate assessment of clinical risk with treatment recommendations resulted in the safest prescribing.

Although all the SPs in the present study had risk factors for GI complications of NSAIDs, the same decision support tool could as well be used to identify patients who are at low risk of adverse GI events and who would not need gastroprotective medication. Given the recent data on the risks of

F i g u r e 2.   Mean proportion of cases per physician with unsafe prescriptions for intervention and control groups at baseline and follow-up.
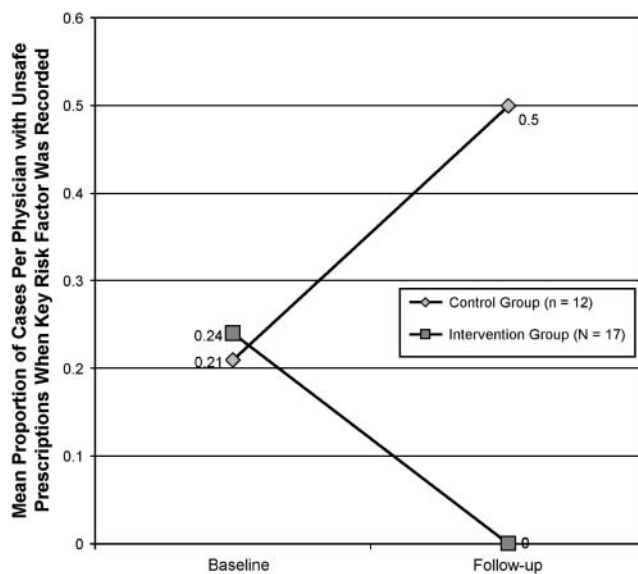
F i g u r e 3.   Mean proportion of cases per physician with unsafe prescriptions for intervention and control groups at baseline and follow-up when key risk factor was recorded.

COX-2 inhibitors[8–10] and the evidence of inappropriate prescribing of these drugs for low-risk patients, as well as the prescribing of nonselective NSAIDs for high-risk patients,[7] decision support tools at the point of care may be especially useful in this area.

There are several possible ways by which the intervention might have influenced performance. Standardized patients reported minimal use of the PDA in their presence, but participants may have used the rule out of the SP's presence since they left the room to consult with the attending physician. Although we initially assumed that the suite of rules on the PDA was complex enough that participants would not memorize the rules, it is possible that the brief training session where they saw the rule or perhaps the use of the rule with other patients prior to the follow-up of SPs was sufficient to sensitize the intervention group participants to the importance of the risk factors and the recommendations for safe prescribing. However, most data show that the impact on performance of even more extensive didactic instruction is minimal.[48]

Although we anticipated that the CDSS would reduce inappropriate prescribing, the effect was more complicated. Instead of the performance of the intervention group improving significantly over time, their performance remained relatively stable, while overall the control group performance degraded over the time period from baseline to follow-up. Other data have shown that housestaff thoroughness, especially for preventive measures, history, or other data not closely related to the acute presenting problem, may decrease as they progress through training.[49] The present study indicated that the CDSS may have minimized that performance degradation in the intervention group.

It is likely that more intensive decision support that does not have to rely on clinicians choosing to use it may be needed to have a stronger effect. Rosenbloom et al.,[50] for instance, found that active decision support was accessed more frequently than the passive decision support that was similar to the type we provided.

Limitations

The limitations of the study were that it was conducted in a single clinic with participants within the same specialty and a small number of cases may limit generalizability. Although formal data on PDA use were not collected on those who chose not to participate, we do know that at least some of them already had PDAs and were not interested in changing to a new one, but it is possible that others were not interested in using the technology. Those who did choose to participate were, by definition, receptive to learning to use the technology. It is possible that this group may have implicitly understood that use was expected, whereas a nonvolunteer sample would not use it as frequently. However, we emphasized that they should use it often as they felt it was needed, and we made the training as minimal as possible to limit this effect. In addition, in regard to the primary outcome of interest in this study, the baseline results showed that the magnitude of unsafe NSAID prescribing was only slightly lower than that found in the study by Tamblyn et al.,[6] which also used SPs.

The absence of reliable data on how the participants actually used the CDSS during the SP visits was a limitation as well.

Finally, our focus was on unsafe prescribing that could *potentially* lead to adverse events. Even with real patients, the errors that occurred would not automatically lead to actual harm to patients, and with proper follow-up and patient education adverse events could be averted. Most of the participants did arrange for reasonable follow-up appointments for the SPs, and several of them provided some instructions about indications for discontinuing the medication.

Although our assessment of unsafe prescriptions was based primarily on prescriptions and other materials actually given to the SP, the assessment of whether the key risk factor was obtained was based on chart review. Our rationale was that these data would indicate that the resident had carefully considered the information, but it is certainly possible that the data may have been obtained and not documented.

Using SPs provided us with realistic, but controlled, case presentations. Research has shown that SPs are a valid method of obtaining data on the process of care, but some studies have shown that detecting SPs could potentially show artificially improved performance.[6,21,51,52] Unlike other studies using SPs, which usually relied on physicians remembering to send in a postcard to the study team when they thought they had detected an SP, we actively queried the residents during the study as to potential SP detection and also conducted a debriefing at the end as to their understanding of the study. Not only was there no significant association between detection and performance, there was no differential detection in intervention and control groups, and participant feedback indicated that even when they detected that the patients were SPs, they did not recognize that the focus of the study was on risk assessment and NSAID prescribing. In addition, participants in both the intervention and control groups had other CDSS and rules within MedDecide that related to other aspects of the cases that may have made the GI risk assessment rule less obvious as a focus of the study. Any heightened diligence brought on by SP detection would likely bias the study against finding the primary effect.

Conclusion

Using simple, easy-to-access programs at the point of care can affect clinician performance in the ambulatory setting. Participants randomized to have a clinical decision support rule on a handheld computer that they could use at the point of care prescribed more safely on relevant cases than participants who did not have the rule. The PDA-based rule appeared to inhibit the increased tendency for unsafe NSAID prescribing that occurred over time for physicians without access to the rule. The participants who had the CDSS overall gathered and interpreted risk factor information more appropriately than those participants without the CDSS. Although a stand-alone decision support system does not have the added benefit of integration with an electronic medical record, at present, this system is far more easily disseminated. Further research should confirm our results and assess the potential reach of such systems.

References ■

1. Garg AX, Adhikari NK, McDonald H, et al. Effects of computerized clinical decision support systems on practitioner performance and patient outcomes: a systematic review. JAMA. 2005;293:1223–38.

2. Bates DW, Leape LL, Cullen DJ, et al. Effect of computerized physician order entry and a team intervention on prevention of serious medication errors. JAMA. 1998;280:1311–6.

3. Kawamoto K, Houlihan CA, Balas EA, Lobach DF. Improving clinical practice using clinical decision support systems: a systematic review of trials to identify features critical to success. BMJ. 2005;330:765.

4. Kucher N, Koo S, Quiroz R, et al. Electronic alerts to prevent venous thromboembolism among hospitalized patients. N Engl J Med. 2005;352:969–77.

5. Teich JM, Osheroff JA, Pifer EA, Sittig DF, Jenders RA, the CDS Expert Review Panel. Clinical decision support in electronic prescribing: recommendations and an action plan. J Am Med Inform Assoc. 2005;12:365–76.

6. Tamblyn R, Berkson L, Dauphinee WD, et al. Unnecessary prescribing of NSAIDs and the management of NSAID-related gastropathy in medical practice. Ann Intern Med. 1997;127:429–38.

7. Solomon DH, Schneeweiss S, Glynn RJ, Levin R, Avorn J. Determinants of selective cyclooxygenase-2 inhibitor prescribing: are patient or physician characteristics more important? Am J Med. 2003;115:715–20.

8. Solomon SD, McMurray JJ, Pfeffer MA, et al. Cardiovascular risk associated with celecoxib in a clinical trial for colorectal adenoma prevention. N Engl J Med. 2005;352:1071–80.

9. Nussmeier NA, Whelton AA, Brown MT, et al. Complications of the COX-2 inhibitors parecoxib and valdecoxib after cardiac surgery. N Engl J Med. 2005;352:1081–91.

10. Bresalier RS, Sandler RS, Quan H, et al. Cardiovascular events associated with rofecoxib in a colorectal adenoma chemoprevention trial. N Engl J Med. 2005;352:1092–102.

11. Bombardier C, Laine L, Reicin A, et al. Comparison of gastrointestinal toxicity of rofecoxib and naproxen in patients with rheumatoid arthritis. VIGOR study group. N Engl J Med. 2000;343:1520–8.

12. Silverstein FE, Faich G, Goldstein JL, et al. Gastrointestinal toxicity with celecoxib vs nonsteroidal anti-inflammatory drugs for osteoarthritis and rheumatoid arthritis: the CLASS study: a randomized controlled trial. Celecoxib long-term arthritis safety study. JAMA. 2000;284:1247–55.

13. Smalley W, Stein CM, Arbogast PG, Eisen G, Ray WA, Griffin M. Underutilization of gastroprotective measures in patients receiving nonsteroidal antiinflammatory drugs. Arthritis Rheum. 2002;46:2195–200.

14. Rosenbloom M. Medical error reduction and PDAs. Int. Pediatr. 2003;18:60–77.

15. McLeod TG, Ebbert JO, Lymp JF. Survey assessment of personal digital assistant use among trainees and attending physicians. J Am Med Inform Assoc. 2003;10:605–7.

16. Miller RH, Hillman JM, Given RS. Physician use of IT: results from the Deloitte research survey. J Healthc Inf Manag. 2004;18:72–80.

17. Rothschild JM, Lee TH, Bae T, Bates DW. Clinician use of a palm-top drug reference guide. J Am Med Inform Assoc. 2002;9:223–9.

18. Cimino JJ, Bakken S. Personal digital educators. N Engl J Med. 2005;352:860–2.

19. Barrett JR, Strayer SM, Schubart JR. Assessing medical residents' usage and perceived needs for personal digital assistants. Int J Med Inf. 2004;73:25–34.

20. Singh G, Ramey DR, Triadafilopolous G. GI score: a simple self-assessment instrument to quantify the risk of serious NSAID-related GI complications in RA and OA [abstract] Arthritis Rheum. 1998;41(Suppl. 9):S75.

21. Luck J, Peabody JW. Using standardised patients to measure physicians' practice: validation study using audio recordings. BMJ. 2002;325:679.

22. Williams JW, Simel DL, Roberts L, Samsa GP. Clinical evaluation for sinusitis: making the diagnosis by history and physical examination. Ann Intern Med. 1992;117:705–10.

23. Whooley MA, Avins AL, Miranda J, Browner WS. Case-finding instruments for depression: two questions are as good as many. J Gen Intern Med. 1997;12:439–45.

24. American Diabetes Association. Summary of revisions for the 2003 clinical practice recommendations. Diabetes Care. 2003;261(Suppl):S33–50.

25. Herman VH, Smith PJ, Thompson TJ, Engelgau MM, Aubert RE. A new and simple questionnaire to identify people at increased risk for undiagnosed diabetes. Diabetes Care. 1995;18:382–7.

26. Lydick E, Cook K, Turpin J, et al. Development and validation of a simple questionnaire to facilitate identification of women likely to have a low bone density. Am J Managed Care. 1998;4:37–48.

27. Stiell IG, Greenberg GH, McKnight RD, et al. Clinical prediction rules for the use of radiography in acute ankle injuries: refinement and prospective validation. JAMA. 1993;269:1127–32.

28. Stiell IG, Greenberg GH, Wells GA, et al. Prospective validation of a clinical prediction rule for the use of radiography in acute knee injuries. JAMA. 1996;275:611–5.

29. Spiegelhalter DJ, Crean GP, Holden R, Knill-Jones RP. Taking calculated risk: predictive scoring systems in dyspepsia. Scand J Gastroenterol. 1987;128(Suppl):152–60.

30. Wells PS, Anderson DR, Rodger M, et al. Derivation of a simple clinical model to categorize patients probability of pulmonary embolism: increasing the models utility with the simpliRED D-dimer. Thromb Haemost. 2000;83:416–20.

31. Cooper RJ, Hoffman JR, Barlett JG, et al. Principles of appropriate antibiotic use for acute pharyngitis in adults. Ann Intern Med. 2001;134:509–17.

32. SPAF Investigators. Predictors of thromboembolism in atrial fibrillation: I. Clinical features of patients at risk. Ann Intern Med. 1992;116:1–5.

33. Martin TP, Hanusa BH, Kapoor WN. Risk stratification of patients with syncope. Ann Emerg Med. 1997;29:459–66.

34. Wigton RS, Hoellerich VL, Ornato JP, Leu V, Mazzotta LA, Cheng IH. Use of clinical findings in the diagnosis of urinary tract infection in women. Arch Intern Med. 1985;145:2222–7.

35. Buchsbaum DG, Buchanan RG, Welsh J, Centor RM, Schnoll SH. Screening for drinking disorders in the elderly using the CAGE questionnaire. J Am Geriatr Soc. 1992;40:662–5.

36. Wells PS, Hirsh J, Anderson DR, et al. Accuracy of clinical assessment of deep-vein thrombosis. Lancet. 1995;345:1326–30.

37. Fine MJ, Auble TE, Yealy DM, et al. A prediction rule to identify low-risk patients with community-acquired pneumonia. N Engl J Med. 1997;336:243–50.

38. U.S. Public Health Service. Treating tobacco use and dependence—a guide for health care administrators, insurers, managed care organizations, and purchasers. November 2000. Available from: http://www.surgeongeneral.gov/tobacco/systems.htm/. Accessed September 1, 2005.

39. Fries JF, Miller SR, Spitz PW, Williams CA, Hubert HB, Bloch DA. Toward an epidemiology of gastropathy associated with nonsteroidal antiinflammatory drug use. Gastroenterology. 1989;96(Suppl):647–55.

40. Simon LS, Hatoum HT, Bittman RM, Archambault WT, Polisson RP. Risk factors for serious nonsteroidal-induced gastrointestinal complications: regression analysis of the MUCOS trial. Fam Med. 1996;28:204–10.

41. Wolfe MM, Lichtenstein DR, Singh G. Gastrointestinal toxicity of nonsteroidal anti-inflammatory drugs. N Engl J Med. 1999;340:1888–99.

42. Cheetham TC, Levy G, Spence M. Predicting the risk of gastrointestinal bleeding due to nonsteroidal antiinflammatory drugs: NSAID electronic assessment of risk. J Rheumatol. 2003;30:2241–4.

43. Fries JF, Bruce B. Rates of serious gastrointestinal events from low dose use of acetylsalicylic acid, acetaminophen, and ibuprofen in patients with osteoarthritis and rheumatoid arthritis. J Rheumatol. 2003;30:2226–33.

44. Silver DS. The use of COX-1-sparing agents in the federal health system. Pharmacy & Therapeutics. 2004;29:454–60.

45. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ. 1999;319:670–4.

46. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum, 1988.

47. Norušis MJ. SPSS 11.0 guide to data analysis. Upper Saddle River, NJ: Prentice Hall, 2002.

48. Davis DA, Thomson MA, Oxman AD, Haynes RB. Changing physician performance: a systematic review of the effect of continuing medical education strategies. JAMA. 1995;274: 700–5.

49. Kogan JR, Reynolds EE, Shea JA. Effectiveness of report cards based on chart audits of residents' adherence to practice guidelines on practice performance: a randomized controlled trial. Teach Learn Med. 2003;15:25–30.

50. Rosenbloom ST, Geissbuhler AJ, Dupont WD, et al. Effect of CPOE user interface design on user-initiated access to educational and patient information during clinical care. J Am Med Inform Assoc. 2005;12:458–73.

51. Tamblyn R, Huang A, Perreault R, et al. The medical office of the 21st century (MOXXI): effectiveness of computerized decision-making support in reducing inappropriate prescribing in primary care. CMAJ. 2003;169:549–56.

52. Kahan M, Liu E, Borsoi D, et al. Family medicine residents' performance with detected versus undetected simulated patients posing as problem drinkers. Med Educ Online. 2004;9:18. Available from: http://www.med-ed-online.org/pdf/res00081.pdf/. Accessed September 1, 2005.

Appendix

The following treatment plans were classified as comparatively safe for the particular cases in terms of minimizing the probability of GI complications: (1) nonpharmacological intervention such as exercise or physical therapy, (2) acetaminophen; (3) nonacetylated salicylates; (4) COX-2 inhibitors; (5) ≤30 days of intended treatment with prescription nonselective NSAIDs; (6) >30 days of treatment with prescription nonselective NSAIDs accompanied by adequate GI prophylaxis (i.e., use of proton pump inhibitor or misoprostol for the entire duration of NSAID treatment); (7) recommendation of over-the-counter NSAID for ≤30 days; (8) recommended over-the-counter NSAID for >30 days with clinical follow-

up scheduled within 30 days. Potentially unsafe treatment plans included (1) use of traditional prescription NSAIDs for >30 days with no gastroprotection; (2) use of traditional prescription NSAIDs for >30 days with inadequate gastroprotection; (3) recommendation of over-the-counter NSAIDs for >30 days with no clear recommendation for clinical follow-up within 30 days. A clear recommendation for over-the-counter use at prescription dose was treated as a prescription. The flow diagram below outlines these decisions. NSAIDS = nonsteroidal anti-inflammatory drugs; s = safe; u = unsafe; f/u = follow-up; OTC = over-the-counter; COX-2 = cyclooxygenase 2; GI = gastrointestinal; P.T. = physical therapy.