Integrating Contextual Factors into Clinical Decision Support to Reduce Contextual Error, Improve Outcomes, and Reduce Misuse, Overuse, and Underuse of Outpatient Primary Care

Saul J. Weiner, MD[1-3]*, Alan Schwartz, PhD[1], Frances Weaver PhD[3-5], William Galanter MD,PhD[1], Sarah Olender MS, RDN[1], Karl Kochendorfer MD[1], Amy Binns-Calvey BA[1-3], Ravisha Saini BA[1,2], Sana Iqbal BS, BA[5], Monique Diaz[1,7], Aaron Michelfelder MD[5,6], Anita Varkey MD[6,8]

* Principal Investigator

1. University of Illinois at Chicago (Primary site)
2. Jesse Brown VA Medical Center
3. Center of Innovation for Complex Chronic Healthcare
4. Edward Hines, Jr. VA Hospital
5. Loyola University Chicago
6. Loyola University Health System (Secondary site)
7. Pacific Central Coast Health Centers, Dignity Health
8. Oak Street Health

Project Period: 07/01/2017 - 04/30/2022
Project Officer: Roland Gamache, roland.gamache@ahrq.hhs.gov

This study was funded by the Agency for Healthcare Research and Quality (AHRQ).
Grant No.: R01 HS25374

Abstract:
Purpose: Assess whether customized clinical decision support tools in the electronic health record can improve clinician contextualization of care and, in so doing, improve patient healthcare outcomes, and reduce cost of inappropriate care.

Scope: Two ambulatory clinics employing two different EHRs. Employ both real patients and unannounced standardized patients (USPs).

Methods: A randomized controlled trial (RCT) and a USP study portraying 4 cases. (1): Outcomes measures for RCT were proportion of contextual red flags that resolved 6 months post-index encounter, proportion of red flags probed, proportion of contextual factors addressed. (2): Outcome measure for the USP study were costs attributed to overuse, misuse, underuse.

Results: Across 452 patients encounter, contextual red flags were not more likely to resolve in the intervention vs. control (AOR 0.97, 95% CI 0.57 – 1.64). However, the intervention increased both contextual probing (AOR 2.1, 1.1-3.9) and contextualization of care (AOR 2.7, 1.3-5.4) controlling for whether a factor was identified by probing or otherwise. Across study arms, contextualized care plans were more likely to result in improvement in the presenting red flag (AOR 2.1, 1.4-3.3). A reduction in cost of inappropriate care was seen in one of four USP cases across 41 visits.

Key Words: Contextual errors, Clinical Decision Support

Purpose

The term *patient context* refers to the myriad *contextual factors* in patients' lives that complicate the application of research evidence to patient care.[1] For instance, the inability of a patient to afford an indicated medication for a clinical condition is a contextual factor. Contextual factors can be addressed when correctly identified. For instance, substituting a low-cost generic for a high-cost brand name medication may enable a patient to afford a medication. Addressing contextual factors in a care plan is termed *contextualizing care*.[2] Conversely, the failure to address a contextual factor when it is feasible to do so is a *contextual error*, because it results in an inappropriate plan of care.[3] In sum, contextual errors are medical errors caused by inattention to patient context. They are common, and linked to both diminished health care outcomes[4] and an increase in health care costs related to overuse and misuse of medical services.[5] These findings were determined using a validated method for coding audio recorded data called Content Coding for Contextualization of Care ("4C")[6] collected during the encounters by both real patients, and by unannounced standardized patients (USPs) employing checklists.[7]

Preventing contextual errors requires enhancing clinician responsiveness to clues that there are contextual factors during the clinical encounter, in real time.[8,9] These clues, termed *contextual red flags* are evident in two sources: the medical record and from patients directly.[10] An unexpected increase in glycosylated hemoglobin is an example of the former; a comment by a patient that "it's been tough adapting to vision loss" reflects the latter. An effective intervention would prompt clinicians to determine whether there are underlying contextual factors that could be addressed in the care plan (e.g., prefilled syringes), averting contextual error. This desirable process is termed *contextual probing*.[6]

While clinical decision support (CDS) has been used to provide physicians with timely biomedical information at the point of care to prevent errors[11-13] and promote appropriate care,[14-16] this technology also has the potential to alert physicians to both contextual red flags and contextual factors to avert contextual errors. In this study we aimed to assess the potential of "contextualized CDS" to improve contextualization of care and health care outcomes through a randomized controlled intervention trial. In parallel we conducted a study employing USPs to measure the effect of the intervention on costs related to overuse, misuse and underuse of medical services. The objective was to pursue these aims by testing three hypotheses that contextualized CDS can:

1. *Reduce contextual error* by informing clinicians of contextual factors and prompting them to explore contextual red flags.
2. *Improve health care outcomes* by facilitating partial or full resolution of the contextual red flag (e.g. elevated HgB A1c) after the index visit.
3. *Reduces avoidable health care costs* by reducing misuse, overuse, and underuse of inappropriate or unnecessary medical services.

Scope

Our team has spent over a decade characterizing contextual errors (what they are and how to detect them), assessing their prevalence in various practice settings, measuring their impact on health care outcomes and costs, and trying to prevent them.[8,9] For the latter we have attempted medical education interventions[4,17] and performance improvement strategies employing audit & feedback.[18] A common theme of all of this work has been that contextual errors occur when physicians overlook essential information about patients' circumstances and behaviors when planning their care, with measurably deleterious consequences for both health care outcomes and costs.[19] Reducing contextual

error rates may require real time strategies, activated during the clinical encounter, that prompt physicians to explore and address patient context in care planning.

Specifically, contextual errors are avoided by *contextualizing care,* a four step process:[6] In the first step the clinician notices clues of unresolved challenges in their patient's health or health care that could be attributed to potentially remediable circumstances in their life situation. These are termed *contextual red flags*. An unexplained loss of diabetes control is a contextual red flag because it suggests a change in diet, activity level or medication adherence may be the reason for the loss of control, all of which are typically related to a change in life circumstances. The second step is asking relevant questions, termed *contextual probes*, e.g. "Mr. Jones, I notice that your diabetes control has deteriorated quite a bit. Is something going on that you think might be the reason for this?" The third is assessing whether the information elicited is, in fact, pertinent to the unresolved challenges, which defines it as a *contextual factor*. For example, were Mr. Jones to reply, "It's been tough managing my insulin ever since I started working the night shift since I eat at different times than I used to," a clinician should recognize that working the night shift is a "contextual factor" because it is pertinent to loss of diabetes control. Finally, the fourth step is arriving at a *contextualized plan of care* that addresses the identified contextual issues. For Mr. Jones it would likely involve a modified medication dosing schedule that is aligned with his altered eating and sleeping cycle and takes into account any practical constraints in his work environment.

Assessing whether a clinician contextualizes care or, conversely, makes a contextual error requires listening in on the visit. There are two practical ways to do this assessment. One approach involves employing unannounced standardized patients (USPs), who present as real patients.[20] The USPs are trained to portray common ambulatory problems complicated by contextual factors as evidenced by contextual red flags. For instance, a USP might present with recent worsening of his asthma symptoms and report that "Boy, it's been tough since I lost my job," a contextual red flag that his worsening asthma symptoms may be related to job loss and possible loss of health insurance. If the doctor probes, the USP reveals that he has in fact recently lost his job and has consequently reduced usage of an expensive brand name inhaled corticosteroid (ICS) so that it lasts longer (a contextual factor). The astute physician will suggest switching the patient to a less costly brand of the same medication and/or apply for medication assistance (contextualizing care). In contrast, a physician who is inattentive to context -- overlooking either the contextual red flag or the contextual factor -- simply increases the dosage or adds another medication to a regimen the patient is already unable to follow (a contextual error). Sending the same USP to many physicians is a strategy for comparing physician performance at contextualizing care under the same conditions.

Following a USP visit, it is possible to see how contextual errors impact care planning by looking at the orders the physician placed (before notifying them that they saw a USP). For instance, in one script, we deployed a 73 year old USP who visited 50 physicians complaining of unexplained weight loss and gave four clues that he was homeless and food insecure (contextual red flags).[7] When physicians identified the underlying cause of his weight loss, they referred him to social services and resources such as Meals on Wheels. When they overlooked the contextual red flags, they typically ordered an extensive work up for malignancy, including a CT scan, colonoscopy and chest X-ray. Such a care plan illustrates how a contextual error can lead to a misuse or overuse of medical services.

The second method for assessing physician performance at contextualizing care is by inviting real patients to carry concealed audio recorders into their visits.[4] Whereas the USP methodology is designed to measure the likelihood that a clinician will make a contextual error when given a standardized opportunity to do so, the real patient method enables measurement of contextual error rates in actual practice. Another benefit is that with real patients it is feasible to compare the outcomes of those who received contextualized care to those whose care plan contained a contextual error.[4] For instance, among patients with poorly controlled diabetes, how might identifying and addressing

contextual factors impact their glycosylated hemoglobin levels over time compared to those whose care plan consists of merely increasing their prescribed insulin or other medications? Answering this question involves prospectively coding encounters with contextual red flags and identified contextual factors for the presence or absence of contextual errors in care planning and following those patients over a pre-determined period of time to determine the disposition of the contextual red flag recorded at the index visit (e.g. did the glycosylated hemoglobin level rise or fall?).

A challenge of employing real patients, however, is that – in contrast to USPs – the criteria for defining whether a care plan is contextualized or contains a contextual error is not predefined. In fact, in many encounters there are no contextual red flags or factors at all. Hence to assess contextualization of care in real patients we developed a coding system that combines chart review with listening to the audio recording, called "Content Coding for Contextualization of Care" or "4C." 4C has nearly 90% inter-rater agreement in coding for the presence of a contextual error.[6]

In sum, USPs and patient collected audio are complementary strategies for studying contextual error: the former is ideal for "apples to apples" comparison of clinician performance, because USPs provide intrinsic risk adjustment, and for assessing the impact of contextual errors on service utilization. The latter enables measurement of contextual error rates in clinical settings with varying proportions of patients with contextual factors impacting their care, and the impact of attending or not attending to those factors on real health care outcomes. Finally, both are powerful tools for measuring the impact of interventions to reduce contextual errors because they are authentic measures of clinician performance since the clinician is unaware at the time that they are being observed.

What we've learned from real patient and USP collected audio about contextual error:

In our research employing real patient collected audio we learned that contextual errors are common. In a study in which 601 patients carried concealed audio recorders into their visits across multiple practice sites, we found that contextual red flags were present in 403 of visits (67%), and that contextual factors were revealed in 208, meaning that in 35% of encounters effective care required identifying and addressing a contextual factor.[4] Physicians were successful about 59% of the time, and responsible for a contextual error in the remaining 41%. In other words, about 14% (0.41 x 35%) of overall care was derailed by a contextual error. When we followed these patients for 9 months, the presenting problem at the time of the index visit was less likely to improve or resolve compared to visits without a contextual error (46% vs 71%; P= 0.002).

In our research employing USPs, we documented similar performance problems, with high contextual error rates.[7] These errors are caused either by inattention to contextual red flags – i.e. not noticing or responding to clues of underlying contextual factors, or not addressing contextual factors in care planning. The cases we developed were designed such that physicians were also challenged to avoid making biomedical errors, e.g. overlooking evidence of gastroesophageal reflux in a patient with asthma presenting with increased symptoms after meals and when recumbent. Before deploying USPs, the cases were iteratively refined until board certified physicians reviewing paper based versions had a low probability of making either a biomedical or contextual error when explicitly informed of the contextual factor.[10] In situ, however, contextual error rates turned out to be both common and more frequent than biomedical errors. In a subsequent analysis we added up the direct service utilization costs of these errors using Medicare cost-based reimbursement data, by tabulating the expenses associated with misuse and overuse of medical services.[5] Over 400 encounters, biomedical errors contributed a mean cost of $30 per encounter, and contextual errors $231 per encounter.

Preventing contextual errors:

Given that contextual errors are frequent, harmful and costly, preventing them should be a priority. Prior to this study we had tried two approaches. The first involves educating clinicians through a

didactic and experiential series of workshops to identify and address contextual factors in care planning. In one study we enrolled fourth year medical students during the internal medicine sub-internships, randomizing them to participate in the workshop series or to the usual curriculum.[17] At the end of the one month block, both intervention and control were assessed in a standardized patient laboratory, with standardized patients presenting the same cases we had used in our USP studies, except the participants knew these were standardized patients (known-SPs) and that they were being assessed. The findings showed a significant improvement in contextualization of care in the intervention compared with the control.

With these favorable findings in hand, we repeated the study with internal medicine residents, but this time we assessed participants both with known-SPs and real patients.[4,21] We discovered that while the known-SP findings were replicated, there was no evidence that participants in the intervention outperformed the control when assessed with real patients. This finding demonstrated a skills to performance gap.[21] In other words, participation in the workshops built skills at contextualizing as measured when clinicians knew they were being tested, but dissipated in the actual practice environment when they were providing their usual care. These and other comparisons of known-SP vs real patient assessment have affirmed the importance of direct and unannounced observation to accurately assess performance.

Based on the findings that an educational intervention is not sufficient to change practice, we introduced an audit and feedback strategy to improve contextualization of care.[18]Since its inception in 2012, thousands of patients within the Department of Veterans Affairs, across seven facilities, have volunteered to carry concealed audio recorders into their visits. As they exit their appointment, they return the audio recorder which is uploaded to a secure server so that it can be reviewed along with the medical record for contextual red flags, contextual factors, and a determination about whether the care plan was contextualized or reflected one or more contextual errors. These data are then shared with the care teams, so that they can observe their own best and worst practices. They are also provided with cumulative trend data illustrating their performance at probing contextual red flags and incorporating contextual factors into their care plan.

From May 1017-May 2019, we conducted a step-wedge trial of the quality improvement program. In the study, which was published in JAMA Network Open Access, patients (mean age, 62.0 years; 92%male) recorded 4496 encounters with 666 clinicians. At baseline, clinicians addressed 413 of 618 contextual factors in their care plans (67%). After either standard or enhanced feedback, they addressed 1707 of 2367 contextual factors (72%), a significant difference (odds ratio, 1.3; 95%CI, 1.1-1.6; P = .01). In a mixed-effects logistic regression model, contextualized care planning was associated with a greater likelihood of improved outcomes (adjusted odds ratio, 2.5; 95%CI, 1.5-4.1; P < .001). In a budget analysis, estimated savings from avoided hospitalizations were $25.2 million (95%CI, $23.9-$26.6 million), at a cost of $337,242 for the intervention.

Although effective, one limitation of audit and feedback is the substantial delay between data collection and corrective information. Weeks pass between the audio recording of an encounter and the return of coded data on performance to participating providers. It is not a mechanism for catching contextual errors in progress and averting them. In addition, auditing is resource intensive and may not be scalable nor lend to continuous improvement. Those limitation led us to propose developing a CDS intervention that would guide clinician to contextualize care at the point of care.

The CDS Innovation

Clinical Decision Support (CDS) provides a set of strategies for both individualizing and timing heightened awareness of patient specific information to inform decision making. CDS integrates patient specific data with a knowledge base and interprets the resulting data with clinical rules and guidelines to provide support for clinicians at various points in the care process.[22] CDS can interact with clinicians in a

variety of ways, from interactive alerts to passive visualization that guides decisions without interrupting clinicians. It can provide support in real time, or asynchronously as a message that can come at a more convenient time for non-urgent information.[23]

To date the knowledge base in CDS systems has been primarily biomedical information, such as laboratory data, pharmaceuticals, diagnosis, patient allergies, age, and gender. We proposed incorporating contextual information into the CDS knowledge base to allow CDS interventions that help clinicians pick up on contextual red flags and prevent contextual errors. The approach embraced the "Five Rights" framework already widely adopted in CDS design.[24] CDS interventions must provide the *right information*, to the *right people*, through the *right channels,* in the right *intervention formats,* at the *right points* in workflow.

In designing the innovation, we asked "how might CDS be deployed to reduced contextual error?" Consider, again, the man with worsening asthma symptoms described above. When we trained and deployed a USP to portray the scenario at 50 different practice sites, physicians failed to probe context 20% of the time, despite the red flags. Among those who did probe and successfully learned that the patient was having trouble affording his medication, fewer addressed the contextual factor in the care plan. Overall, only 30% of visits concluded with a contextualized care plan, such as a prescription for a less costly medication.

We considered how CDS could serve as a corrective: This time, when the patient scheduled his appointment, he would be invited to complete a questionnaire embedded in the patient portal to complete, accessible on a smart phone or any web browser. The questionnaire would ask him to respond "yes" or "no" to seven questions. Any affirmative response would constitute a contextual red flag. In the asthma example, the patient would respond "yes" to the following question: "Are you having any difficulty taking medications the way you have been told to take them?" His response would then take him to a checklist of contextual factors related to medication non-adherence. He would select: "My medication is too expensive." When he then saw the physician, he would again report his worsening asthma and comment that "it's been tough since I lost my job." The CDS would then serve as corrective, were the physician not to probe. For instance, if the physician, pre-occupied with data entry, missed the comment and attempted to enter an order to add additional medications, an alert would appear indicating that the patient had reported that he is not taking his medication as directed because he could not afford it. The alert would ask the physician if they still want to order the medication.

In addition to eliciting contextual information directly from the patient via a questionnaire, contextualized CDS could also cull data from the patient's medical record utilizing a set of rules designed to identify common contextual red flags, such as frequent missed appointments, loss of control of a chronic condition, or lapses in medication refills, all indicators that a patient may be struggling with contextual factors complicating their care.

Hence, we envisioned a contextualized CDS that elicits information about patient context both directly from the patient and from their medical record and utilizes it to drive CDS at the point of care. Alerting a physician at a point where they are initiating an action that is likely inappropriate and that indicates they have overlooked essential information in order to point them in a different direction when and where they have the tools to alter their plan, exemplifies the Five Rights model.

In the following section, we outline the plan we employed for incorporating and assessing patient contextual information (contextual red flags and contextual factors) into CDS, and assessing its impact on contextual error rates, health care outcomes and the misuse, overuse, and underuse of medical services, drawing on methods of measurement developed, validated and extensively employed in our prior research.

Methods

We describe the study design, organized around three components to test the hypotheses specified in the aims section above, through the execution of an RCT and a USP study, conducted in parallel with the same physician subject participants:

1. Intervention: A description of the contextualized CDS design
2. RCT: The randomization of real patients to receive usual care versus care augmented by contextualized CDS, for the purpose of assessing contextualized CDS on both contextual error rates and patient health care outcomes;
3. Unannounced Standardized Patients (USPs) cost study: USPs receive usual care versus care augmented by contextualized CDS, for the purpose of assessing contextualized CDS on costs of misuse, overuse, and underuse of medical services.

Intervention:

The intervention design, which followed the Five Rights of CDS framework, consisted of two elements: A "contextual care box" (CCB) that appeared in the clinician's note at the start of the visit notifying them of contextual factors based on patient questionnaire responses, and selected passive and active interruptive alerts intended to direct them towards a contextualized care plan and away from a contextual error.[23] Both were populated by two pre-visit sources of data: the questionnaire and a set of algorithms that extract contextual factors from the patient's medical record.

The questionnaire included seven questions designed in a prior study to elicit a broad range of contextual red flags pertaining to medication, appointment, laboratory, and test adherence; declining recommended treatments, tests, and procedures; repeated visits to the emergency department (ED); difficulty accessing equipment or supplies; challenges carrying out activities important to staying healthy, and other challenges related to self-managing care.[25] An affirmative response to any item prompted the respondent to select one or more contextual factors if present.

In addition, the following contextual red flags were extracted from the patient's medical record based on rules programmed into the EHR and activated for intervention patients: missed appointments, missed tests and procedures, multiple ED visits, loss of control of either diabetes or hypertension while on medication, and self-pay status. Each was determined by a set of parameters, such as number of missed appointments in a specific time frame (examples in table 1).

Table 1: Prospectively determined outcomes based on the presenting contextual red flag

Red Flag (sources)	**Criteria**	**Good Outcome (Red flag documented to improve/resolve)**	**Bad Outcome (Red flag documented to worsen/persist)**	**No change (Red flag documented at same level or not documented as changed)**
1 - Uncontrolled chronic condition for which the patient is being treated (EHR, audio)	Hemoglobin A1c (Hb A1c) has increased since prior measurement by > 1% point; blood pressure has increased since prior measurement. Systolic Blood Pressure (SBP) and/or diastolic BP (DBP) by > 10 mmHg	Any decrease in HbA1c level; Any decrease in SBP or DBP	Any increase in HbA1c level; Any increase in SBP or DBP	No change in HbA1c; no change in SBP or DBP

Red Flag (sources)	Criteria	Good Outcome (Red flag documented to improve/resolve)	Bad Outcome (Red flag documented to worsen/persist)	No change (Red flag documented at same level or not documented as changed)
2 - Appointment non-adherence: clinics, labs, imaging, procedures (EHR, patient questionnaire, audio)	Missed ≥ 2 clinical encounters in past 4 months; Missed ≥ 1 laboratory and/or scheduled study in past 4 months	Patient misses fewer appointments during the next four months; Patient completes scheduled laboratory tests and/or scheduled studies	Patient misses more appointments in the next four months; Patient completes fewer scheduled laboratory tests and/or scheduled studies	Patient misses same number of appointments in next four months; patient miss same proportion of scheduled tests or studies
3 - ED visits (EHR, patient questionnaire, audio)	≥ 2 in past 4 months	Patient has fewer ED visits	Patient has more ED visits	Patient has same number of ED visits
4 - Medication non-adherence (patient questionnaire, audio)	Answers "yes" to "Are you having any difficulty taking medications the way you have been told to take them?" and/or mentions it during visit.	Patient takes medications as prescribed	Patient does not take medications as prescribed	Medical adherence not documented
5 – Missed preventive care (patient questionnaire, audio)	Answers "yes" to "In the past six months, have you declined any treatments, tests, or procedures that your provider recommended; like vaccines, blood tests, a colonoscopy?" and/or mentions it during visit.	Patient receives recommended treatments, tests, or procedures	Patient does not receive recommended treatments, tests, or procedures	Patient receipt of treatments, tests, procedures not documented
6 - Weight gain/loss (audio)	At least 10 lbs gained or lost since last appointment	If overweight, weight is lower; if underweight, weight is higher.	If overweight, weight is higher; if underweight, weight is lower.	Patient weight stays the same
7 - Unaware of diagnosis/results (audio)	Patient mentions that they are unaware of diagnosis/test results that should have been communicated to patient	Patient is aware of diagnosis/results	Patient is still unaware of diagnosis/results	Patient awareness or unawareness not documented

Red Flag (sources)	Criteria	Good Outcome (Red flag documented to improve/resolve)	Bad Outcome (Red flag documented to worsen/persist)	No change (Red flag documented at same level or not documented as changed)
8 – Difficulty with equipment (patient questionnaire, audio)	Unable to use/decline to use medical equipment/using someone else's equipment	Using own equipment	Not using (own) equipment	Patient use of equipment not documented
9 – General statements by patients that are concerning such as "I'm not eating" (patient questionnaire, audio)	Individualized	Individualized and prospectively determined	Individualized and prospectively determined	Individualized, prospectively determined outcome not documented

For those randomized to the intervention, contextual red flags and factors appeared in the CCB. Interruptive alerts were used either to initiate non-medical interventions that were likely to be beneficial or to re-direct an apparently misguided plan of care.[27] The CDS system did not operate during visits by patients in the control group.

RCT (real patients)

Over 27-months, primary care attending clinicians were recruited at two academic health centers in Chicago utilizing two different EHRs, developed by Cerner (site 1) and Epic (site 2). Clinicians were informed that the purpose of the project was to assess whether enhanced clinical decision support that provides information about patient contextual factors could improve clinical decision making and healthcare outcomes. They were informed that if they participated, a small number of patients would audio record their visits.

All English-speaking adult patients of participating clinicians who could be contacted in advance of their appointments were eligible to participate. They were informed that if they participated, they would complete a questionnaire about challenges that might impact their care. They could access it directly in the patient portal or complete it with assistance from a research assistant. When they arrived for their appointment, they received a digital audio recorder to carry into their visit. They were told that it was preferable to conceal the audio, but that they could reveal it if they felt more comfortable doing so. They were informed that a member of the research team would access their medical record first to note contextual red flags and factors self-identified by patients on the questionnaire, and then several months later to see if key healthcare indicators noted at the visit had improved. Finally, they were told that their doctor might or might not receive the information they provided, based on random assignment. They received a $20 recruitment incentive. The trial protocol was approved by the Institutional Review Boards at both sites and patients provided written informed consent.

End points

The primary study end point was proportions of improved/resolved (vs. worsened or no change) contextual red flags 6 months after the visit (Table 1). Secondary end points were the proportion of contextual red flags probed in the visit and the proportion of contextual factors identified during the visit that were addressed in the visit care plan i.e., the contextualization of care rate. The nine categories

of contextual red flags, along with sub-categories, were developed in previous research.[28,29] Exploratory outcomes were the effects of the EHR alerts on probing of contextual red flags that patients did not themselves bring up during the visit and the effect of the alerts on the red flags 6 months later.

Data were recorded in REDCap. Following each encounter, research assistants, blind to the assignment of each patient to intervention or control, accessed the EHR and listened to the audio recording to identify contextual red flags (and determine whether each was probed by the clinician) and contextual factors (and determine whether each was incorporated into the care plan by the clinician). Red flags and contextual factors were noted by the coders from the patient questionnaire, EHR, or visit audio; factors noted on audio emerged either as a result of clinician probing or spontaneous disclosure by the patient. For each red flag, coders prospectively defined the conditions under which the red flag would be considered improved or worsened in the future.

Six months after each encounter, research assistants, again blind to the assignment of patients to intervention or control, applied the predefined conditions to code each of the patients' red flags from a chart review as improved, worsened, unchanged or undocumented at follow-up.

Sample Size

Based on prior studies[4,30], we assumed that contextual red flags with associated factors would be present in 50% of recorded visits, and that clinicians, unaided, would probe 50% of contextual red flags and contextualize care plans in 50% of visits with contextual factors. Based on the number of contextual red flags and factors we expected, to obtain 80% power to detect an absolute increase in contextualization rates when factors were present from 50% to 75%, we determined that we would need recordings from 192 intervention and 288 control patients (which also provides at least 80% power to detect the expected increase in probing). We chose a 2:3 intervention:control ratio as there would be fewer opportunities to contextualize care plans in the control group if fewer red flags were probed.

Randomization and Blinding

Patients were randomized 2:3 to either contextualized CDS or usual care, utilizing computer-generated randomization. Randomized assignments were available to research personnel responsible for activating the CDS tools immediately prior to the encounter, but those involved in coding visits and tracking outcomes were blinded. Although patients were blinded as to whether they were in the intervention or control group, it was not possible to blind clinicians because the intervention is a set of novel CDS tools that are not available in usual care.

Statistical Methods

We applied a logistic mixed effects modeling approach to examine the impact of the study intervention. We fitted two models to the 6-month changes in the red flags. One model examined whether the red flag had improved (aka "resolved") (vs. worsened or unchanged; primary end point); the other examined whether the red flag had worsened (vs. improved or unchanged). Each included fixed effects of study arm, study site, and whether the red flag's associated factor had been incorporated into the care plan, and random effects of visit and clinician. As a sensitivity analysis, we also fitted a mixed ordered logit model to the outcomes categorized as worsened, unchanged/mixed, or improved.

For each red flag noted by coders, we fitted a model to whether the red flag was probed by the clinician (secondary end point). In addition to a fixed effect of study arm (intervention vs. control), we included fixed effect covariates for study site, whether a relevant factor had been selected on the pre-visit questionnaire, and whether the audio recorder was concealed or disclosed, and random effects of visit (patient) and clinician to adjust for clustering of red flags within visits and visits within clinicians.

Similarly, for each contextual factor noted by coders, we fitted a model to whether the factor was incorporated in the care plan by the clinician (secondary end point). In addition to a fixed effect of study arm, we included fixed effects of study site, whether the factor was identified by probing of a red flag, selected in the patient questionnaire, or revealed spontaneously by the patient, and whether the audio recorder was concealed or visible, and random effects of visit and clinician.

Our exploratory analyses of the impact of CDS CCB/alerts were limited to intervention visits with red flags that could be presented in the CCB (that is, red flags available prior to the visit itself), and examined whether the presence of a CCB/alert for a red flag affected the likelihood of probing, incorporating a contextual factor in a plan if one was present, or having an improved or worsened red flag at four months, adjusting for site and random effects of visit and clinician. We also examined this question using all visits and including covariates for intervention vs. control group and the interaction of intervention vs. control group and presence of CCB/alert (which could only happen in the intervention groups).

All analyses were conducted on an intention-to-treat basis using R 3.6 (R Core Team, Vienna, Austria).

Unannounced Standardized Patients

Since USPs are, by definition, constructed and standardized, they enable an experimental assessment of how a group of clinicians differ when seeing the same patient against an ideal standard. In this study we introduced just one variable – contextualized CDS – enabling us to isolate its effect. We constructed 4 USP scripts with embedded contextual red flags and factors, drawn from our library of such cases were selected. The actors' training and deployment was managed by the UIC Simulation and Integrative Learning Institute (SAIL) Center, which has extensive prior USP experience.[7,26] The scripts were modified and customized to assess the efficacy of the selected CDS innovations such that failure of CDS to prevent inattention to contextual red flags or factors in USP cases could result in misuse, overuse or underuse of medical services. Each script was to be portrayed at 10 control visits without CDS support and 10 intervention visits with CDS support, divided across the two sites, for a total of 80 USP visits.

Analytic Framework

Utilizing our previously published methods,[5] we adopted the economic perspective of the patient and their third party payer, if any, with a time horizon of the expected consequences of care during the 30 days following the consultation. We considered only the direct consequences of care associated with diagnosis or misdiagnosis. We did not consider downstream costs beyond the initial recommendations from the consultation, and we did not consider societal costs not incurred by the patient or payer, such as lost productivity. We included only resources related to the immediate diagnostic and therapeutic management at the index visit. Resources are direct medical costs in the case of unnecessary treatment and foregone direct medical costs in the case of under treatment.

Sample size (USPs): In our past work with USPs, physicians made contextual errors approximately 80% of the time.[7] Assuming that the contextualized CDS enhances physician attention to red flags and leads them to probe substantially more often (e.g. increasing probe rate from 50% to 75%) and attend to identified information (e.g. increasing plan rate from 50% to 75%), we expected overall contextual errors to occur no more than 45% of the time, and that 28 control and 28 intervention USP visits would provide 80% power to detect such a difference and test hypothesis 1.

In our past work, we found an overall median cost of error of $194 when cases presented with contextual red flags, based on a median cost of $231 when contextual errors occurred and a median cost of $0 when contextual errors did not occur.[5] Based on bootstrapped simulation from our cost data in that study, 40 control and 40 intervention USP visits provide 83% power to detect the expected cost reduction (a median of $156) due to reduced contextual errors using a Wilcoxon rank-sum test with a significance level of $p<.05$. Accordingly, we planned to conduct 40 control and 40 intervention USP visits

to provide sufficient power to test both study hypotheses. As the study would comprise 4 USP visits (2 control, 2 intervention) per physician, we would recruit 20 physicians for this portion of the study.

Results

Real Patients (RCT)

Study Participants and Recruitment

From September 2018 to March 2021, 452 adult patients (65% female) with upcoming primary care appointments to 39 physicians completed the pre-visit questionnaire (95% with assistance) and carried a concealed audio recorder into their visit. As shown in Figure 1, 275 were randomly allocated to the control group and 177 to the intervention.

Figure 1: Flow of Patients Through the Clinical Decision Support to Prevent Contextual Errors Trial

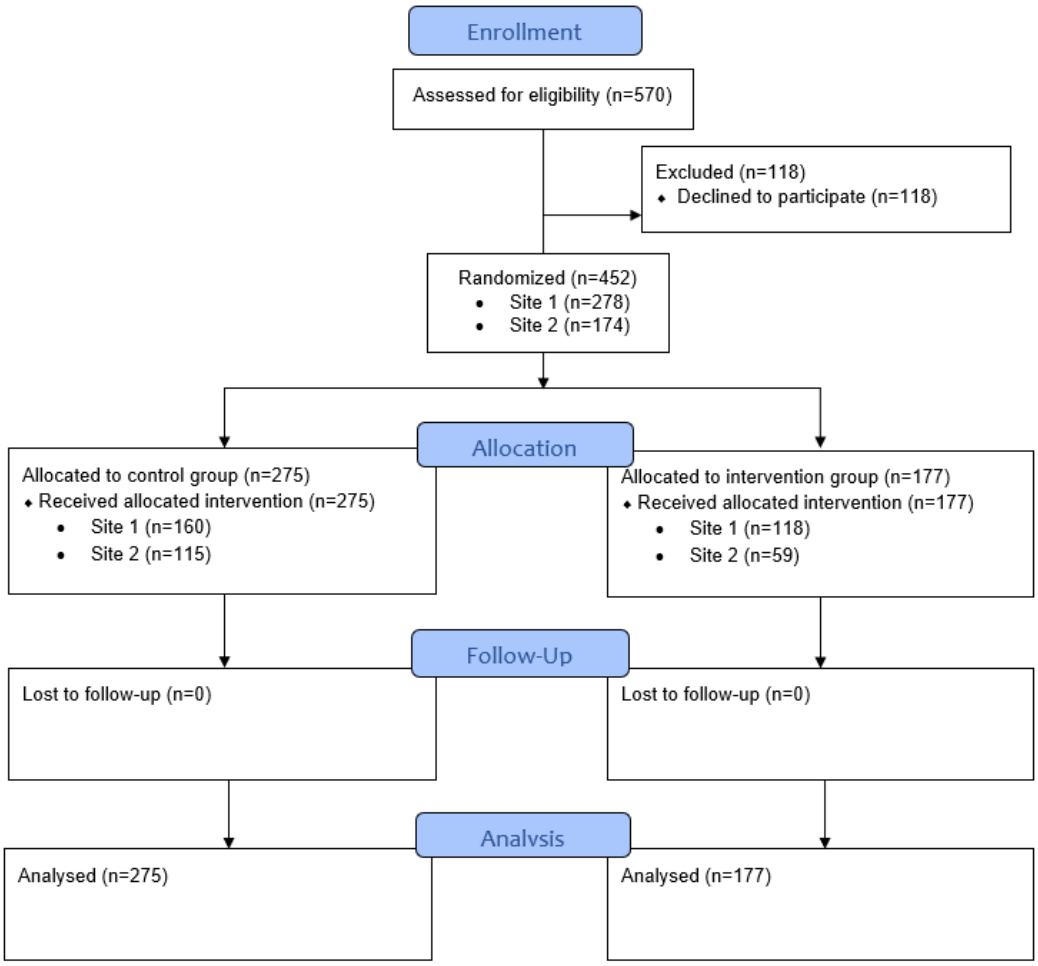


Table 2 presents patient characteristics by study group; patients did not differ between groups in site, gender, numbers of red flags available before or during the visit, or whether they chose to reveal the audio recorder or keep it concealed.

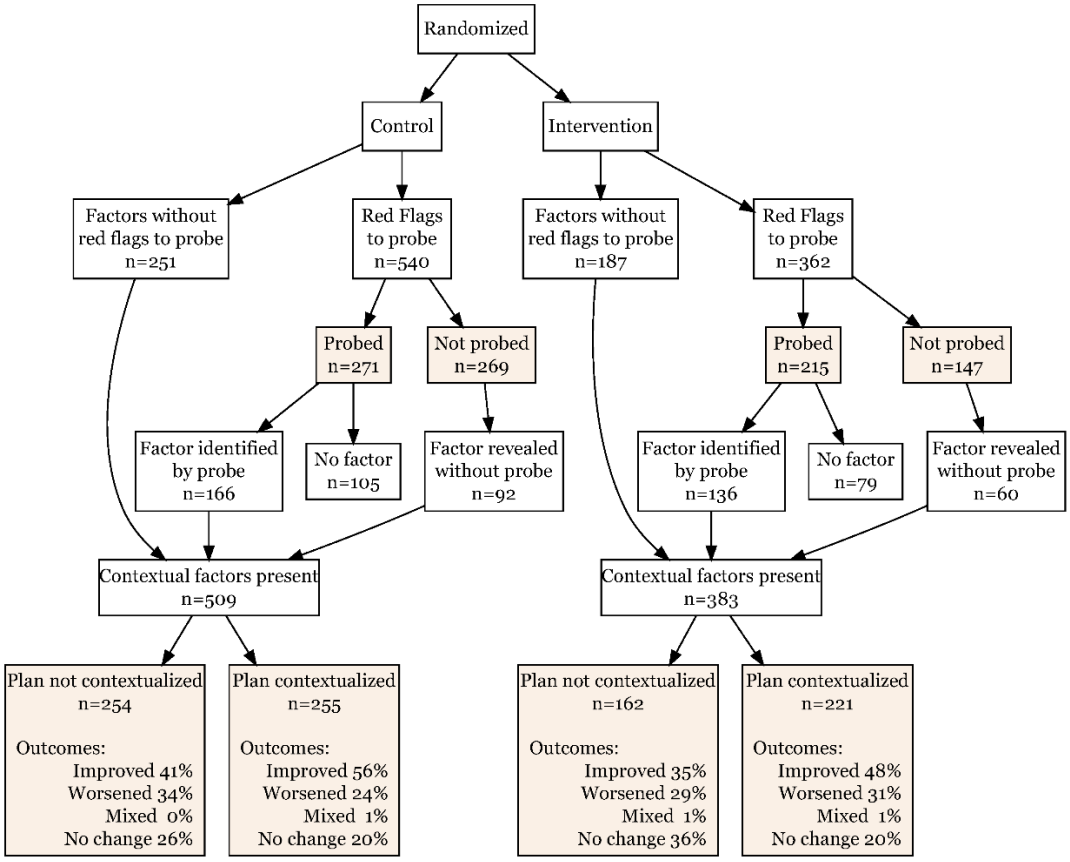
Table 2: Participant Characteristics

	Control (N=275)	**Intervention (N=177)**	**Total (N=452)**
Site			
1	160 (58.2%)	118 (66.7%)	278 (61.5%)
2	115 (41.8%)	59 (33.3%)	174 (38.5%)
Gender			
Female	182 (66.2%)	111 (62.7%)	293 (64.8%)
Male	93 (33.8%)	66 (37.3%)	159 (35.2%)
Number of red flags per visit, mean (sd)	3.9 (2.4)	4.0 (2.3)	3.9 (2.4)
Only reported contextual red flag on patient questionnaire	78 (28.4%)	57 (32.3%)	135 (29.9%)
Only had a contextual red flag that could be extracted from the EHR	33 (12.0%)	18 (10.2%)	51 (11.3%)
Contextual red flag both reported on questionnaire and extracted from EHR	140 (50.9%)	87 (49.2%)	227 (50.2%)
Revealed audio recorder			
Yes	28 (10.2%)	19 (10.7%)	47 (10.4%)
No	247 (89.8)	158 (89.3%)	405 (89.6%)

Outcomes

Figure 2 provides raw data on the disposition of all contextual red flags and contextual factors in both trial groups, including contextual factors disclosed by patients without a clinician probe ("Factors without red flags to probe").

Figure 2: Raw study outcomes by arm. "No change" includes cases where evidence of change was not available.



These disclosures occurred either because the patient reported the contextual factors on the pre-visit questionnaire or revealed them spontaneously during the encounter. In addition, red flag outcomes are shown. Red flags could be improved, worsened, unchanged, or mixed. A red flag outcome could be mixed when there was more than one element -- for example, when a good outcome would require that a patient had engaged in both diet modification and exercise but at follow-up had adopted one but not the other.

Table 3 summarizes the main outcomes of the study: Patient outcomes, clinician probing, and clinician contextual care planning, by study arm. For the primary end point, study arm did not have additional significant impact on outcomes beyond its impact on plans (p=0.91, p=0.25 for improvement and worsening respectively; table 3, row 1). However, across arms, contextualized plans were associated with a significantly higher likelihood of an improved red flag outcome (AOR=2.14, p=0.001), and a significantly lower likelihood of a worsened red flag outcome (AOR=0.66, p=0.04); The ordered logit model produced the same conclusions (proportional odds AOR associated with contextualized plan = 1.77, 95% CI 1.24 – 2.52, p=0.002; data not shown).

Table 3: Patient outcomes, clinician probing, and clinician contextual care planning, by study arm*

	Control rate, unadjusted	Intervention rate, unadjusted	Effect size (odds ratio), adjusted*	95% CI for effect size	P
Outcome: Improvement or resolution of red flags at 4-6 months, adjusted for whether provider incorporated contextual factor	247 / 509 (48%)	163 / 383 (43%)	0.97	0.57 – 1.64	0.91
Probing: Provider probes contextual red flags	271 / 540 (50%)	215 / 362 (59%)	2.09	1.13 – 3.86	0.02
Planning: Provider incorporates contextual factors into care plan	255 / 509 (50%)	221 / 383 (58%)	2.67	1.32 – 5.41	0.002

*All models adjusted for random effects of visit (patient) and provider and fixed effects of study site. Probing model also adjusted for whether patient indicated a contextual factor on pre-visit questionnaire and whether patient chose to make recorder visible to provider. Planning model also adjusted for whether patient indicated a contextual factor on pre-visit questionnaire, whether patient chose to make recorder visible to provider, whether provider identified the contextual factor by probing a red flag, whether patient spontaneously revealed the contextual factor, and the interaction between whether provider probed and whether the recorder was visible.

Clinicians in the intervention arm were more likely to probe red flags than clinicians in the control arm (adjusted odds ratio (AOR)=2.09, p=0.02; Table 3, row 2). In addition, red flags selected on the pre-visit questionnaire were less likely to be probed than those expressed orally during the visit (AOR=0.15, p<0.001). There was no effect of whether the recorder was visible or concealed (p=0.33). Table 4 shows the proportion of contextual red flags probed by a clinician, the proportion of contextual red flags for which a contextual factor was noted by the 4C coders, and the proportion of those contextual factors addressed in the care plan i.e., the contextualization of care rate by types of flags in the control and intervention groups.

Table 4: Clinician probing, identification of contextual factors, and contextual care plan for each category of contextual red flag in control and treatment

	Total red flags				Probed (out of total)				Factors with red flags (out of total)				Plans (out of factors with red flags)			
	Control		Intervention		Control		Intervention		Control		Intervention		Control		Intervention	
Red Flag	N	%	N	%	N	%	N	%	N	%	N	%	N	%	N	%
1 – Uncontrolled chronic condition	80	15%	37	10%	49	61%	31	84%	15	19%	12	32%	9	60%	11	92%
2 – Appointment non-adherence (clinics, labs, imaging, procedures)	81	15%	59	16%	33	41%	23	39%	42	52%	26	44%	28	67%	20	77%
3 – Under/over Resource Utilization	23	4%	14	4%	6	26%	4	29%	12	52%	9	64%	7	58%	8	89%
4 - Medication non-adherence	166	31%	106	29%	74	45%	62	58%	52	31%	52	49%	39	75%	43	83%
5 – Plan of Care non-adherence	100	19%	68	19%	43	43%	38	56%	80	80%	51	75%	54	68%	38	75%
6 – Significant weight gain/loss	7	1%	4	1%	6	86%	3	75%	3	43%	2	50%	2	67%	2	100%
7 - Lack of knowledge of health or health care status	3	1%	6	2%	3	100%	5	83%	3	100%	3	50%	3	100%	2	67%
8 - Medical equipment/supplies non-adherence	19	4%	19	5%	13	68%	12	63%	15	79%	15	79%	11	73%	10	67%
9 - Other	61	11%	49	14%	44	72%	37	76%	36	59%	26	53%	27	75%	22	85%
Total	540	100%	362	100%	271	50%	215	59%	258	48%	196	54%	180	70%	156	80%

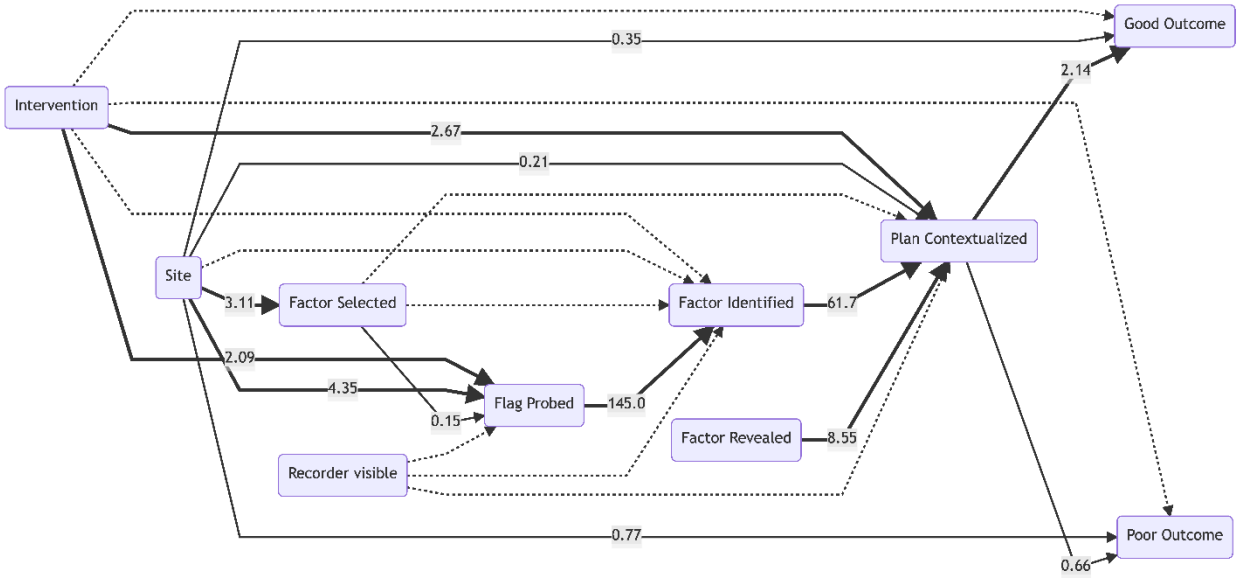
Overall, 59% of contextual red flags were probed in the intervention group, compared with 50% in the control. Contextual factors were identified in 54% compared with 48% of them, respectively. Overall, 80% of contextual factors with presenting red flags (58% of all contextual factors) were elicited in the intervention compared to 70% (50%) in the control.

Factors in the intervention arm had a higher likelihood of incorporation into the care plan (AOR=2.67 p=.0002; table 3, row 3). As has been found in past studies, factors identified through

probing were additionally much more likely to be incorporated into care plans than those that were not (AOR=61.7, p<.001).[31] Factors spontaneously revealed by patients were also more likely to be incorporated (AOR=8.6, p<.001), but factors selected by patients on the pre-visit questionnaire were not (p=0.45). There was no effect of whether the recorder was visible or concealed (p=0.79).

Figure 3 summarizes the results of the models as a path analysis and illustrates the direct and indirect effects of the intervention on contextual probing, identification of contextual factors, contextualization of care plans, and patient outcomes.

Figure 3: Path analytic summary of multivariate models. Each node with arrows pointing to it represents a separate mixed effects model in which the node's outcome is regressed on the nodes pointing to it in the diagram. Illustrates the direct and indirect effects of the intervention on contextual probing, identification of contextual factors, contextualization of care plans, and patient outcomes.



Impact of EHR Alerts

When the EHR populated the CCB or activated alerts for red flags prior to the start of the visit (i.e., from the patient questionnaire or EHR triggers), the likelihood of the clinician probing the red flag was significantly increased (AOR=3.6, 95% CI 1.2-11.2, p=0.02), and when a factor was present, the likelihood of incorporating it into the care plan also increased (AOR=11.3, 95% CI 2.3 – 55.9, p=0.003). In addition, when the red flag was the subject of an alert, it was less likely to have worsened four months later (AOR=0.19, 95% CI 0.05-0.73, p=0.02) but not more likely to have been improved (AOR=0.75, p=0.55), beyond the effect of whether the care plan had been contextualized. Supplement E tables 2 and 3 present the regression coefficients for models exploring the effect of EMR box/alerts in intervention arm visits and all visits, respectively.

<u>USP Study</u>

Because of the onset of Covid-19, the USP study was suspended in March 2020. At that point we had completed 32 visits at Site 1 and 9 visits at Site 2, for a total of 41 out of total of 80 planned visits. We were unable to resume data collection because the benefit:risk of sending actors disguised as real patients to engage in face-to-face encounters with physicians was not again favorable throughout the remaining duration of the study.

The intervention did not increase probing rates in the USP visits. Regardless, across study arms and cases, probing was associated with a greater likelihood of contextualization of plans. The intervention did increase contextualization of plans, however, over and above the effect of probing. Intervention encounters had a significantly higher proportion of contextualized items per case than control encounters, adjusted for site, case, and repeated measures within physicians (odds ratio 2.4, 95% CI 1.3 - 4.7, p=0.008).

Cost Analysis:

The primary aim of the cost analysis (Aim 3 of study) was to test the hypothesis that Contextualized CDS *reduces avoidable health care costs* by reducing misuse, overuse, and underuse of inappropriate or unnecessary medical services. See Analytic framework, "economic perspective" above for methodology. The analysis was conducted by sorting each action by a clinician (generally an order or message to a clinical care team member), across each of the four USP cases into one of three categories: overuse, underuse, or appropriate use of medical services (there were no observed instances of misuse). Appropriate use was defined at the time of the USP design and operationalized in the CDS guidance. Specifically, a clinician's acceptance of CDS recommendations always constituted appropriate care. For the purpose of the analysis, appropriate care was set at zero cost, meaning the cost was neither higher nor lower than its economic value from the perspective of the patient and their third-party payer. Both overuse and underuse were valued as the cost of the resource inappropriately ordered or inappropriately not ordered, including medications, tests, procedures, consults, and medical equipment. To standardize the analysis, all medications were based on 2017 Medicare expenditure for a 30-day refill. For instance, ordering a medication that was either not necessary or contraindicated (overuse or misuse) was assigned a cost equivalent to a 30-day refill, as was neglecting to order a medication that was indicated and recommended by the contextualized CDS.

Figure 4: Adjusted costs of misuse, overuse, and underuse of medical services in care of USPs in by clinicians with contextualized CDS vs usual care.

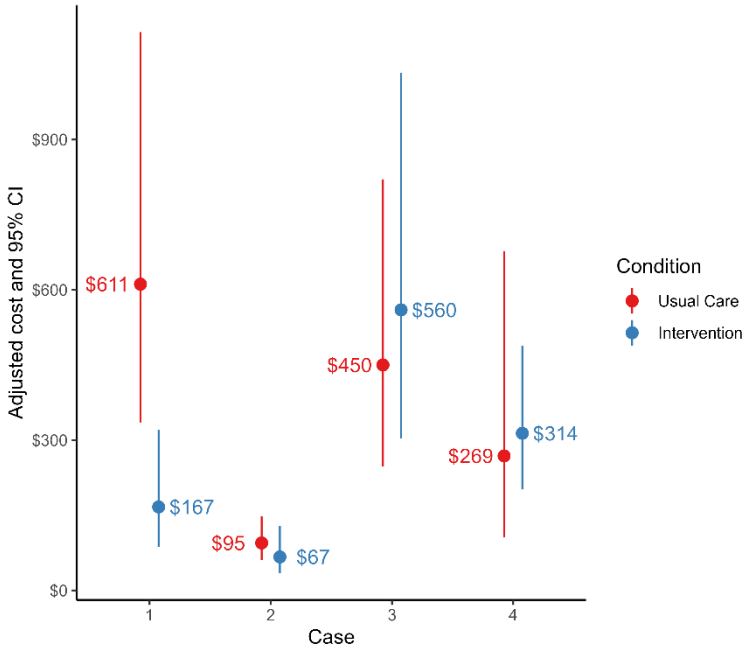


Figure 4 shows adjusted costs based on a log-linked gamma regression, adjusting for repeated visits to the same physician and for site. As shown in the figure, the intervention resulted in lower improper costs in case 1 (p=.0025) and no significant difference in other cases. The effect in case 1 was driven by two prescriptions of one medication that should have been discontinued because the patient had indicated they could not afford it and replaced with an affordable alternative, as recommended by the CDS.

Overall, whether plan elements were contextualized did not significantly affect the cost of care. However, the cost of care was lower in the intervention than control encounter plans, adjusting for site,

case, contextualized elements, and repeated measures within physicians. The primary driver of this impact was that "usual care" encounters at site 2 were particularly costly (estimated mean cost $860 [$347 - $2131] for usual care vs. $350 [$153 - $800] for intervention encounters (p=<.001). All encounters at site 1 were similar in cost to intervention encounters at site 2 (site 1 usual case $333 [$202 - $551], site 1 intervention $327 [$197 - $542], p=.88).

Thus, based on the partial USP data we could collect prior to the pandemic, we could detect greater recorded contextualization in the intervention vs. control encounters, but not a consistent reduction in inappropriate costs.

Discussion

In this randomized clinical trial, CDS tools increased the likelihood that a clinician would address relevant patient life context in their care plan. As in prior studies, contextualization of care was associated with improved outcomes. However, despite this association, the CDS intervention did not have an additional significant impact on outcomes. This counter-intuitive finding appears to have occurred because contextualized care plans (specifically, those for which there was an outcome that could be prospectively defined) in the control group were more likely to positively effect outcomes than contextualized care plans in the intervention (specifically those for which there was an outcome that could be prospectively defined).

In the USP study, CDS tools were also effective at facilitating contextualization of care. Since the CDS tools had this effect without increasing probing rates, it's likely that they worked by alerting clinicians to contextual factors that they otherwise would have missed. This USP study did not demonstrate, however, that contextualized CDS significantly reduced unnecessary costs of care. However, data collection was cut short by the pandemic resulting in an underpowered analysis based on just 41 USP encounters.

Although it has been previously established that contextual errors are common, adversely affect patient outcomes, and that clinicians can learn to make fewer of them with feedback,[32] this is the first study to demonstrate that CDS tools built into the EHR can decrease contextual errors. CDS has the advantage of being readily scalable. Just as CDS can guide biomedically focused decision making by drawing on data from evidence-based guidelines and other forms of research evidence, it can also guide contextually informed decision making by drawing on data specific to the life circumstances and behaviors of individual patients.[33]

Limitations

It was not possible to blind clinicians to study arm since they can tell when they are receiving contextualized CDS.[34] It's possible that they try harder in the intervention, independent of the CDS. We attempted to mitigate any effect by explaining to clinicians that there would be no analysis of how they individually perform. Furthermore, the lack of an effect of the visible recorder on clinician probing suggests that they weren't motivated to change their behavior just because they knew they were in a study. Finally, the study was underpowered to ascertain whether contextualized CDS can reduce the costs of inappropriate care.

Conclusions:

Information about patients' life circumstances and behaviors, i.e., patient life context, is essential to identify and address in care planning to achieve desired outcomes. However, it is often overlooked. The findings of this study support the use of clinical decision support tools that draw on

data elicited directly from the patient prior to the visit and from the EHR to facilitate decision-making that improves contextualization of care. Although contextualized CDS in this study increased contextualization of care, and contextualization of care across arms was associated with improved outcomes, the intervention did not have an additive effect. Further work is needed to ascertain if contextualized CDS improved health care outcomes. Also, because of pandemic effects on the protocol this study was underpowered to ascertain whether contextualized CDS reduces the cost of inappropriate care.

Publications

A manuscript describing the findings of the RCT has been revised and resubmitted to JAMA Network Open Access. Trial results have also been submitted to clinicaltrials.gov to update the registry records of the trial.

References

1. Weiner SJ. Contextualizing medical decisions to individualize care: lessons from the qualitative sciences. *J Gen Intern Med.* 2004;19(3):281-285.
2. Weiner SJ. From research evidence to context: the challenge of individualizing care. *ACP J Club.* 2004;141(3):A11-12.
3. Weiner SJ. Contextual Error. In: Kattan M, ed. *Encyclopedia of Medical Decision Making*: SAGE; 2009:198-202.
4. Weiner SJ, Schwartz A, Sharma G, et al. Patient-centered decision making and health care outcomes: an observational study. *Ann Intern Med.* 2013;158(8):573-579.
5. Schwartz A, Weiner SJ, Weaver F, et al. Uncharted territory: measuring costs of diagnostic errors outside the medical record. *BMJ Qual Saf.* 2012.
6. Weiner SJ, Kelly B, Ashley N, et al. Content Coding for Contextualization of Care: Evaluating Physician Performance at Patient-Centered Decision Making. *Med Decis Making.* 2013.
7. Weiner SJ, Schwartz A, Weaver F, et al. Contextual errors and failures in individualizing patient care: a multicenter study. *Ann Intern Med.* 2010;153(2):69-75.
8. Weiner SJ SA. Contextual Errors in Medical Decision Making: Overlooked and Understudied. *Acad Med.* 2016;91(5):657-662.
9. Weiner SJ, Schwartz A. *Listening for what matters : avoiding contextual errors in health care.* Oxford ; New York: Oxford University Press; 2016.
10. Weiner SJ, Schwartz A, Yudkowsky R, et al. Evaluating physician performance at individualizing care: a pilot study tracking contextual errors in medical decision making. *Med Decis Making.* 2007;27(6):726-734.
11. Galanter WL, Bryson ML, Falck S, et al. Indication alerts intercept drug name confusion errors during computerized entry of medication orders. *PLoS One.* 2014;9(7):e101977.
12. Galanter W, Falck S, Burns M, Laragh M, Lambert BL. Indication-based prescribing prevents wrong-patient medication errors in computerized provider order entry (CPOE). *Journal of the American Medical Informatics Association : JAMIA.* 2013;20(3):477-481.
13. Galanter WL, Didomenico RJ, Polikaitis A. A trial of automated decision support alerts for contraindicated medications using computerized physician order entry. *Journal of the American Medical Informatics Association : JAMIA.* 2005;12(3):269-274.
14. Galanter WL, Thambi M, Rosencranz H, et al. Effects of clinical decision support on venous thromboembolism risk assessment, prophylaxis, and prevention at a university teaching hospital. *American journal of health-system pharmacy : AJHP : official journal of the American Society of Health-System Pharmacists.* 2010;67(15):1265-1273.

15. Galanter WL, Polikaitis A, DiDomenico RJ. A trial of automated safety alerts for inpatient digoxin use with computerized physician order entry. *Journal of the American Medical Informatics Association : JAMIA.* 2004;11(4):270-277.
16. Kucher N, Koo S, Quiroz R, et al. Electronic alerts to prevent venous thromboembolism among hospitalized patients. *N Engl J Med.* 2005;352(10):969-977.
17. Schwartz A, Weiner SJ, Harris IB, Binns-Calvey A. An educational intervention for contextualizing patient care and medical students' abilities to probe for contextual issues in simulated patients. *JAMA.* 2010;304(11):1191-1197.
18. Weiner SJ SA, Sharma G, Binns-Calvey A, Ashley N, Kelly B, Weaver FM. Patient collected audio for performance assessment of the clinical encounter. *Jt Comm J Qual Patient Saf.* 2015;42(6):273-278.
19. Weiner SJ, Schwartz A, Weaver F, et al. Overlooking Contextual Information When Individualizing Care: A Source of Medical Error and Avoidable Cost. *Journal of General Internal Medicine.* 2009;24:130-130.
20. Weiner SJ, Schwartz A. Directly Observed Care: Can Unannounced Standardized Patients Address a Gap in Performance Measurement? *J Gen Intern Med.* 2014.
21. Schwartz A, Weiner SJ, Binns-Calvey A. Comparing announced with unannounced standardized patients in performance assessment. *Jt Comm J Qual Patient Saf.* 2013;39(2):83-88.
22. Beeler PE, Bates DW, Hug BL. Clinical decision support systems. *Swiss Med Wkly.* 2014;144:w14073.
23. Osheroff JA, Healthcare Information and Management Systems Society. *Improving outcomes with clinical decision support : an implementer's guide.* 2nd ed. Chicago, IL: HIMSS; 2012.
24. Sirajuddin AM, Osheroff JA, Sittig DF, Chuo J, Velasco F, Collins DA. Implementation pearls from a new guidebook on improving medication use and outcomes with clinical decision support. Effective CDS is essential for addressing healthcare performance improvement imperatives. *J Healthc Inf Manag.* 2009;23(4):38-45.
25. Weaver FM, Binns-Calvey A, Gonzalez B, et al. Alerting Doctors About Patient Life Challenges: A Randomized Control Trial of a Previsit Inventory of Contextual Factors. *MDM Policy Pract.* 2019;4(1):2381468319852334.
26. Dr. Allan L. and Mary L. Graham Clinical Performance Center. http://chicago.medicine.uic.edu/grahamcpc/.
27. Co Z, Holmgren AJ, Classen DC, et al. The tradeoffs between safety and alert fatigue: Data from a national evaluation of hospital medication-related clinical decision support. *Journal of the American Medical Informatics Association : JAMIA.* 2020;27(8):1252-1258.
28. Binns-Calvey AE, Sharma G, Ashley N, Kelly B, Weaver FM, Weiner SJ. Listening to the Patient: A Typology of Contextual Red Flags in Disease Management Encounters. *J Patient Cent Res Rev.* 2020;7(1):39-46.
29. Weiner, SJ, Ashley N, Binns-Calvey A, Kelly B, Sharma G, Schwartz A. Content Coding for Contextualization of Care. Version 12.0, Released May 27, 2021. Harvard Dataverse Network Project [on line] https://dataverse.harvard.edu/dataverse/4C.
30. Weiner SJ, Schwartz A, Sharma G, et al. Patient-collected audio for performance assessment of the clinical encounter. *Jt Comm J Qual Patient Saf.* 2015;41(6):273-278.
31. Schwartz A, Weiner SJ, Binns-Calvey A, Weaver FM. Providers contextualise care more often when they discover patient context by asking: meta-analysis of three primary data sets. *BMJ Qual Saf.* 2016;25(3):159-163.
32. Weiner SJ, Schwartz A. Contextual Errors in Medical Decision Making: Overlooked and Understudied. *Acad Med.* 2016;91(5):657-662.

33. Sutton RT, Pincock D, Baumgart DC, Sadowski DC, Fedorak RN, Kroeker KI. An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ Digit Med.* 2020;3:17.
34. McCambridge J, Witton J, Elbourne DR. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. *J Clin Epidemiol.* 2014;67(3):267-277.