# Development, Evaluation and Implementation of Chief Complaint Groupings to Activate Data Collection

## A Multi-Center Study of Clinical Decision Support for Children with Head Trauma

S. J. Deakyne[1]; L. Bajaj[2]; J. Hoffman[3]; E. Alessandrini[4]; D. W. Ballard[5]; R. Norris[6]; L. Tzimenatos[7]; M. Swietlik[8]; E. Tham[2]; R. W. Grundmeier[9]; N. Kuppermann[7]; P. S. Dayan[10] for the Pediatric Emergency Care Applied Research Network (PECARN)

[1]Children's Hospital Colorado, Department of Research Informatics, Aurora, Colorado, United States; [2]University of Colorado, Department of Pediatrics, Section of Emergency Medicine, Aurora, Colorado, United States; [3]Nationwide Children's Hospital, Columbus, Ohio, United States; [4]Children's Hospital Medical Center, Cincinnati, Ohio, United States; [5]Kaiser Permanente, San Rafael Medical Center, San Rafael, California, United States; [6]Kaiser Permanente, Sacramento Medical Center, Sacramento, California, United States; [7]University of California Davis School of Medicine, Departments of Emergency Medicine and Pediatrics, Sacramento, California, United States; [8]Children's Hospital Colorado, Department of Clinical Application Services, Aurora, Colorado, United States; [9]Children's Hospital of Philadelphia and Perelman School of Medicine, Philadelphia, Pennsylvania, United States; [10]Columbia University College of Physicians and Surgeons, Department of Pediatrics, Division of Emergency Medicine, New York, New York, United States

**Summary**
**Background:** Overuse of cranial computed tomography scans in children with blunt head trauma unnecessarily exposes them to radiation. The Pediatric Emergency Care Applied Research Network (PECARN) blunt head trauma prediction rules identify children who do not require a computed tomography scan. Electronic health record (EHR) based clinical decision support (CDS) may effectively implement these rules but must only be provided for appropriate patients in order to minimize excessive alerts.
**Objective:** To develop, implement and evaluate site-specific groupings of chief complaints (CC) that accurately identify children with head trauma, in order to activate data collection in an EHR.
**Methods:** As part of a 13 site clinical trial comparing cranial computed tomography use before and after implementation of CDS, four PECARN sites centrally developed and locally implemented CC groupings to trigger a clinical trial alert (CTA) to facilitate the completion of an emergency department head trauma data collection template. We tested and chose CC groupings to attain high sensitivity while maintaining at least moderate specificity.
**Results:** Due to variability in CCs available, identical groupings across sites were not possible. We noted substantial variability in the sensitivity and specificity of seemingly similar CC groupings between sites. The implemented CC groupings had sensitivities greater than 90% with specificities between 75–89%. During the trial, formal testing and provider feedback led to tailoring of the CC groupings at some sites.
**Conclusion:** CC groupings can be successfully developed and implemented across multiple sites to accurately identify patients who should have a CTA triggered to facilitate EHR data collection. However, CC groupings will necessarily vary in order to attain high sensitivity and moderate-to-high specificity. In future trials, the balance between sensitivity and specificity should be considered based on the nature of the clinical condition, including prevalence and morbidity, in addition to the goals of the intervention being considered.

**Correspondence to:**

Sara J. Deakyne,
13123 E. 16th Avenue, Aurora, CO 80045, Box B321
Tel.: 720–777–4703
Email: Sara.deakyne@childrenscolorado.org

# 1. Introduction

Traumatic brain injury (TBI) is the leading cause of death and disability in children younger than 18 years worldwide [1, 2]. Fortunately, of the more than 450 000 emergency department (ED) visits annually in the US for head trauma in children, most are for minor trauma [1, 2]. Investigators in the Pediatric Emergency Care Applied Research Network (PECARN) derived and validated two prediction rules that accurately identify children with minor head trauma who are at very-low risk for clinically-important TBI and who do not typically require computed tomography [2]. Subsequently, the PECARN conducted a multicenter trial in order to study the effect of implementing the head trauma prediction rules [3]. The main intervention in the PECARN trial was electronic health record (EHR)-based computerized clinical decision support (CDS), which provided clinicians with an automated risk assessment for clinically-important TBI and a recommendation regarding computed tomography use.

In the PECARN trial, chief complaints (CCs) were used to activate a clinical trial alert (CTA) which then prompted clinicians to populate a head trauma template in the EHR. Little has been published about formal methods to develop, evaluate and refine the accuracy of CC groupings as criteria to trigger a CTA in real-time in the ED setting. Accurate trigger criteria are important in order to minimize the number of times clinicians ignore seemingly excessive alerts (known as "alert fatigue") [4–16]. In prior studies, criteria for initiating alerts, both clinical alerts and CTAs, have included more well-defined criteria such as age, medication ordered, physical attributes of a patient such as weight or body mass index, social history, medical diagnoses, test results or location of patient encounter [17–25]. These criteria are discretely documented in the EHR and their use in research is more clearly defined than are CCs. In some EHRs, including the ones used in this study, CCs are discretely documented from structured lists, but with substantial variability in CCs from which to choose.

# 2. Objective

Our objective was to develop, implement and evaluate groupings of nurse-assigned CCs across multiple sites that accurately identified children with head trauma, in order to activate a CTA and specific data collection in an EHR.

# 3. Methods

As part of the 13 site clinical trial comparing cranial computed tomography use before and after implementation of CDS, four PECARN sites (three freestanding children's hospital EDs and one general ED with a separate pediatric ED) participated in a process to centrally develop and then locally implement CC groupings to trigger the completion of specific data points in an ED head trauma data collection template. All four sites used Epic Systems® (Verona, WI) as their ED EHRs. The other sites participating in the trial also used CTAs to activate the template; however, they chose to use local processes to create their CC groupings rather than central development.

▶Figure 1 details the expected process from the time of patient presentation in the ED to the provision of CDS. Based on ED evaluations and workflow analyses conducted prior to the trial, CCs were found to be one of the first events that occurred in the workflow, and assigned prior to ordering of imaging. Therefore, we believed that CCs were likely the best criteria available in the EHR to determine the appropriate patients for whom the CTA should be activated in real-time. We designed the CTA so that if the patient had a CC entered that was among the CC list (grouping), the provider would be prompted via the CTA (here forward termed the "trigger CTA") to answer a question regarding whether a patient had head trauma within the preceding 24 hours (and thus should receive CDS).

At the four sites, we used an iterative, six-phase process for the development, testing, validation and refinement of the CC groupings, which is the focus of this manuscript (▶Figure 2). An Epic Clarity (Clarity is Epic's analytical database) certified analyst led the centralized process. In Phase 1,

the analyst (SD) and lead study investigator (PD) compiled a comprehensive list of ICD-9 codes to identify patients with head trauma or other injuries. We used ICD-9 codes as the reference standard for defining head trauma as these are assigned for all patients and reflect the final diagnoses. The ICD-9 codes were selected by the lead investigator (PD) and reviewed by the analyst. We chose to use a broad inclusion of ICD-9 codes to ensure that any patient that may have experienced head trauma was included in our study population. For example, we included ICD-9 code 920: contusion of the face, scalp and neck except eye(s). At three of the four sites, we used billing ICD-9 diagnoses to define our head trauma population. However, one site did not use the EHR billing module, and thus lacked billing ICD-9 diagnoses within the EHR. For this site, we used clinician-assigned encounter ICD-9 diagnoses instead.

Based on the ICD-9 list, the analyst developed site-specific Epic Clarity reports that were distributed to sites and run in their local EHRs. In these reports, we collected the CC frequencies for patient encounters corresponding to the selected ICD-9 codes. Each site collected data retrospectively for a one-year period from January through December 2010. The CCs were routinely assigned by triage nurses as part of standard practice. A nurse could select multiple CCs for a given patient encounter. The CCs were discrete words or phrases selected from dropdown lists and free text was only allowed as comments added to the chosen CCs. The comments were not collected in the frequency distributions. Furthermore, physician-assigned, free-text CCs were not considered. Although each site used a drop-down list of CCs, the list was different at each of the sites.

In Phase 2, we used these CC frequencies to construct multiple CC groupings for testing, including the top 15 and 30 CCs at each site. In addition, the lead investigators developed combinations based on these lists by removing individual and groups of CCs that were expected to have low specificity based on clinical experience. We aimed to create groupings that, upon testing, would have at least 90% sensitivity to identify those with head trauma and to attain as high as possible specificity to limit alert fatigue.

In Phase 3, the analyst created site-specific reports to assess the potential accuracy of selected CC groupings. Reports contained variables to indicate the presence or absence of a head trauma ICD-9 code and whether the patient was assigned any of the CCs from each of the groupings tested. The reports were run at each site, including all patient encounters (regardless of diagnosis or CC), during the same time period as that used to create the frequencies in Phase 1. De-identified data from the sites were exported and provided to the central analyst.

In Phase 4, we identified the potentially optimal CC grouping for each site by inspection of the test characteristics. The lead study investigators (PD, NK) and analyst (SD) reviewed the groupings' sensitivities and specificities. We chose to favor CC groupings with high sensitivity, balanced with as high as possible specificity. For CC groupings with similar test characteristics, we favored the CC groupings in which individual CCs were more specific for head trauma. At some sites, additional groupings were tested for further optimization.

After initiation of the clinical trial, we formally retested the accuracy of the CC groupings (Phase 5) at two sites that received feedback from the clinicians that the CTAs were "over-firing." To do so, we completed the same processes as in Phases 3 and 4 to generate reports and calculate the test characteristics of the CC groupings based on data collected in the trial up to that point. We calculated both the predicted and actual sensitivities of the initially implemented CC groupings based on the test and reference standard variables noted in ▶Table 1. The actual sensitivity and specificity was based on whether the trigger CTA was truly activated, while the predicted was based on whether the patient had the CC.

We anticipated that individual ED workflows would not always follow the exact order noted in ▶Figure 1. Thus, we expected that reassessment of the sensitivities and specificities (and positive predictive values) of the CC groupings would differ from the initial assessment, particularly because a trigger event was also required to lead to the trigger CTA activation. The trigger event is an event in the EHR that must occur in order for the CCs assigned by the nurse to be evaluated against the CCs in the grouping. Importantly, we needed to choose a trigger event that would occur after CC assignment and before the decision on whether to order a head computed tomography scan. Based on the nursing workflow analysis (▶Figure 1), we selected the filing of vital signs as the trigger event. Selection of this trigger event was also important to prevent falsely firing the trigger CTA prior to the time when the CC entry would occur in the workflow.

In phase 6, we attempted to improve the accuracy of the CC groupings (Phase 6). At three sites, including the two that had conducted formal reassessments of the CC groupings' sensitivities and specificities, we assessed some modified versions of the CC groupings. This assessment included creating new reports and calculating test characteristics as completed in phases 3 and 4, but with new groupings, using data from the first year of the trial. One site felt their CTA was firing as needed, so did not undergo this reassessment. At the three sites that desired to improve accuracy, we obtained informal feedback from the clinical staff to create modified versions of the CC groupings for further testing. Based on the test characteristics of the modified CC groupings, we implemented specific changes at individual sites, such as removing select CCs from the grouping.

We created all reports using Crystal Reports® 2008 and calculated test characteristics using SAS 9.3 (SAS Institutes, Cary, NC).

# 4. Results

►Table 2 displays the top ten nurse-assigned CCs associated with the head trauma ICD-9 codes at each site. Note that this is not an exhaustive list of the patients' self-reported complaints, but rather the nurses' interpretations of the children's complaints categorized according to the available CC choices at the given site. A head trauma CC was either the most frequent or second most frequent CC for patients with actual head trauma at each site. Although all had head trauma CCs among their ten most frequent, the verbatim wording of the CCs available in the EHR systems and selected by the nurse differed, as did the proportion of patients who were identified by specific CCs. Two sites captured more than 40% of their patients with the head trauma CCs, whereas the other two sites captured less than 25%. All sites had a laceration CC among the ten most common CCs; however, some sites had more specific terminology, such as "head laceration" or "laceration forehead", while others had more general terminology, such as "laceration". Furthermore, all sites had a CC regarding fall mechanisms and headache in the top ten CCs. Other common CCs for head injured patients included symptoms, such as vomiting, loss of consciousness or dizziness. Additionally, sites frequently used general trauma categories such as trauma or minor trauma and specific trauma mechanisms, such as motor vehicle-related and bicycle-related injuries. Differences across sites included two sites that used CCs that alerted a trauma team (e.g. "critical trauma" and "trauma alert") among the ten most common CCs and another had patient encounters without any CC (termed throughout manuscript as 'null') selected.

Using the CC frequencies identified in Phase 1, we constructed the CC groupings in Phase 2 to maximize sensitivity, balanced with at least moderate specificity. The study PIs and analyst created several broad CC groupings, such as those with all trauma-related CCs, all head trauma CCs, the most frequent 15 and 30 CCs and combinations of these groups. We created many additional combinations to potentially improve specificity, removing a single or groups of CCs in the process. All CC groupings included a null CC, in order to ensure that patients without a CC were still assessed for the presence of head trauma.

In Phase 3, each site tested 17–26 combinations of CC groupings. Due to the variability in the CCs available across sites, we were unable to create identical groupings for testing at each site. The varying numbers of combinations at each site were necessary to account for varied use of different types of CCs, including symptom-related CCs and general CCs, as well as varied patterns of CC use. Additionally, some sites used few CCs for patients with head trauma, and others used a wide range of CCs for these patients.

On review of the CC groupings' test characteristics (Phase 4), sensitivities ranged from 77% to 98%, and specificities from 71% to 94%. There was substantial variation in the sensitivities and specificities across seemingly similar CC groupings. For example, one site had a specificity of 78% for the top 15 CCs at their site while the three other sites had specificities near 85%. ►Table 3 displays the test characteristics of the top two CC groupings by site, including the one that was ultimately chosen and the second choice, which was not used. The CC groupings chosen were similar with regard to typical head trauma CCs but varied in their inclusion of more general CCs. The variations in the chosen groupings among sites were often due to the differences in the available CCs. For example, at one site, 17% of head-injured patients had a CC of either 'MT (minor trauma) injury' or

'MT (minor trauma) other'; other sites either did not have these somewhat general CCs available or used more specific CCs instead.

In Phase 5, with the clinical trial underway, we re-evaluated the CC grouping test characteristics at two sites for the trial period. ►Table 4 displays the results of this re-evaluation, noting a lower sensitivity compared to the baseline data in ►Table 3 for both site 1 and 2 and a lower specificity than baseline for site 2. We also examined positive predictive values, which were low for both sites due to the low prevalence of head trauma in those identified by the grouping. At site 1, 1 034 (38%) of 2 693 encounters with head injury ICD-9 codes had either general CCs (e.g. 'MT (minor trauma) Injury') or non-specific symptom CCs (e.g. 'Neuro Headache') entered rather than CCs specific to head trauma. The negative predictive value, however, was high for both site 1 (99.3%) and site 2 (99.7%). Thus, >99% of the patients identified by the CC grouping as not having a head injury did not, in fact, have a head injury. Additionally, we found differences between the expected sensitivity and specificity compared to the actual. The differences resulted from the CTA requirement for two events to occur in order to actually be triggered, including selection of the CC and either entering vital signs or verifying allergies. If these events did not occur as expected, then the CTA was either incorrectly triggered or not triggered at all.

In Phase 6, we assessed the potential effect of refining the CC groupings at three sites, mainly in an attempt to improve specificity and limit excess alerts. Site provider and investigator feedback led to testing the effect of removing the null CC and removing more general trauma complaints that were most often identifying patients without head trauma. ►Table 5 demonstrates the test characteristics of the initial CC grouping without the null and select new combinations that were tested. Removing the null at site 3 improved specificity by 2.9% (to 91.6%), with the sensitivity dropping slightly to 91.2%. Therefore, this site removed the null CC from their grouping. At site 1, although the specificity increased 4% (to 86.8%), removing the null dropped sensitivity to 87.9%; therefore, the null was retained in the CC grouping. At site 1, removing less specific CCs such as 'MT (minor trauma) Injury' improved specificity modestly, but decreased sensitivity by more than 50%. In contrast, to attain modest increases in specificity at sites 2 and 3, sensitivity decreased substantially less. In fact, both sites 2 and 3 were able to achieve a corresponding sensitivity and specificity in the 90s, whereas site 1's highest specificity with a corresponding sensitivity in the 90s was 85.9%. The larger decrease in sensitivity to achieve high specificity and the inability to maintain both sensitivity and specificity in the 90s at site 1 was due to the frequent use of general and symptomatic CCs at this site that were not specific to head trauma. Changes to the CC groupings were implemented at site 3, while site 1 decided to instead educate staff to select CCs that were more specific to the problem for which the patient presented to the ED. Site 2 did not make any changes.

Importantly, at site 3 the provider feedback also uncovered that the trigger CTA was activated for patients who did not have a CC in the grouping. On closer examination, we found that CC synonyms were assigned to some of the CCs in that particular grouping. For example, "ear injury" was included in one CC grouping, but "ear pain" was in the system as its synonym. Due to this synonym relationship, the trigger CTA was activated for many patients with non-trauma related ear pain, including ear infections. Prior to removing these CCs from the grouping, we encouraged this site to request that the synonyms be removed from the CCs in their EHR. The site was unable to implement this change during the trial, so this CC was removed from their grouping.

At one site, it was discovered that patient workflow had changed, resulting in inappropriate activation of the trigger CTA for a number of patients. In this new workflow, a 'first nurse' was added during the busiest hours to perform a brief assessment prior to triage, including obtaining vital signs, but without assigning a CC. This site previously chose vital sign entry as the event that would cause the trigger CTA to assess a patient's CCs against the grouping and open the head trauma question if a grouping CC was present. Because this site included a null CC in the grouping for the trigger CTA, the 'first nurse' workflow change led to an excess of alerts fired. The site changed the trigger event that caused the trigger CTA to evaluate the CCs to allergy verification in order to account for the workflow change.

## 5. Discussion

Across multiple sites, we successfully developed and implemented highly-sensitive and moderately-to-highly specific CC groupings to activate a CTA that identified children who presented to the ED with head trauma. The chosen CC groupings necessarily differed across sites to achieve high sensitivity and to minimize falsely identifying those without head trauma. Importantly, despite the moderately high specificities, the positive predictive values were low. Furthermore, the number and types of chosen CCs differed across sites, mainly due to the lack of consistency among staff and across sites in the use of CCs for similar patients, and the frequent use of general trauma and general symptom CCs rather than those specific to head trauma. The actual sensitivities and specificities of the CC groupings differed from those predicted due to changes in the expected workflow, such as the addition of the 'first nurse', and unanticipated design features in the EHR, such as CC synonyms. Although we might have tried to enforce a specific workflow to prevent misfiring of the CTA trigger, we chose a pragmatic approach under the real-world conditions of an ED, in which unexpected changes in workflow occur due to varying patient volumes and the emergent nature of some patients' conditions.

Little has been published about formal methods to develop, evaluate and refine the accuracy of CC groupings as criteria for a trigger CTA in the ED setting. In one prior study, investigators used CCs, among other variables, to trigger an alert reminding primary care physicians of research studies in which a patient may qualify for enrollment [25]. This prior study chose to favor specificity over sensitivity in order to minimize alert fatigue at the expense of missing eligible patients. In this study, we chose to favor sensitivity in order to decrease the chance of missing eligible patients. However, the feedback provided by clinicians during our reassessment of the CC groupings (Phase 6), as well as prior literature, emphasize the importance of minimizing CTA and clinical alert firing when not necessary [4, 5, 9, 11]. In fact, the European Regional Development Fund provided priorities to maximize the effectiveness of alerts in CDS for computerized order entry including: 1) determining the optimal sensitivity and specificity of alerts; 2) determining whether personalization of alerts will reduce alert fatigue; and 3) determining whether appropriate timing of an alert will reduce alert fatigue [27]. Alert fatigue may lead to overriding or ignoring the content of the alert, even when the evidence to support the recommendation is strong [4, 5, 9, 11]. In one recent randomized trial, the investigators noted a progressive decrease in response rate to a CTA from 50% to 35% over a 36 week period [5]. The clinical importance of overriding alerts and alert fatigue varies, but can be important, with one recent publication noting a misdiagnosis of a drug allergy due to an alert override [7, 8, 26].

Most prior studies have used objective measures to trigger CTAs, such as age, medication ordered, patient physical attributes, test results, prior diagnoses and patient location [17–25]. These criteria can be highly accurate, in particular when all inclusion criteria are objective, discrete and entered into the EHR prior to the time at which the alert criteria are assessed. We were not able to use objective trigger criteria in our study for several reasons. First, our study included an intervention for an acute condition with no prior diagnoses relevant to the current head trauma available in the patient medical records to use as criteria for an alert. Second, we sought to identify all patients with head trauma, regardless of severity and whether any diagnostic tests were ordered. Therefore, we were unable to base the identification of patients with head trauma on a clinician order, such as a cranial computed tomography scan, which is only ordered routinely for patients with more severe head trauma. Furthermore, we aimed to collect data about head trauma patients prior to diagnostic test decision-making; thus, there were no pre-existing specific clinical data that could have been used to identify these patients.

Finally, our results suggest methods to improve the across-site accuracy, consistency and use of CCs to identify patients for whom a CTA or CDS is appropriate. Efforts may focus at the staff or EHR levels. Specifically, staff training should include emphasizing the important role of assigning CCs and the need for consistency in assignment. Alternatively, focus should be placed on the EHR design by encouraging EHR vendors to create specific CCs that are consistent across EHR systems. This would require input from clinicians and potentially use methods to limit the number of CCs available, similar to the process used to streamline one pediatric diagnosis-based severity classification system [28]. The large variability in the CCs available at sites and the variation in assigning

CCs resulted in the need to include more than 30 CCs in each site's grouping in order to obtain greater than 90% sensitivity. This variability not only leads to an increase in the effort required and complexity in developing groupings to trigger CDS, but also limits our ability to understand trends in CCs for ED visits at the local, national and international levels. The lack of specificity in assigning CCs leads to difficulty in determining denominator data needed to assess diagnostic accuracy for patients with undifferentiated illnesses presenting to an acute care setting [29].

Our study had certain limitations. We based the assessment of the CCs for our groupings on ICD-9 codes, which may not identify all patients with head trauma and, conversely, falsely indicate that patients sustained head trauma. A more comprehensive method to identify patients with head trauma would have been through chart review; however, this was felt to be overly labor intensive and not reproducible outside of research. Although the PECARN prediction rules vary based on age, we did not assess age differences in CCs. Additionally, only two sites formally assessed the actual performance of the groupings. These sites made changes based on re-assessment, either modifying the CCs used in their CC groupings or educating staff. However, no formal reassessment was undertaken to determine the effect of these changes. Finally, we did not use statistical modelling techniques, such as recursive partitioning or area under the receiver operating characteristic curves, to create the CC groupings. Although these techniques may allow for a more precise, analytical approach to assessing the accuracy and ideal selection of potential groupings, use of these techniques requires specific training and may not be generalizable.

# 6. Conclusions

CC groupings based on ICD-9 codes can be successfully developed and implemented across multiple sites to accurately identify research study patients who should have a CTA triggered and facilitate data collection in the EHR for children with head trauma. The balance between sensitivity and specificity in CC groupings must be carefully considered based on the nature of the clinical condition, including its prevalence and its morbidity, as well as the goals of the intervention being considered. Further efforts are needed to assure that specific CCs are available and similarly used across EHRs and medical systems.

**Clinical Relevance**
The findings in this manuscript have implications for the methods by which informaticists develop and refine CTAs and clinical alerts to identify intended patients based on available data within the EHR at the time of clinical care decisions. To improve the accuracy of groupings of CCs to identify appropriate patients, we highlight the importance of uniformity in the available choices of CCs, the need for designing CC lists with specific rather than general CCs within the ED (where CCs are often the only available source for defining the denominator of patients), and the usefulness of reassessing and refining the CC groupings to maximize accuracy.

**Conflicts of Interest**
The authors declare that they have no conflicts of interest in the research.

**Protection of Human Subjects**
The clinical trial was reviewed by each individual sites' institutional review board and either was granted a waiver of consent and HIPAA authorization or was deemed non-human subjects research.

**Fig. 1**    Expected process from the time of patient presentation in the ED to the provision of CDS.

CC: chief complaint; TBI: traumatic brain injury; CTA: clinical trial alert

*Sites used either filing vital signs or verifying allergies as the event that would lead to assessing the CCs against the grouping.
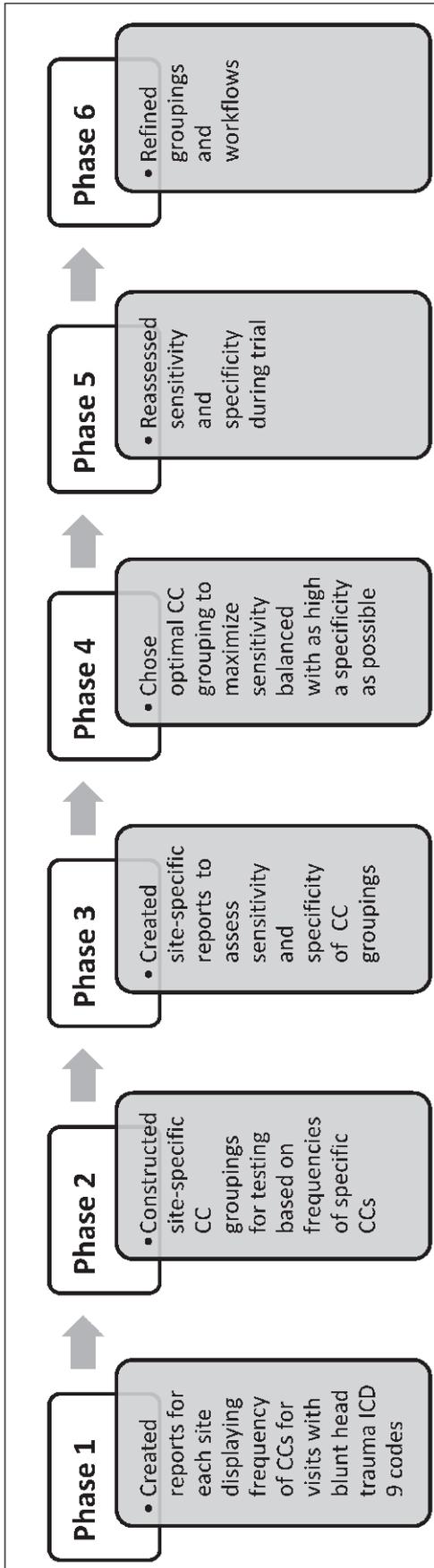
**Fig. 2** An iterative, six-phase process for the development, testing, validation and refinement of the CC groupings.

CC: chief complaint

**Table 1** Definitions of variables used for calculating sensitivity and specificity of initially–implemented CC groupings at two sites during the trial (Phase 5).

|  | Test Variable | Reference Variable | Purpose |
|---|---|---|---|
| Predicted | Presence of any grouping CC | ICD-9 code for head trauma | Determine whether the assignment of CCs during the study period was similar to the period used to create the groupings |
| Actual | Trigger CTA fired | ICD-9 code for head trauma | Determine whether the CC groupings performed as predicted |

CC: chief complaint; CTA: clinical trial alert

**Table 2** Top ten most frequent CCs (verbatim from EHR system) at each site for head trauma encounters based on ICD-9 (Phase 1).

|  | Site 1 | | Site 2 | | Site 3 | | Site 4 | |
|---|---|---|---|---|---|---|---|---|
|  | Verbatim CC | % of enounters* | Verbatim CC | % of encounters* | Verbatim CC | % of encounters* | Verbatim CC | % of encounters* |
| 1 | MT (Minor trauma) fall | 25% | Head injury | 44% | Head injury | 42% | Critical trauma level III blunt | 23% |
| 2 | MT CHI (closed head injury) | 22% | Laceration | 30% | Head laceration | 15% | Head trauma minor no LOC (loss of consciousness) | 21% |
| 3 | Laceration forehead | 12% | Fall | 17% | Fall | 9% | Head trauma positive LOC but awake | 13% |
| 4 | MT injury | 11% | Level 2 trauma alert | 5% | Facial laceration | 9% | Critical trauma level II blunt | 10% |
| 5 | MT other | 6% | Vomiting | 4% | Laceration | 6% | Critical trauma level I blunt | 5% |
| 6 | Laceration face | 5% | Headache | 3% | Vomiting | 3% | Laceration simple | 4% |
| 7 | (Null – no CC entered) | 3% | Motor vehicle related injury | 2% | Fall down stairs | 3% | Falls | 4% |
| 8 | Neurology headache | 2% | Facial injury | 2% | Headache | 2% | Trauma, minor complaint | 3% |
| 9 | MT MVC (motor vehicle collision) | 2% | Level 1 trauma alert | 2% | Motor vehicle crash | 2% | Headache | 2% |
| 10 | Head injury | 2% | Head pain | 1% | Fall from furniture | 2% | Altered mental status acute | 2% |

*Columns do not add to 100%, as not all CCs are listed; patients may also have more than one CC during a given encounter
CC: chief complaint; MT: minor trauma; MVC: motor vehicle collision; LOC: loss of consciousness; CHI: closed head injury

**Table 3**   Top two chief complaint grouping test characteristics prior to trial initiation (Phase 4).

|  |  | Sensitivity | Specificity | # CCs |
|---|---|---|---|---|
| Site 1 | Chosen | 95.1% | 83.4% | 50 |
|  | Second Choice | 94.0% | 85.6% | 16 |
| Site 2 | Chosen | 97.8% | 83.8% | 37 |
|  | Second Choice | 95.1% | 89.1% | 28 |
| Site 3 | Chosen | 96.6% | 88.6% | 34 |
|  | Second Choice | 95.5% | 91.0% | 33 |
| Site 4 | Chosen | 95.5% | 75.3% | 31 |
|  | Second Choice | 94.9% | 78.4% | 24 |

CC: chief complaint

**Table 4**   Reassessment of chief complaint groupings at two sites during the trial (Phase 5).

| Site | Expected* | | | | Actual* | | | |
|---|---|---|---|---|---|---|---|---|
|  | Sensitivity | Specificity | PPV | NPV | Sensitivity | Specificity | PPV | NPV |
| 1 | 90.8% | 82.8% | 23.9% | 99.3% | 91.5% | 74.0% | 17.4% | 99.7% |
| 2 | 94.6% | 84.2% | 20.8% | 99.3% | 94.7% | 83.7% | 20.4% | 99.7% |

*Data were gathered during the trial period. Expected test characteristics were determined using the entered chief complaint as the test variable and ICD-9 diagnosis as the reference. Actual test characteristics were determined based on whether the CTA was triggered as the test variable and ICD-9 diagnosis as the reference.

**Table 5**   Sensitivity and specificity of additional CC groupings tested during refinement phase of the trial (Phase 6).

|  | Site 1 | | | Site 2 | | | Site 3 | | |
|---|---|---|---|---|---|---|---|---|---|
|  | Sensitivity | Specificity | PPV | Sensitivity | Specificity | PPV | Sensitivity | Specificity | PPV |
| Initial CC grouping without null | 87.9% | 86.8% | 27.5% | Not completed | | | 91.2% | 91.9% | 32.9% |
| Refined grouping with highest specificity (with associated sensitivity)* | 43.5% | 96.1% | 37.6% | 75.6% | 95.8% | 54.3% | 75.8% | 97.2% | 58.9% |
| Refined grouping with highest specificity with sensitivity at least 90% | 91.8% | 85.9% | 26.1% | 90.4% | 92.6% | 45.7% | 90.0% | 95.8% | 52.8% |

*Of the combinations tested. Clinically-relevant combinations of CCs were created in order to improve specificity while maintaining a high sensitivity.
CC: chief complaint

# References

1. National Center for Injury Prevention and Control (NCIPC). Traumatic brain injury in the United States: assessing outcomes in children. Available at: http://www.cdc.gov/ncipc/tbi/tbi report/index.htm. Accessed May 7, 2010.

2. Kuppermann N, Holmes JF, Dayan PS, Hoyle JD Jr, Atabaki JR, Holubkov R, Nadel FM, Monroe D, Stanley RM, Borgialli DA, Badawy MK, Schunk JE, Quayle KS, Mahajan P, Lichenstein R, Lillis KA, Tunik MG, Jacobs ES, Callahan JM, Gorelick MH, Glass TF, Lee LK, Bachman MC, Cooper A, Powell EC, Gerardi MJ, Melville KA, Muizelaar JP, Wisner DH, Zuspan SJ, Dean JM, Wootton-Gorges SL for the Pediatric Emergency Care Applied Research Network (PECARN). Identification of children at very low risk of clinically-important brain injuries after head trauma: a prospective cohort study. Lancet 2009; 374: 1160–1170.

3. Dayan PS. Traumatic Brain Injury – Knowledge Translation (TBI-KT). In: ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US) 2000 [cited 2014 Oct 20]. Available from: https://clinicaltrials.gov/ct2/show/NCT01453621 NLM Identifier: NCT01453621.

4. Sittig DF, Wright A, Osheroff JA, Middleton B, Teich JM, Ash JS, Campbell E, Bates DW. Grand challenges in clinical decision support. J Biomed Inform 2008; 41(2): 387–392.

5. Embi PJ, Leonard AC. Evaluating alert fatigue over time to EHR-based clinical trial alerts: findings from a randomized controlled study. J Am Med Informatics Assoc 2012; 19(e1): e145–e148.

6. Karsh BT. Clinical practice improvement and redesign: how change in workflow can be supported by clinical decision support. AHRQ Publication. Rockville, Maryland: Agency for Healthcare Research and Quality 2009.

7. Isaac T, Weissman JS, Davis RB. Overrides of medication alerts in ambulatory care. Arch Intern Med 2009; 169: 305–311.

8. Weingart SN, Toth M, Sands DZ, Aronson MD, Davis RB, Phillips RS. Physicians' decisions to override computerized drug alerts in primary care. Arch Intern Med 2003; 163: 2625–2631.

9. Glassman PA, Belperio P, Simon B, Lanto A, Lee M. Exposure to automated drug alerts over time: effects on clinicians' knowledge and perceptions. Med Care 2006; 44: 250–256.

10. Payne TH, Nichol WP, Hoey P, Savario J. Characteristics and override rates of order checks in a practitioner order entry system. Proc AMIA Symp 2002; 602–606.

11. Ahearn MD, Kerr SJ. General practitioners' perceptions of the pharmaceutical decision-support tools in their prescribing software. Med J Aust 2003; 179(1): 34–37.

12. Abookire SA, Teich JM, Sandige H, Paterno MD, Martin MT, Kuperman GJ, Bates DW. Improving allergy alerting in a computerized physician order entry system. Proc AMIA Symp 2000; 2–6.

13. Shah NR, Seger AC, Seger DL, Fiskio JM, Kuperan GJ, Blumenfel B, Recklet EG, Bates DW, Gandhi TK. Improving acceptance of computerized prescribing alerts in ambulatory care. J Am Med Inform Assoc 2006; 13(1): 5–11.

14. Paterno MD, Maviglia SM, Gorman PN, Seger DL, Yoshida E, Seger AC, Bates DW, Gandhi TK. Tiering drug-drug interaction alerts by severity increases compliance rates. J Am Med Inform Assoc 2009; 16: 40–46.

15. Osheroff JA [editor]. Improving medication use and outcomes with clinical decision support: a step-by-step guide. Chicago, IL: Health Information and Management Systems Society 2009.

16. Tamblyn R, Huang A, Taylor L, Kawasumi Y, Bartlett G, Grad R, Jacques A, Dawes M, Abrahamowicz M, Perreault R, Winslade N, Poissant L, Pinsonneault A. A randomized trial of the effectiveness of on-demand versus computer-triggered drug decision support in primary care. J Am Med Inform Assoc 2008; 15: 430–438.

17. Embi PJ, Jain A, Clark J, Bizjack S, Hornung R, Harris CM. Effect of a clinical trial alert system on physician participation in trial recruitment. Arch Intern Med. 2005; 165(19): 2272–2277.

18. Heinemann S, Thuring S, Wedeken S, Schafer T, Scheidt-Nave C, Ketterer M, Himmel W. A clinical trial alert tool to recruit large patient samples and assess selection bias in general practice research. BMC Med Res Methodol 2011; 11: 16.

19. Koplan KE, Brush AD, Packer MS, Zhang F, Senese MD, Simon SR. "Stealth" alerts to improve warfarin monitoring when initiating interacting medications. J Gen Intern Med 2012; 27(12): 1666–1673.

20. Ledwich LJ, Harrington TM, Ayoub WT, Sartorius JA, Newman ED. Improved influenza and pneumococcal vaccination in rheumatology patients taking immunosuppressants using an electronic health record best practice alert. Arthritis Rheum 2009; 61(11): 1505–1510.

21. Tang JW, Kushner RF, Cameron KA, Hicks B, Cooper AJ, Baker DW. Electronic tools to assist with identification and counseling for overweight patients: a randomized controlled trial. J Gen Intern Med 2012; 27(8): 933–939.

22. Haerian K, McKeeby J, Dipatrizio G, Cimino JJ. Use of clinical alerting to improve the collection of clinical research data. AMIA Annu Symp Proc 2009; 218–222.
23. Mathias JS, Didwania AK, Baker DW. Impact of an electronic alert and order set on smoking cessation medication prescription. Nicotine Tob Res 2012; 14(6): 674–681.
24. Jain A, McCarthy K, Xu M, Stoller JK. Impact of a clinical decision support system in an electronic health record to enhance detection of $\alpha_1$-antitrypsin deficiency. Chest 2011; 140(1): 198–204.
25. Grundmeier RW, Swietlik M, Bell LM. Research subject enrollment by primary care pediatricians using an electronic health record. Amia Annu Symp Proc 2007: 289–293.
26. Carspecken CW, Sharek PJ, Longhurst C, Pageler NM. A clinical case of electronic health record drug alert fatigue: consequences for patient outcome. Pediatrics 2013; 131(6): e1970–e1973.
27. Coleman JJ, van der Sijs H, Haefeli WE, Slight SP, McDowell SE, Seidling HM, Eiermann B, Aarts J, Ammenwerth E, Slee A, Ferner RE. On the alert: future priorities for alerts in clinical decision support for computerized physician order entry identified from a European workshop. BMC Med Inform Decis Mak 2013; 13(1): 111.
28. Alessandrini EA, Alpern ER, Chamberlain JM, Shea JA, Holubkov R, Gorelick MH, Pediatric Emergency Care Applied Research Network. Developing a diagnosis-base severity classification system for use in emergency medical services for children. Acad Emerg Med 2012; 19(1): 70–78.
29. Iyer S, Reeves S, Varadarajan K, Alessandrini E. The Acute Care Model: A New Framework for Quality Care in Emergency Medicine. Clinical Pediatric Emergency Medicine 2011; 12(2): 91–101.