Checking it twice: an evaluation of checklists for detecting medication errors at the bedside using a chemotherapy model

Rachel E White, Patricia L Trbovich, Anthony C Easty, Pamela Savage, Katherine Trip, Sylvia Hyland

ABSTRACT
Objective To determine what components of a checklist contribute to effective detection of medication errors at the bedside.
Design High-fidelity simulation study of outpatient chemotherapy administration.
Setting Usability laboratory.
Participants Nurses from an outpatient chemotherapy unit, who used two different checklists to identify four categories of medication administration errors.
Main outcome measures Rates of specified types of errors related to medication administration.
Results As few as 0% and as many as 90% of each type of error were detected. Error detection varied as a function of error type and checklist used. Specific step-by-step instructions were more effective than abstract general reminders in helping nurses to detect errors. Adding a specific instruction to check the second individual’s identification improved error detection in this category by 65 percentage points. Matching the sequence of items on the checklist with nurses’ workflow had a positive impact on the ease of use and efficiency of the checklist.
Conclusions Checklists designed with explicit step-by-step instructions are useful for detecting specific errors when a care provider is required to perform a long series of mechanistic tasks under a high cognitive load. Further research is needed to determine how best to assist clinicians in switching between mechanistic tasks and abstract clinical problem solving.

INTRODUCTION
Adverse drug events have long been established as a significant cause of patient harm, and severe incidents with intravenous drug administration have been well documented. Thus, for certain medications, safety groups are increasingly recommending the use of independent double checks for detecting errors.

Independent double checking is a process involving two individuals, in which the responsibility of the second individual is to verify the work performed by the first. The word ‘independent’ refers to having a second person follow a series of steps to arrive at a calculation or setting; those steps are performed with no prior knowledge of any previous calculation or setting. This approach is thought to reduce the possibility of ‘confirmation bias,’ which occurs when the person checking the medication is likely to see what they expect to see, even if an error has occurred.

METHODS
It was essential to begin this research with a solid understanding of the context in which error checking takes place, as well as the specific risks associated with administering chemotherapy by AIP. Hence, using the approach of contextual enquiry, we observed 15 registered nurses in the unit for 50 h. We then identified all failures that could occur and classified these into four categories (table 1).

CHECKLIST DESIGNS
Two checklists were compared in this study. One had been in use in the unit for several months. The new checklist was a revision of the old checklist and was based on our observations of how it was being used.

Old checklist
This checklist had been designed to ensure that the second nurse had no knowledge of the medication
order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.

The second nurse checks the drug label against the order before viewing the pump settings,\textsuperscript{9,15} and was meant to be conducted as follows (figure 1):

The first nurse checks the ‘five rights’ between the order and the drug label, and programmes the pump according to the label. The second nurse copies all values directly from the pump screen to the checklist, and then verifies the checklist against the label.
Errors in patient identification
Detection of identification errors with the new checklist (80%; 16/20) was significantly higher than with the old checklist (15%; 3/20) (p<0.01). Thus, the addition of the specific item on the checklist (ie, ‘check MRN (medical record number) and name from armband to medication label’) had a positive impact on error detection.

Mismatches between order and label
Overall, detection of mismatch errors was low, and there was no significant difference between the old checklist (45%; 9/20) and the new checklist (60%; 12/20) (p>0.05). Since no changes were made to this error detection task on the checklist, we were not surprised by this result.

Clinical decision errors
Neither checklist helped nurses to identify clinical errors (none of these was detected; 0/30). Thus, the addition of the general reminder to stop and think critically had no impact on error detection (p>0.05).

Efficiency
On average, it took nurses 2:16 (minutes:seconds) to complete a check with the old checklist and 1:55 with the new checklist. This 21s improvement was not statistically significant (p>0.05). Thus, despite the addition of a step for verifying patient identity that required the nurse to travel to the bedside, there was no difference in efficiency between the two checklists. Further, nurses commented that the new checklist seemed easier to use.

DISCUSSION
In this study we aimed to determine which checklist features impacted nurses’ ability to detect different types of errors in a double checking process. We found a wide range in error-detection rates: as few as 0% and as many as 90% of errors were identified, depending on the type of error and the checklist used.

SPECIFIC REMINDERS ARE EFFECTIVE BUT NOT FAILSAFE
When the checklist included specific instructions detailing what to look for and where to look (such as the pump programming...
items on both checklists and the patient identification item on the new checklist), detection rates were much higher (80–90%) than when the instructions were less specific. For example, the items for detecting mismatches between the order and label simply instructed nurses to check the medication label against the original order but did not specify which data points to check;
the detection rates for this type of error were quite low (45–60%). The addition of a general reminder to think critically did not help nurses detect any clinical errors. However, adding a specific reminder to check identification from the patient’s armband to the drug label improved error detection by 65 percentage points. Thus, specific checklist items can help clinicians identify well-defined, specific errors. Although this is an encouraging finding, it is important to note that 10–20% of errors still went undetected when they were accompanied by specific instructions, highlighting that human checking processes are not failsafe.

Mechanistic versus abstract tasks
To detect most of the medication errors in this study, nurses had to mechanistically compare data from one source (eg, rate in ml/h on a drug label) against data from another source (eg, rate on an infusion pump in ml/h) to determine if the two matched. In contrast, detection of clinical decision errors required nurses to compare data from one tangible source (eg, dosage in mg from a physician’s order) against their abstract clinical knowledge of chemotherapy protocols, in the context of a specific patient, to determine if all the details of the order were appropriate. Nurses were much better at detecting errors which required mechanistic tasks than those which required critical thought. This difference in performance is consistent with the idea that it is more difficult to detect strategic mistakes than tactical mistakes and that execution errors are more easily detected than method errors.24 25

Further, whereas the use of specific instructions helped nurses with mechanistic error-detection tasks, the addition of a general reminder to think critically and remember the ‘five rights’ of medication administration did not help with the abstract task. Expecting clinicians to mechanistically review several specific aspects of medication administration and then to switch to thinking abstractly about clinical appropriateness may be unrealistic.26 If abstract thinking tasks are essential to the final medication administration process, it may be necessary to separate these from the mechanistic tasks and their associated checklists, and to develop other strategies for supporting abstract clinical thought. Further research is needed in this area.

### Table 2 Errors used in laboratory experiment

<table>
<thead>
<tr>
<th>Error type</th>
<th>Explanation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pump-programming error</td>
<td>First nurse programmed the pump incorrectly, Error in rate, volume or lock level.</td>
<td>Rate has been programmed as 0.3 ml/h. The correct rate is 3 ml/h as shown on order and label.</td>
</tr>
<tr>
<td>Patient-identification error</td>
<td>Two patients in the unit have similar names and/or identifiers.</td>
<td>Order and label are for Ross Kelly, MRN (medical record number) #7004591. Patient wristband reads Kelly Ross, MRN #7004591.</td>
</tr>
<tr>
<td>Mismatch between drug label and order</td>
<td>Physician changed the order after pharmacy prepared the drug, or pharmacy made a transcription error.</td>
<td>Order shows correct dose of 56 mg. Medication label shows 50 mg.</td>
</tr>
<tr>
<td>Clinical decision error</td>
<td>Physician made error during order entry which pharmacist did not detect.</td>
<td>Physician ordered dose 10x too high. Order and label show 1000 mg. Correct clinical value is 100 mg.</td>
</tr>
</tbody>
</table>

### Figure 5 Error-detection rates by error type and checklist type.

Our revised approach to double checking was designed to encourage nurses to review the pump settings before looking at the prescription details, thus ensuring ‘independence’ through the elimination of confirmation bias. We believed this approach would result in higher error-detection rates, but found no difference between the old and new checklists in detection of errors in pump programming. Upon analysis, we realised that the new design did not in fact eliminate confirmation bias; instead, confirmation bias was

### Table 3 Steps for developing a checklist for detecting errors

<table>
<thead>
<tr>
<th>Step</th>
<th>Task</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Determine the errors with high risk or high probability that could reach the bedside, using a technique such as failure modes and effects analysis.18</td>
</tr>
<tr>
<td>2</td>
<td>Develop specific checklist instructions for each predictable error. Include details of what information to check (eg, dose in mg) and from what sources (physician’s order and drug label). Keep the list short17 20 by omitting items with lower risk and lower probability.</td>
</tr>
<tr>
<td>3</td>
<td>If the possibility of an error is abstract or general (eg, error in physician’s dosage choice), but the error itself has a high severity or probability, break the error down into smaller, more specific steps that can be added to the instructions (eg, drug dosage on medication order against hospital drug formulary of appropriate adult doses).</td>
</tr>
<tr>
<td>4</td>
<td>Determine the workflow of the first and second nurses by observing them working in their natural environment using a technique such as contextual enquiry.21 22</td>
</tr>
<tr>
<td>5</td>
<td>To encourage efficiency and adoption, assemble the itemised instructions into a checklist that corresponds with their workflow, and use language and terms that match their existing tools such as the infusion pump screen prompts.</td>
</tr>
<tr>
<td>6</td>
<td>To test and improve the usability of the checklist, recruit a small sample of end users (three to six people) to use the checklist while you observe. If they become confused, use the checklist in a way that is not anticipated, or readily miss errors, refine the design of the form to be more intuitive, and repeat the testing process.</td>
</tr>
<tr>
<td>7</td>
<td>For each potential error not included on the checklist, develop alternate strategies to prevent it from reaching the bedside. Continue to develop additional strategies for eliminating all possible errors, even those that can be identified with the checklist, since no human checking process is failsafe.</td>
</tr>
</tbody>
</table>

---

The detection rates for this type of error were quite low (45–60%). The addition of a general reminder to think critically did not help nurses detect any clinical errors. However, adding a specific reminder to check identification from the patient’s armband to the drug label improved error detection by 65 percentage points. Thus, specific checklist items can help clinicians identify well-defined, specific errors. Although this is an encouraging finding, it is important to note that 10–20% of errors still went undetected when they were accompanied by specific instructions, highlighting that human checking processes are not failsafe.

### Mechanistic versus abstract tasks

To detect most of the medication errors in this study, nurses had to mechanistically compare data from one source (eg, rate in ml/h on a drug label) against data from another source (eg, rate on an infusion pump in ml/h) to determine if the two matched. In contrast, detection of clinical decision errors required nurses to compare data from one tangible source (eg, dosage in mg from a physician’s order) against their abstract clinical knowledge of chemotherapy protocols, in the context of a specific patient, to determine if all the details of the order were appropriate. Nurses were much better at detecting errors which required mechanistic tasks than those which required critical thought. This difference in performance is consistent with the idea that it is more difficult to detect strategic mistakes than tactical mistakes and that execution errors are more easily detected than method errors.24 25

Further, whereas the use of specific instructions helped nurses with mechanistic error-detection tasks, the addition of a general reminder to think critically and remember the ‘five rights’ of medication administration did not help with the abstract task. Expecting clinicians to mechanistically review several specific aspects of medication administration and then to switch to thinking abstractly about clinical appropriateness may be unrealistic.26 If abstract thinking tasks are essential to the final medication administration process, it may be necessary to separate these from the mechanistic tasks and their associated checklists, and to develop other strategies for supporting abstract clinical thought. Further research is needed in this area.

### Error management

The detection rates for this type of error were quite low (45–60%). The addition of a general reminder to think critically did not help nurses detect any clinical errors. However, adding a specific reminder to check identification from the patient’s armband to the drug label improved error detection by 65 percentage points. Thus, specific checklist items can help clinicians identify well-defined, specific errors. Although this is an encouraging finding, it is important to note that 10–20% of errors still went undetected when they were accompanied by specific instructions, highlighting that human checking processes are not failsafe.

### Mechanistic versus abstract tasks

To detect most of the medication errors in this study, nurses had to mechanistically compare data from one source (eg, rate in ml/h on a drug label) against data from another source (eg, rate on an infusion pump in ml/h) to determine if the two matched. In contrast, detection of clinical decision errors required nurses to compare data from one tangible source (eg, dosage in mg from a physician’s order) against their abstract clinical knowledge of chemotherapy protocols, in the context of a specific patient, to determine if all the details of the order were appropriate. Nurses were much better at detecting errors which required mechanistic tasks than those which required critical thought. This difference in performance is consistent with the idea that it is more difficult to detect strategic mistakes than tactical mistakes and that execution errors are more easily detected than method errors.24 25

Further, whereas the use of specific instructions helped nurses with mechanistic error-detection tasks, the addition of a general reminder to think critically and remember the ‘five rights’ of medication administration did not help with the abstract task. Expecting clinicians to mechanistically review several specific aspects of medication administration and then to switch to thinking abstractly about clinical appropriateness may be unrealistic.26 If abstract thinking tasks are essential to the final medication administration process, it may be necessary to separate these from the mechanistic tasks and their associated checklists, and to develop other strategies for supporting abstract clinical thought. Further research is needed in this area.

### Error management

The detection rates for this type of error were quite low (45–60%). The addition of a general reminder to think critically did not help nurses detect any clinical errors. However, adding a specific reminder to check identification from the patient’s armband to the drug label improved error detection by 65 percentage points. Thus, specific checklist items can help clinicians identify well-defined, specific errors. Although this is an encouraging finding, it is important to note that 10–20% of errors still went undetected when they were accompanied by specific instructions, highlighting that human checking processes are not failsafe.

### Mechanistic versus abstract tasks

To detect most of the medication errors in this study, nurses had to mechanistically compare data from one source (eg, rate in ml/h on a drug label) against data from another source (eg, rate on an infusion pump in ml/h) to determine if the two matched. In contrast, detection of clinical decision errors required nurses to compare data from one tangible source (eg, dosage in mg from a physician’s order) against their abstract clinical knowledge of chemotherapy protocols, in the context of a specific patient, to determine if all the details of the order were appropriate. Nurses were much better at detecting errors which required mechanistic tasks than those which required critical thought. This difference in performance is consistent with the idea that it is more difficult to detect strategic mistakes than tactical mistakes and that execution errors are more easily detected than method errors.24 25

Further, whereas the use of specific instructions helped nurses with mechanistic error-detection tasks, the addition of a general reminder to think critically and remember the ‘five rights’ of medication administration did not help with the abstract task. Expecting clinicians to mechanistically review several specific aspects of medication administration and then to switch to thinking abstractly about clinical appropriateness may be unrealistic.26 If abstract thinking tasks are essential to the final medication administration process, it may be necessary to separate these from the mechanistic tasks and their associated checklists, and to develop other strategies for supporting abstract clinical thought. Further research is needed in this area.

### Table 2 Errors used in laboratory experiment

<table>
<thead>
<tr>
<th>Error type</th>
<th>Explanation</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pump-programming error</td>
<td>First nurse programmed the pump incorrectly, Error in rate, volume or lock level.</td>
<td>Rate has been programmed as 0.3 ml/h. The correct rate is 3 ml/h as shown on order and label.</td>
</tr>
<tr>
<td>Patient-identification error</td>
<td>Two patients in the unit have similar names and/or identifiers.</td>
<td>Order and label are for Ross Kelly, MRN (medical record number) #7004591. Patient wristband reads Kelly Ross, MRN #7004591.</td>
</tr>
<tr>
<td>Mismatch between drug label and order</td>
<td>Physician changed the order after pharmacy prepared the drug, or pharmacy made a transcription error.</td>
<td>Order shows correct dose of 56 mg. Medication label shows 50 mg.</td>
</tr>
<tr>
<td>Clinical decision error</td>
<td>Physician made error during order entry which pharmacist did not detect.</td>
<td>Physician ordered dose 10x too high. Order and label show 1000 mg. Correct clinical value is 100 mg.</td>
</tr>
</tbody>
</table>

### Figure 5 Error-detection rates by error type and checklist type.

Our revised approach to double checking was designed to encourage nurses to review the pump settings before looking at the prescription details, thus ensuring ‘independence’ through the elimination of confirmation bias. We believed this approach would result in higher error-detection rates, but found no difference between the old and new checklists in detection of errors in pump programming. Upon analysis, we realised that the new design did not in fact eliminate confirmation bias; instead, confirmation bias was
shifted from the point of checking the pump to the point of checking the newly completed checklist against the prescription details. Hence, the nurses might still have seen what they expected to see on the order when they verified the (sometimes incorrect) values from the checklist against the prescription. More research is needed to determine how to reduce or eliminate confirmation bias from a human verification process.

For double checking processes, we feel that the most important factor is the completion, by a second individual, of a well-designed, easy-to-use checklist with a specific item for each specific high-risk error.

DEVELOPING A CHECKLIST FOR DETECTING ERRORS

On the basis of our findings, we recommend the following steps when developing a checklist for detecting errors (table 3).

**Study limitations**

The clinical focus of this research was the detection by nurses of errors related to administration of chemotherapy and did not include other clinical areas. However, we anticipate that our findings and recommendations will be applicable to any context where a care provider must perform a series of mechanistic tasks under a high cognitive load.

For ethical reasons, we were not able to conduct this controlled experiment of error detection in a live clinical environment. Due to the short-term nature of the simulation, we could not determine the long-term effect of the new checklist on nurses’ behaviour. Further, participants had been using the ‘old’ checklist in their unit for some time. We had expected nurses would detect more errors with the old checklist as a result of their familiarity with it, but this was not the case. Another limitation was that interruptions were not scripted but instead flowed naturally from the actors to create a sense of realism, and therefore were more subject to random error than with a script.

**CONCLUSIONS**

This research has highlighted that checklists incorporating specific step-by-step instructions are useful for detecting certain errors. Although technological systems can achieve a much higher accuracy and reliability than any human processes, fully integrated and automated systems for complex healthcare are not likely to exist in the near future, and we question whether this goal is ultimately achievable. As such, checklists remain a necessary safety tool for clinicians performing long series of mechanistic tasks and must be designed to support this activity. However, further research is needed to determine how best to assist clinicians in switching between mechanistic tasks and abstract problem solving.

**Acknowledgements**

The following people supported this research: H Colbert, A Tosine, D Inczekol, J Stewart, S Ladak, R Lopez, A Chagpar, C Banez, S Savage, D Grosse Wentrup, C Masino, colleagues at the Centre for Global eHealth Innovation and the study participants.

**Funding**

This research was conducted under a research grant from the Canadian Patient Safety Institute (Grant # RF/A0506284).

**Competing interests**

None.

**Ethics approval**

Ethics approval was provided by the University Health Network, Research Ethics Board, Toronto, Ontario, Canada.

**REFERENCES**

Corrections


There are two errors in the results section of this article. The authors state that “the new checklist helped nurses to detect more errors of any type (55%; 71/130) than the old checklist (38%; 49/130)*”. These fractions should not have been included because they are not a logical statistic to report. There were different numbers of planted errors in each category, making the sum of total errors unbalanced: error types which happened to have more planted errors get more weight in the fraction than those with fewer errors. The percentages reported are accurate because the authors took the average error detection percentage across each of the four types- giving them equal weight.

The authors also state that 51/60 errors in pump programming were detected with the old checklist, when it should read 54/60. The percentage value reported was correct (90%).


The authors names were incorrectly cited in this paper. The author list should have been as follows; C Snijders, T W van der Schaaf, H Klip, R A van Lingen, W P F Fetter, A Molendijk.