

## Original Article

## Open Access

Mark L. Graber\*, Asta V. Sorensen, Jon Biswas, Varsha Modi, Andrew Wackett, Scott Johnson, Nancy Lenfestey, Ashley N.D. Meyer and Hardeep Singh

# Developing checklists to prevent diagnostic error in Emergency Room settings

## Abstract

**Background:** Checklists have been shown to improve performance of complex, error-prone processes. To develop a checklist with potential to reduce the likelihood of diagnostic error for patients presenting to the Emergency Room (ER) with undiagnosed conditions.

**Methods:** Participants included 15 staff ER physicians working in two large academic centers. A rapid cycle design and evaluation process was used to develop a general checklist for high-risk situations vulnerable to diagnostic error. Physicians used the general checklists and a set of symptom-specific checklists for a period of 2 months. We conducted a mixed methods evaluation that included interviews regarding user perceptions and quantitative assessment of resource utilization before and after checklist use.

**Results:** A general checklist was developed iteratively by obtaining feedback from users and subject matter experts, and was trialed along with a set of specific checklists in the ER. Both the general and the symptom-specific checklists were judged to be helpful, with a slight preference for using symptom-specific lists. Checklist use commonly prompted consideration of additional diagnostic possibilities, changed the working diagnosis in approximately 10% of cases, and anecdotally was thought to be helpful in avoiding diagnostic errors. Checklist use was prompted by a variety of different factors, not just diagnostic uncertainty. None of

the physicians used the checklists in collaboration with the patient, despite being encouraged to do so. Checklist use did not prompt large changes in test ordering or consultation.

**Conclusions:** In the ER setting, checklists for diagnosis are helpful in considering additional diagnostic possibilities, thus having potential to prevent diagnostic errors. Inconsistent usage and using the checklists privately, instead of with the patient, are factors that may detract from obtaining maximum benefit. Further research is needed to optimize checklists for use in the ER, determine how to increase usage, to evaluate the impact of checklist utilization on error rates and patient outcomes, to determine how checklist usage affects test ordering and consultation, and to compare checklists generally with other approaches to reduce diagnostic error.

**Keywords:** checklist; clinical reasoning; diagnosis; diagnostic error; emergency medicine.

DOI 10.1515/dx-2014-0019

Received March 23, 2014; accepted May 5, 2014; previously published online June 19, 2014

## Introduction

Diagnostic errors represent a largely unaddressed concern in all healthcare settings, including the Emergency Room (ER) [1, 2]. According to current estimates, one in every 10 diagnoses is likely to be wrong and one in every 1000 ambulatory encounters engenders the risk of harm from a diagnostic error [3, 4]. In aggregate, diagnostic errors are estimated to account for 40,000–80,000 deaths annually in the US [5]. In the ER setting, studies of malpractice claims identify multiple breakdowns in the diagnostic process, and cognitive factors were involved in 96% of these instances [6].

Recent reviews have summarized both the system-related and cognitive interventions that have been proposed to reduce diagnostic error [7–11]. A wide range of well-justified interventions has been considered,

\*Corresponding author: Mark L. Graber, MD FACP, Senior Fellow, RTI International; Professor Emeritus, SUNY Stony Brook School of Medicine; Founder and President, Society to Improve Diagnosis in Medicine, New York, USA, Phone: +919 667-8585, Fax: +631 862-6992, E-mail: mark.graber@va.gov

Asta V. Sorensen and Nancy Lenfestey: RTI International, Research Triangle Park, NC, USA

Jon Biswas: Emergency Department, Michael E DeBakey Veterans Affairs Medical Center, Houston, TX, USA

Varsha Modi, Ashley N.D. Meyer and Hardeep Singh: Houston Veterans Affairs Center for Innovations in Quality, Effectiveness and Safety; Michael E DeBakey Veterans Affairs Medical Center and Baylor College of Medicine, Houston, TX, USA

Andrew Wackett and Scott Johnson: Department of Emergency Medicine, University Hospital, SUNY Stony Brook, NY, USA

including personnel-focused programs, educational interventions, structured process systems, second reviews, and technology-based solutions [9]. Relatively few of these have been evaluated rigorously in practice, although second review programs are now increasingly common in both pathology and radiology programs, and decision support tools to help with differential diagnosis [12, 13], along with online access to medical knowledge resources are increasingly being utilized in direct patient care.

Although the diagnostic process is clearly influenced by many factors inherent in healthcare systems, diagnosis is fundamentally a cognitive process, often based on intuition or gestalt [14]. Successful diagnosis depends on clinical reasoning to synthesize all of the available information and relevant knowledge to arrive at plausible diagnostic possibilities. With an ever-expanding number of diseases to be considered (8000–12,000 according to estimates), the variability in how the disease will manifest and be described and perceived, the substantial uncertainty attending every step of the process, and ultimately the inherent limitations of human memory and cognitive processing, diagnosis is a challenge. The typical ER, with the pressures of workload and time, repeated distractions, and constant multitasking, contributes to this challenge.

Checklists have been developed to address complex and error-prone processes, especially for situations where the outcome is too critical to rely on human memory [15]. Although checklist usage in medical settings is a relatively new endeavor [16], they have already produced impressive reductions in the infection rate from central line placement [17, 18], and in surgical complications [19, 20]. However, the process of diagnosis is not as well defined, and it is not known if checklists will be as effective in this application. The goal of this project was to develop and pilot test a checklist to assist with diagnosis and ultimately help reduce the risk of diagnostic error. To complement symptom-specific checklists that have been previously developed and trialed [21, 22], we sought to develop a more general checklist, and performed initial evaluations of both of these products in an ER setting.

## Methods

### Design

Rapid cycle design and development was used to draft, develop, evaluate, and refine a general checklist for use in ER medical diagnostic evaluations. We solicited input from subject matter experts and physician users on content and

design elements. We used qualitative methods to evaluate usability, physician satisfaction and impact of checklist usage. We also quantitatively assessed test and consult ordering pre- and post-checklist use. The protocols and products were reviewed and approved by a central Institutional Review Board (IRB) at RTI International and separately by the IRB's at the two participating medical centers.

### Setting

We recruited 16 staff-level ER physicians from two academic centers who participated with informed consent. One physician dropped out due to an acute illness. Physicians who completed the project and participated in all of the evaluations were offered a \$500 incentive reward.

### Checklist development

Development of a symptom-specific set of checklists for primary care has been previously described [21, 22]. We focused on developing a usable “general” diagnostic process checklist:

- Version 1: To derive a novel, general checklist for diagnosis we used published checklist design principles and recommendations [23–25] and input from a content-matter expert advisory panel (see acknowledgements) to modify a draft prototype (Version 1, see Appendix) we had previously developed as a starting point [22, 26]. We conducted 45–60 min, structured individual telephone discussions with the seven subject matter experts to obtain input on the ideal components of a general checklist, feedback on Version 1, suggestions on the proposed process for refining and piloting the checklist, and key outcomes to consider. Interviews were transcribed and common themes and major suggestions were identified.
- Version 2a and 2b: Based on input from the subject matter experts and the participating ER physicians, two different general checklists were developed, designated 2a and 2b (see Appendix, Checklists 2a and 2b). Experts recommended we include a complete set of symptom-specific checklists along with the general checklist [22]. Both were then published as a spiral-bound 4×6 inch booklet. Physicians at the two ER sites used the booklet during two work shifts. We then performed cognitive testing to obtain feedback through interviews in person or by telephone. Based on this input and further consultation with the subject matter experts, a revised final “Version 3” checklist was developed.

- The Version 3 checklist, along with reformatted and updated versions of the symptom-specific checklists, (Figure 2A, B) were again published as a spiral-bound booklet sized to fit in a lab coat pocket. Over the next 2 months, the 15 participating physicians were encouraged to use the booklet containing Version 3 twice per shift, focusing on patients they believed had high risk symptoms or “don’t miss” diagnoses. Usage was specifically encouraged for patients with chest or abdominal pain. Physicians were also encouraged to review the appropriate symptom-specific checklist with the patient, and were shown a video demonstrating this approach (<http://tiny.cc/ElyDemo> or <http://www.youtube.com/watch?v=uHpieuyP1w0>).

Physicians completed a post-shift questionnaire (Supplemental Data, which accompanies the article at <http://www.degruyter.com/view/j/dx.2014.1.issue-3/issue-files/dx.2014.1.issue-3.xml> – Checklist Impact Evaluation) examining how and why they used the checklist and whether this had any impact on their clinical decisions.

An overview of the development and evaluation process is illustrated in Figure 1 and the final (Version 3) checklist and one example of a symptom-specific checklist are illustrated in Figure 2.

## Outcome assessment and data analysis

### Qualitative data

Upon completing development of Versions 1 and 2 (2a and 2b), we conducted cognitive interviews with participating physicians to elicit their impressions of the relative advantages, feasibility, usability, satisfaction, usefulness, and checklist usage fit within their clinical workflow. Interviews

were performed using a structured format (see Supplemental Data – Cognitive Testing of Checklist Content and Format) and responses were recorded and coded by two team members (AS, NL) using NVivo 9 software (QSR-International) to identify emergent themes. Themes were identified when at least three individual respondents introduced a topic [27, 28]. A final round of structured interviews was conducted and analyzed similarly after the physicians had used Version 3 for 2 months to evaluate checklist appropriateness, usability, and impact.

### Chart reviews

At one of the test sites, chart reviews were performed to evaluate any major trends in resource utilization as a result of using the checklists. We randomly selected and reviewed a total of 186 medical records relating to ER visits for patients with chest pain or abdominal pain seen by physicians participating in the study. We recorded the number of laboratory tests, imaging exams, and consultations ordered. A total of 104 charts were reviewed of patients seen before the physicians were exposed to any of the checklists, and 82 charts were reviewed of patients seen after they had used the Version 3 checklist for 2 months. These data were compared by  $\chi^2$  analysis.

### Feedback during checklist developmental

The ER user group considered the content of Versions 2a and 2b to be appropriate and helpful in the initial testing, but users suggested replacing some of the items they did not find helpful (such as, “take your own history”) with items identifying specific high risk situations and actions to take in those situations. Combined with other suggestions, this resulted in the revised content used in Version 3.

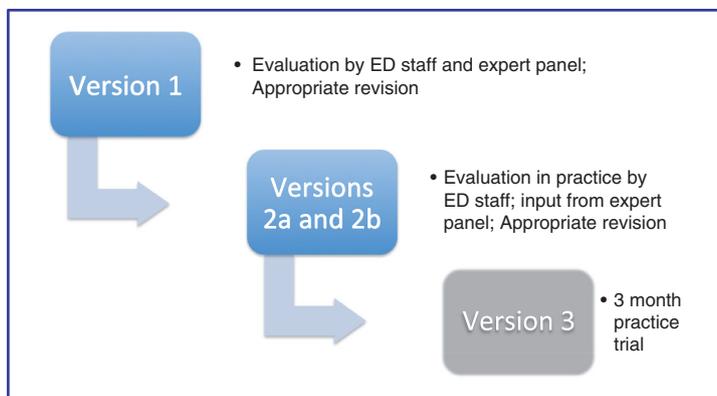


Figure 1 The process used to develop and evaluate a general checklist for diagnosis.

**A**

**HIGH RISK SITUATIONS FOR DIAGNOSTIC ERROR**

**(A “Yes” response to any of the questions puts you at high risk for error)**

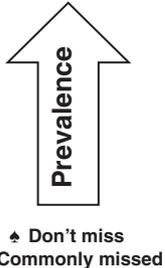
- Are there “must-not-miss” diagnoses that need consideration ?
- Did I just accept the first diagnosis that came to mind ?
- Was the diagnosis suggested to me by the patient, nurse or another MD ?
- Is there data about this patient I haven’t obtained and reviewed ? Old records? Family? Primary care provider ?
- Are there any pieces that don’t fit ?
- Did I read the X-ray myself ?
- Was this patient handed off to me from a previous shift ?
- Was this patient seen in the ER or clinic recently for the same problem ?
- Was I interrupted/distracted/cognitively overloaded while evaluating this patient ?
- Is this a patient I don’t like for some reason ? Or like too much ? (friend, relative)

**What to do in high risk situations:**

1. Pause to reflect - Take a diagnostic “time out”
2. Consider the universal antidote: What else could this be ?
3. Make sure the patient knows when and how to get back to you if necessary

**B**

Dyspepsia, peptic gastritis, esophagitis	Inflammatory bowel disease
Viral gastroenteritis	♣Intussusception
Constipation	♣Carbohydrate malabsorption
Medications	♣Sickle cell crisis
*Psychiatric (depression, sexual abuse)	♣Lead poisoning
Irritable bowel syndrome	♣Hypercalcemia
♣Diverticulitis	♣Gallbladder disease
♣Pancreatitis	Abdominal migraine
Mittelschmerz	Narcotic bowel syndrome
Abdominal wall pain	♣Fitz-Hugh-Curtis syndrome
♣*Appendicitis	Porphyria
♣Kidney stone	Familial Mediterranean Fever
Herpes zoster	Periarteritis nodosa
Giardia	♣Hyper or Hypothyroidism
♣Hernia	Eosinophilic gastroenteritis
Gastroparesis	♣Leukemia
Bacterial overgrowth	Epiplioic appendagitis
♣*Myocardial infarction	Superior mesenteric artery syndrome
Celiac disease	♣Spleen wandering, abscess, infarction
Food poisoning	Hereditary angioedema
♣*Diabetic ketoacidosis	Celiac artery compression
♣*Pneumonia, empyema	Iliocostal friction syndrome
Mesenteric lymphadenitis	Black Widow Spider bite
♣Hepatitis, hepatic abscess	
♣Pelvic inflammatory disease, endometritis	
Leiomyoma, torsion, infarction	
Sphincter of Oddi dysfunction	
♣Ovarian cyst, ovarian torsion	
♣Bowel obstruction, partial obstruction	
♣*Pregnancy complication, ectopic	
Endometriosis, adenomyosis	
♣*Intraabdominal tumor, lymphoma	
Bladder distension, urinary tract infection	
♣*Aortic dissection, ruptured aneurysm	
♣Mesenteric ischemia, infarction	
♣Perforated viscus, peritonitis	
♣Meckel's diverticulitis	

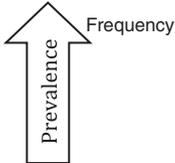


♣ Don't miss  
\* Commonly missed

**C**

**Chest pain**

- Chest wall pain
- Gastroesophageal reflux disease
- Pleurisy, pleuritis
- \*Psychiatric
- ♣\*Coronary disease (myocardial infarction, angina)
- ♣\*Pulmonary embolus
- Herpes zoster
- ♣\*Pneumonia, bronchitis
- ♣Pneumothorax
- Esophageal spasm
- Pericarditis, myocarditis
- ♣Aortic stenosis
- ♣\*Dissecting aortic aneurysm
- ♣\*Tumor
- ♣Esophageal rupture
- ♣Mediastinitis
- ♣\*Spinal cord compression



Frequency

♣ Don't miss \* Commonly missed

**Figure 2** Checklist examples. (A) The Version 3 General checklist. (B) The John Ely specific checklists for abdominal pain. (C) The John Ely specific checklist for chest pain.

## Results

### Quantitative data from post-shift responses

The responses physicians provided after using the checklists 348 times during their last shift are shown in Table 1.

The major findings were 1) Both the general and the specific checklists were used, and the specific checklists were used somewhat more commonly than the general one; 2) Checklist usage was prompted by a variety of different factors, not just diagnostic uncertainty; and 3) In the majority of cases, using the checklist helped confirm the original considerations and had no major impact on the

**Table 1** Physician responses after using checklist usage. Responses reflect data from 348 patient encounters. Not every question was answered in every instance and some questions allowed multiple responses.

	Overall responses	
	n	%
1. Which checklist did you use? ( <i>more than one was selected when appropriate</i> )		
Specific (Checklist symptoms)	166	47.7
High-Risk Situations Checklist (page 1 of the checklist booklet)	117	33.6
Both	62	17.8
2. What prompted you to use the checklist? ( <i>more than one was selected when appropriate</i> )		
Compliance with study protocol	151	43.4
Patient presenting high risk symptoms	139	39.9
Uncertainty in my initial diagnosis	126	36.2
Complex patient history	101	29.0
Teaching tool for resident training	8	7.6
Frequent chief complaint by same patient	4	1.1
Handed off from previous shift	3	0.9
Other (list reasons) ( <i>free text</i> )	2	0.6
3. How did the checklist influence your diagnostic assessment? ( <i>more than one was selected when appropriate</i> )		
It helped me confirm my own working diagnoses	219	62.9
It had no impact on my diagnostic decisions	99	28.4
I changed my working diagnoses as a result of using the Checklist	37	10.6
It slowed me down	21	6.0
It widened differential	6	1.7
It negatively impacted my usual ability to form a diagnosis	3	0.9
4. Did using the checklist help you to consider important diagnoses you hadn't thought of?		
In Part	121	34.8
Yes	118	33.9
No	106	30.5
5. How did using the checklist impact your management plans? ( <i>more than one was selected when appropriate</i> )		
It had no effect on my management plans	247	71.0
I plan to order an additional diagnostic test	77	22.1
Re-examined/re-evaluation	10	2.9
I plan to cancel a pending order for a diagnostic test	8	2.3
I plan to order an additional medication	7	2.0
I plan to cancel a pending order for a medication	4	1.1
It had a negative effect on my management plans (please specify) ( <i>free text</i> )	1	0.3
6. Did the checklist influence your referral decisions? If so, how? ( <i>more than one was selected when appropriate</i> )		
No impact on my referral	298	85.6
Will still refer, but helped me refer more appropriately	20	5.7
I asked for subspecialist opinion	14	4.0
No longer needed to refer	10	2.9
I asked a colleague for second opinion	7	2.0
It impacted my usual referral process in a negative way (please specify) ( <i>free text</i> )	2	0.6

final diagnosis rendered or the management plan chosen. However, using the checklists commonly helped in considering novel diagnostic possibilities (approximately one-third of usages), and in a small but important fraction of cases, using the checklists changed the working diagnosis (37 instances, roughly 10% of usages). In terms of resource utilization, both positive and negative effects were identified: In 21 instances using the checklist was judged to have slowed the physician down and in 77 instances additional diagnostic tests had to be ordered. However, some planned tests were cancelled, there was no substantial

increase in the plans to obtain subspecialty consultation and sometimes there were plans to not refer, refer more appropriately or ask a colleague. Thus, although use of the checklists did not have an impact on management or referrals in most cases, in some cases it led to resource optimization and in some to additional resource utilization. We did not evaluate whether checklist usage was ultimately beneficial to the patient's diagnostic process, or whether any new diagnostic considerations were in fact correct.

Of the 15 physicians, 10 had more than 3 years of experience in the ER ("senior clinicians"). The more

junior clinicians were just as likely to use the general and specific checklists (51 and 52 instances, respectively), whereas the more senior clinicians favored the specific checklists (65 and 115 instances, respectively;  $\chi^2$  5.1; Yates  $p$  value=0.02, Fisher's exact test  $p<0.0001$ ). Experience was not a significant factor in any of the other response categories.

## Qualitative data from final interviews

Several common themes emerged from interviews of the ER physicians after they had used the Version 3 general checklist and the set of symptom-specific checklists for 2 months (see Table 2). Primarily, the comments

corroborated the quantitative findings from the post-shift questionnaires, that the checklists were generally helpful and could help prompt consideration of additional diagnostic possibilities. Several providers noted specific patients where a critical diagnosis would have been missed had the checklists not been used. None of the physicians had read the instructions provided, and for a variety of reasons none had used the toolkit in front of the patient, or with the patient's participation. Several users commented on the usefulness of the general checklist as a way to introduce concepts relating to diagnostic error to trainees. A comprehensive listing of impressions, including comments on barriers and facilitators for usage and suggestions for future improvements is included in the Supplemental Data.

**Table 2** Common themes identified from interviews after checklist exposure.

Theme	Examples from interviews
Checklist usage was generally helpful	<ul style="list-style-type: none"> <li>– “It is definitely helpful; especially for many patients you're not sure of the diagnosis or feel you're missing something”;</li> <li>– “That's when I found it most helpful, particularly when I'm stumped”;</li> <li>– “I like that aspect of expanding my differential diagnosis because in the ER you don't have much time to think unfortunately. My favorite part was going through the lists and tweaking my differential.”</li> </ul>
Checklist usage occasionally prompted consideration of a diagnosis that would have been missed	<ul style="list-style-type: none"> <li>– “I had one very profound situation. A guy came in with flank pain and I didn't actually look at the differential at the time, but it only has 5–6 different things. He actually had a ruptured abdominal aneurism. I ended up scanning him and he had a ruptured aneurism. I would have never thought of that.”</li> <li>– “Someone came in with abdominal pain. He also said he couldn't urinate and was vomiting. I went through the checklist and did a CAT scan to find out he was having a kidney stone. I would have called a surgeon, which wasn't necessary, if I hadn't known he was having a kidney stone.”</li> <li>– “I catch maybe one routine obvious differential that I would have missed if I hadn't looked at the checklist. ... there were two cases. One, I wasn't particularly thinking about thyroid disorder – maybe something like ear pain and I looked through the checklist and it said thyroid so I read up on the differentials and lo and behold there was a thyroid issue. I wouldn't have done that otherwise because it wasn't something I knew right off the bat.”</li> </ul>
The checklists were useful as teaching tools	<ul style="list-style-type: none"> <li>– “I always tell them to go online and find a differential but I'm not aware of a really good site for differential diagnoses. This seems like a really good one but may not be inclusive enough”</li> <li>– “If all the attendings in a department adopted these rules for teaching it would help their own practice because it would be in their consciousness. You can't help but learn it yourself. (It would help) If we all adopted the culture of following these and make everyone refer to it to make sure you're thinking in the best possible way”</li> </ul>
Checklists weren't used in front of, or with the patient	<ul style="list-style-type: none"> <li>– “... if you need to review it you probably don't know what's going on and it doesn't instill confidence in the patient.”</li> <li>– “Would not feel comfortable doing that in front of the patient. I remember as a little boy our family doctor pulled down Merck and would read though it in front of you. Gave you no confidence in his ability.”</li> <li>– “I try to make my own differential then think about if everything fits in there, then look at the book to see if there are things I'm missing. I don't tell the patient those things because I feel it would make him more anxious”</li> </ul>

## Quantitative data from chart reviews pre- and post-checklist use

Comparing the actions of the ER physicians before and after using the Version 3 and symptom-specific checklists on patients with chest or abdominal pain, there was a tendency post-exposure to list more items in the differential diagnosis (1.60 items/patient vs. 1.43,  $p=0.08$ ), to order more laboratory tests (7.05 tests/patient vs. 6.10,  $p=0.17$ ), imaging tests (1.40 tests/patient vs. 1.05,  $p=0.03$ ) and consults (0.22 consults/patient vs. 0.14,  $p=0.24$ ), although the significance of these comparisons were limited due to the small sample size.

## Discussion

The majority of diagnostic errors in medicine reflect cognitive errors [29], and thus the goal of this project was to develop a checklist that would help optimize the clinical reasoning process and help physicians recognize situations that predispose to these errors. To develop a general checklist, we began with some of the major shortcomings identified in the medical literature, then used an iterative process of evaluation and refinement based on input from subject matter experts and end-users. Rather than serving as a how-to guide, the resulting general checklist may cue physicians to pause and reflect in the process for formulating a diagnosis in the ER setting, and to recognize situations prone to cognitive biases.

We found that using the resulting general checklist along with a set of previously developed symptom-specific checklists was generally well accepted by practicing emergency room physicians. Using the checklists was reported to be helpful in confirming initial impressions, and in almost a third of instances the checklist helped the physician consider important diagnoses they had not previously considered. Using the checklists was identified as having prevented several important diagnostic errors, and there were no reports that usage induced new errors. The checklist also encouraged physicians to engage patients in ways that help prevent diagnostic errors or help catch these errors at an early stage, to avoid or minimize harm (“Make sure the patient knows when and how to get back to you if necessary”).

Although both the general and the specific checklists were judged to be valuable, users preferred using the specific checklists for most situations. Shimizu and colleagues have evaluated the same symptom-specific checklists that we used, along with a general debiasing

checklist (our Version 1 checklist) in 188 Japanese medical students asked to evaluate five cases of increasing clinical complexity [30]. Using the general checklist did not increase the proportion of students choosing the correct diagnosis. In contrast, using the symptom-specific checklist was beneficial, at least for the more difficult cases. A similar result was found in a study using checklists to reduce nursing errors in programming IV infusion pumps: a highly specific, step-by-step checklist outperformed a general checklist that encouraged nurses to think critically and consider the ‘five rights’ of medication administration [31]. In our study, physician users were not receptive to a step-by-step approach that emphasized basic elements of the diagnostic process, viewing this as too simplistic. A general checklist that instead focused on high-risk situations therefore evolved, and the results confirmed that high-risk situations were one of the most common reasons to review the checklists. This acknowledges the reality that the synthesis phase of diagnosis, where diagnostic hypotheses are generated from integrating all the available information, is simply not amenable to the step-by-step approach that most checklists employ. Thus, the use of general and symptom specific checklists in the ER setting needs to be explored further, because they appear to be serving different purposes and each may play a role in unique situations. We also believe that the specific checklists used in this study, designed for patients seen in primary care, could be improved by customizing them for patients with urgent and emergency conditions, which often invokes a different differential diagnosis set. Finally, symptom-specific checklists in general are not ideal for patients presenting with multiple problems, or patients whose differential diagnosis is influenced by contextual factors. For example, the differential diagnosis for a healthy patient with weakness would be prioritized differently than the same differential for a patient on chemotherapy for metastatic disease.

A critical issue regarding interventions to improve diagnostic reliability is the extent to which reductions in errors and error-related harm are offset by unintended consequences. In the case of checklists for diagnosis, the goal of prompting consideration of a wider differential diagnosis could potentially lead to additional diagnostic testing and consultation, and thus additional costs in terms of time, money, and possibly complications of these investigations. Although our study was not powered to investigate this definitively, there did appear to be a tendency for providers to order additional diagnostic evaluations post-exposure, albeit the effect appeared to be small, and to some extent offset by cases where tests and consults were cancelled after checklist review. Other

negatives were occasionally noted by users, including a sense in some cases that checklist usage ‘slowed me down’ or was not worth the time invested.

Although the success of checklists to improve surgical safety and to reduce central line infections seems adequately established, the impact of checklists in other medical settings has been mixed [16]. The possibility that checklists could improve diagnostic performance is suggested by the findings of Shimizu discussed above, and recent studies by Sibbald et al. who evaluated the impact of a simple checklist for interpreting electrocardiograms [32]. Even though this checklist included simply the standard steps of ECG interpretation, using the checklist improved accuracy in a group of expert interpreters (senior Cardiology fellows) without increasing the cognitive load or causing expert reversal. Similarly, use of a checklist improved the reliability of identifying key findings on the cardiovascular examination [33].

Our study had several limitations. It was not designed to evaluate whether checklist usage reduced the incidence of diagnostic error or error-related harm, the major outcomes of interest. We also make no assertion that the final general checklist that evolved from our process is the final tool or is superior to any others, and several other general checklists and ‘tips sheets’ for diagnosis exist. A key issue regarding checklist usage in ERs is whether it should be used in every case, or just when use is judged to be “appropriate”. We did not mandate checklist use in every patient, even though many diagnostic errors arise from situations where the diagnosis is established quickly and automatically using ‘System 1’ cognitive processing, and physicians seem generally to be unaware of which of their diagnoses are incorrect [34–36]. We would predict that usage in every case would decrease diagnostic errors and might also increase diagnostic testing, hypotheses that would need to be evaluated in subsequent studies. Generalizing our findings is limited by the small size of the study sites (2) and physicians (15).

The ultimate success of checklists to improve clinical performance will reflect the complex interaction of many different socio-adaptive influences beyond simply ticking off boxes [37–39]. Improving communication amongst team members, education, providing feedback after usage, using clinical champions, endorsement by local leaders, and establishing a cultural expectation of improved safety are among the other elements relevant to whether such checklists can decrease diagnostic error and harm. Stakeholder buy-in is also critically important [24] and the checklists must be appropriately customized in relation to local needs.

In conclusion, the general checklist derived in this project has been extensively evaluated and refined, and in pilot studies was perceived as being helpful in the diagnostic process. Whether checklists can reduce diagnostic error will require further study. Given the complex nature of the diagnostic process, improving diagnostic reliability will no doubt require a set of multi-faceted interventions. Our pilot study suggests that checklists for diagnosis should be considered for further development, implementation and evaluation, perhaps with additional complementary strategies to reduce diagnostic error.

**Acknowledgments:** We sincerely appreciate advice provided by the subject matter experts: Pat Croskerry, John Ely, Robert Wears, Peter Pronovost, Atul Gawande, Key Dismukes, and Dan Boorman. We also thank John Ely for providing his set of symptom-specific checklists.

#### Conflict of interest statement

**Authors’ conflict of interest disclosure:** The authors stated that there are no conflicts of interest regarding the publication of this article. Research funding played no role in the study design; in the collection, analysis, and interpretation of data; in the writing of the report; or in the decision to submit the report for publication.

**Research funding:** This work was supported by an ACTION II contract award from the Agency for Healthcare Research and Quality to RTI International (Contract Number HHSA290200600001I, Prism No. HHSA29032005T, Task Order #8). Drs. Meyer and Singh and Ms. Modi are supported by in part by the Houston VA Center for Innovations in Quality, Effectiveness and Safety (CIN 13-413).

**Employment or leadership:** None declared.

**Honorarium:** None declared.

## Appendix

#### Checklist versions

- Version 1
- Version 2a and 2b
- Version 3 General Checklist and examples of specific checklists.

#### Protocols

- Expert Interview Guide
- Cognitive Testing of Checklist Content and Format
- Checklist Impact Evaluation

**Additional results:** Post Checklist Interview Responses.

## References

- Berner ES, Graber ML. Overconfidence as a cause of diagnostic error in medicine. *Am J Med* 2008;121:S2–23.
- Graber M. The incidence of diagnostic error. *BMJ Qual Saf* 2013;22, Part 2:ii21–ii7.
- Singh H, Traber D, Meyer A, Forjuoh S, Reis M, Thomas E. Types and origins of diagnostic errors in primary care settings. *JAMA Intern Med* 2013;173:418–25.
- Newman-Toker D, Makary M. Measuring diagnostic errors in primary care – the first step on a path forward. *JAMA Intern Med* 2013;173:425–6.
- Leape L, Berwick D, Bates D. Counting deaths from medical errors. *J Am Med Assoc* 2002;288:2405.
- Kachalia A, Gandhi T, Puopolo A, Yoon C, Thomas EJ, Griffey R, et al. Missed and delayed diagnoses in the emergency department: a study of closed malpractice claims from 4 liability insurers. *Ann Emerg Med* 2006;19:196–205.
- Graber M, Kissam S, Payne V, Meyer AN, Sorenson A, Lenfestey N, et al. Cognitive interventions to reduce diagnostic error: a narrative review. *BMJ Qual Saf* 2012;21:535–57.
- Singh H, Graber M, Kissam S, Sorensen AV, Lenfestey NF, Tant EM, et al. System-related interventions to reduce diagnostic errors: a narrative review. *BMJ Qual Saf* 2012;21:160–70.
- McDonald K, Matesic B, Contopoulos-Iannidis D, Lonhart J, Schmidt E, Pineda N, et al. Patient safety strategies targeted at diagnostic errors – a systematic review. *Ann Int Med* 2013;158:381–9.
- El-Kareh R, Hasan O, Schiff G. Use of health information technology to reduce diagnostic error. *BMJ Qual Saf* 2013;22ii:40–4.
- Croskerry P, Singhal G, Mamede S. Cognitive debiasing 2: impediments to and strategies for change. *BMJ Qual Saf* 2013;22ii:65–72.
- Thomas N, Ramnarayan P, Bell M, Maheshwari P, Wilson S, Nazarian EB, et al. An international assessment of a web-based diagnostic tool in critically ill children. *Technol Health Care* 2008;16:103–10.
- Ramnarayan P, Cronje N, Brown R, Negus R, Coode B, Moss P, et al. Validation of a diagnostic reminder system in emergency medicine: a multi-centre study. *Emerg Med J* 2007;24:619–24.
- Cervellin G, Borghi L, Lippi G. Do clinicians decide relying primarily on Bayesian principles or Gestalt perception? Some pearls and pitfalls of Gestalt perception in medicine. *Intern Emerg Med* 2014. DOI 10.1007/s11739-014-1049-8.
- Gawande A. *The checklist manifesto: how to get things right*, 1st ed. New York: Metropolitan Books, 2010.
- Ko H, Turner T, Finnigan M. Systematic review of safety checklists for use by medical care teams in acute care hospitals – limited evidence of effectiveness. *BMC Health Serv Res* 2011;11:211.
- The Matching Michigan Collaboration and Writing Team. Matching Michigan: a 2-year stepped intervention programme to minimise central venous catheter-blood stream infections in intensive care units in England. *BMJ Qual Saf* 2013;22:110–23.
- Pronovost P, Needham D, Berenholz S, Sinopoli D, Chu H, Cosgrove S, et al. An intervention to decrease catheter-related blood-stream infections in the ICU. *N Engl J Med* 2006;355:2725–32.
- Treadwell J, Lucas S, Tsou A. Surgical checklists: a systematic review of impacts and implementation. *BMJ Qual Saf* 2014;23: 299–318.
- Haynes AB, Weiser TG, Berry WR, Lipsitz SR, Breizat AH, Dellinger EP, et al. A surgical safety checklist to reduce morbidity and mortality in a global population. *N Engl J Med* 2010;360:491–9.
- Ely JW, Osheroff JA. Checklists to prevent diagnostic errors. Abstract, Presented at: Diagnostic Error in Medicine Annual Conference, Phoenix-AZ, 2008.
- Ely JW, Graber M, Croskerry P. Checklists to reduce diagnostic errors. *Acad Med* 2011;86:307–13.
- Hales B, Terblanche M, Fowler R, Sibbald W. Development of medical checklists for improved quality of patient care. *Int J Qual Health Care* 2008;20:22–30.
- Thomassen O, Espeland A, Softeland E, Lossius H, Heltne J, Brattebo G. Implementation of checklists in health care; learning from high-reliability organisations. *Scand J Trauma Resusc Emerg Med* 2011;19:53.
- Degani A, Weiner E. Cockpit checklists: concepts, design, and use. *Human Factors* 1993;35:28–43.
- Graber M. Educational interventions to reduce diagnostic error – can you teach this stuff? *Adv Health Sci Educ Theory Pract* 2009;14:63–9.
- Yin RK. *Case study research: design and methods*. Thousand Oaks, CA: Sage Publications, Inc., 1994.
- Miles M, Huberman A. *Qualitative analysis: an expanded sourcebook*, 2nd ed. Thousand Oaks, CA: Sage Publications, 1994.
- Graber ML, Franklin N, Gordon RR. Diagnostic error in internal medicine. *Arch Int Med* 2005;165:1493–9.
- Shimizu T, Kentaro M, Tokuda Y. Effects of the use of differential diagnosis checklist and general de-biasing checklist on diagnostic performance in comparison to intuitive diagnosis. *Med Teacher* 2013;35:e1218–e29.
- White R, Trbovich PL, Easty AC, Savage P, Trip K, Hyland S. Checking it twice: an evaluation of checklists for detecting medication errors at the bedside using a chemotherapy model. *Qual Saf Health Care* 2010;19:562–7.
- Sibbald M, de Bruin A, van Merriënboer J. Checklists improve experts' diagnostic decisions. *Med Educ* 2013;47:301–8.
- Sibbald M, de Bruin ABH, Cavalcanti RB, van Merriënboer JGG. Do you have to re-examine your diagnosis? Checklists and cardiac exam. *BMJ Qual Saf* 2013;22:333–338.
- Meyer AN, Payne VL, Meeks DW, Rao R, Singh H. Physicians' diagnostic accuracy, confidence, and resource requests. *JAMA Internal Med* 2013;173:1952–8.
- Podbregar M, Voga G, Krivec B, Skale R, Pareznik R, Gabrsek L. Should we confirm our clinical diagnostic certainty by autopsies? *Intensive Care Med* 2001;27:1750–5.
- Friedman CP, Gatti GG, Franz TM, Murphy GC, Wolf FM, Hecklering PS, et al. Do physicians know when their diagnoses are correct? Implications for decision support and error reduction. *J Gen Intern Med* 2005;20:334–9.
- Chopra V, Shojania K. Recipe for checklists and bundles – one part active ingredient and one part measurement. *Qual Saf Health Care* 2013;22:93–6.
- Dixon-Woods M, Bosk C, Aveling E, Goeschel C, Pronovost P. Explaining Michigan: developing an ex-post theory of a quality improvement program. *Milbank Q* 2011;89:167–205.
- Bosk CL, Dixon-Woods MS, Goeschel CA, Pronovost P. Reality check for checklists. *Lancet* 2009;374:444–5.

**Supplemental Material:** The online version of this article (DOI: 10.1515/dx-2014-0019) offers supplementary material, available to authorized users.