Reduced Verification of Medication Alerts Increases Prescribing Errors

David Lyell[1]   Farah Magrabi[1]   Enrico Coiera[1]

[1] Centre for Health Informatics, Australian Institute of Health Innovation, Macquarie University, Sydney, New South Wales, Australia

Address for correspondence David Lyell, BA, MBA, Centre for Health Informatics, Australian Institute of Health Innovation, Macquarie University, New South Wales 2109, Australia (e-mail: david.lyell@mq.edu.au).

Abstract

Objective Clinicians using clinical decision support (CDS) to prescribe medications have an obligation to ensure that prescriptions are safe. One option is to verify the safety of prescriptions if there is uncertainty, for example, by using drug references. Supervisory control experiments in aviation and process control have associated errors, with reduced verification arising from overreliance on decision support. However, it is unknown whether this relationship extends to clinical decision-making. Therefore, we examine whether there is a relationship between verification behaviors and prescribing errors, with and without CDS medication alerts, and whether task complexity mediates this.

Methods A total of 120 students in the final 2 years of a medical degree prescribed medicines for patient scenarios using a simulated electronic prescribing system. CDS (correct, incorrect, and no CDS) and task complexity (low and high) were varied. Outcomes were omission (missed prescribing errors) and commission errors (accepted false-positive alerts). Verification measures were access of drug references and view time percentage of task time.

Results Failure to access references for medicines with prescribing errors increased omission errors with no CDS (high-complexity: $\chi^2(1) = 12.716$; $p < 0.001$) and incorrect CDS (Fisher's exact; low-complexity: $p = 0.002$; high-complexity: $p = 0.001$). Failure to access references for false-positive alerts increased commission errors (low-complexity: $\chi^2(1) = 16.673$, $p < 0.001$; high-complexity: $\chi^2(1) = 18.690$, $p < 0.001$). Fewer participants accessed relevant references with incorrect CDS compared with no CDS (McNemar; low-complexity: $p < 0.001$; high-complexity: $p < 0.001$). Lower view time percentages increased omission ($F(3, 361.914) = 4.498$; $p = 0.035$) and commission errors ($F(1, 346.223) = 2.712$; $p = 0.045$). View time percentages were lower in CDS-assisted conditions compared with unassisted conditions ($F(2, 335.743) = 10.443$; $p < 0.001$).

Discussion The presence of CDS reduced verification of prescription safety. When CDS was incorrect, reduced verification was associated with increased prescribing errors.

Conclusion CDS can be incorrect, and verification provides one mechanism to detect errors. System designers need to facilitate verification without increasing workload or eliminating the benefits of correct CDS.

Background and Significance

Prescribing errors are a leading cause of preventable adverse drug events.[1] A common cause of prescribing errors is a lack of knowledge about medicines and the patients for whom they are being prescribed.[2] Clinical decision support (CDS) within electronic prescribing (e-prescribing) systems has been shown to reduce adverse events by alerting clinicians to potential errors such as drug–drug interactions.[3–5] However, CDS is not a perfect substitute for information about medicines: not all potential problems are alerted,[6] malfunctions can occur,[7–9] and alerts are frequently overridden.[10,11]

Verification is the process of establishing the truth or correctness of something by the investigation or evaluation of data.[12] Prescribing errors could be avoided by verification of prescriptions and testing their correctness (safety and appropriateness) against information published in drug references. Inadequate verification is considered an indicator of complacency in overseeing automation, such as decision support.[13–15]

Of specific concern, clinicians may overrely on CDS and consequently reduce their verification efforts, which could lead to errors when CDS is incorrect. This overreliance is known as automation bias and occurs when CDS alerts are used as a "heuristic replacement for vigilant information seeking and processing."[16] Omission errors occur when clinicians fail to address problems because they were not alerted to the problem by CDS, whereas commission errors occur when incorrect CDS advice is acted upon.[16–18] Reduced verification has been associated with automation bias errors in the heavily automated domains of aviation and process control in supervisory control tasks,[13–15,19–22] but it has not yet been tested for CDS medication alerts, where tasks, decision support, and task complexity are likely to differ.[23]

The evidence for higher task complexity increasing automation bias errors is mixed.[17,24–26] However, high-complexity tasks typically have more information to verify[27] and therefore might result in increased reliance on CDS.[23]

While verification could have a key role in reducing prescribing errors, this relationship has not yet been directly studied. Accordingly, this study examines the following: (1) the relationship between verification and prescribing errors with and without CDS medication alerts and (2) whether task complexity mediates this relationship. We are especially interested in the impact of incorrect CDS, which creates the potential for automation bias errors.

Methods

This study presents an analysis of verification data collected as part of a previously reported e-prescribing experiment.[17] An earlier study reported significant evidence of automation bias, with overreliance on incorrect CDS resulting in significantly more errors than when there was no CDS. A second analysis evaluated whether high cognitive load was a cause of automation bias but instead found that participants who made omission errors experienced significantly lower cognitive load than those who did not make errors.[28] This third study extends the prior studies by examining how the presence of CDS and automation bias impact participants' verification and how those changes might contribute to errors.

Participants

The study included students enrolled in the final 2 years of a medical degree at Australian universities, who would typically have received training in rational and safe prescribing and completed the National Prescribing Curriculum, a series of online modules based on the principles outlined in the World Health Organization's Guide to Good Prescribing.[29]

Experiment Design

The analysis had two within-subject factors: quality of CDS (correct, incorrect, and no CDS) and scenario complexity (low and high). The control involved scenarios with no CDS. The original experiment included an interruption condition, which was excluded from this analysis as participants were interrupted while verifying.[17] All participants performed one scenario in each of the six conditions (►Fig. 1).

Outcome Measures

- Omission error (yes/no): participants made an *omission error* if they prescribed a designated medication containing a prescribing error, indicating that they had failed to detect it. If the error was corrected, it was not scored as an error.
- Commission error (yes/no): participants made a *commission error* if they wrongly acted on a false-positive alert by not prescribing a medication that was unaffected by prescribing errors.

Verification Measures

- Access (accessed/not accessed): whether the participant accessed the drug reference for the medicine with the prescribing error (omission error) or the medicine triggering the false-positive alert (commission error).
- View time percentage: the percentage of task time viewing drug references. The conversion to a percentage of task time allowed for comparisons between low- and high-complexity conditions, which differed in the number of prescription requested. High-complexity scenarios requested five more prescriptions than low-complexity scenarios. Task and drug reference view time were expected to increase as a function of the number of prescriptions requested.

Experimental Task

Participants were provided with patient scenarios presenting a brief patient history and a list of medications for them to prescribe using a simulated e-prescribing system (►Fig. 2). One of the listed medicines was contraindicated, posing a sufficiently severe risk of harm to the patient that its use should be avoided. All other requested medication orders were unaffected by prescribing errors. Participants were instructed to prescribe all medications except those they believed contain a prescribing error. Of interest was whether participants would detect the prescribing error. See the appendices in the study by Lyell et al[28] for examples of the patient scenarios and a summary of the errors inserted in the scenarios.
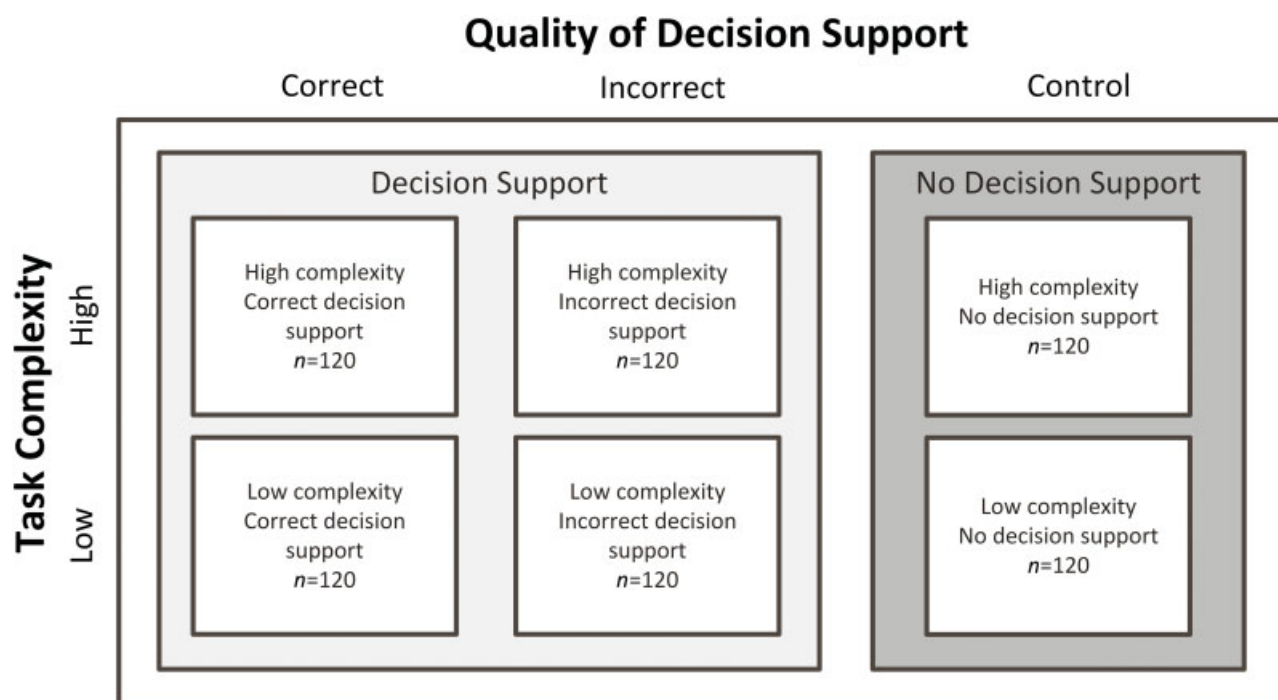
Quality of Decision Support

Fig. 1 Experimental design with the number of participants in each condition. (Adapted from Lyell et al[17] and reproduced under CC BY 4.0.)

Verification of Prescriptions

Participants were able to verify the safety of prescriptions independently of CDS and the correctness of CDS by accessing a drug reference viewer built into the e-prescribing system. The drug reference was easily accessible and displayed monographs from the Australian Medicines Handbook,[30] an evidence-based reference widely used in Australian clinical practice.[31] Participants were instructed the following: (1) CDS could be incorrect, (2) how to verify using the drug reference, (3) rely on the drug reference over CDS if there was a discrepancy, and (4) refer only to the provided drug reference.

Drug references were checked by M. Z. R. (a pharmacist) and D. L. to ensure that they provided clear and sufficient information to enable prescribing errors to be identified. A log recorded access to drug references and view times.

Fig. 2 Example of the experimental task showing the e-prescribing system (left) and patient scenario (right). (Adapted from Lyell et al[17] and reproduced under CC BY 4.0.). No personally identifying information was displayed to participants or reported in this article. The patients presented in the prescribing scenarios were fictional. The biographical information was made up for this experiment in order to present participants with the information they would expect in such patient cases.

Fig. 3 Clinical decision support medication alert. (Adapted from Lyell et al[17] and reproduced under CC BY 4.0.)

Clinical Decision Support Alerts

CDS displayed alerts (►**Fig. 3**) when a medication order containing a prescribing error was entered and required resolution either by removing the prescription or by overriding the alert with a reason. For examples of the override reasons provided by participants, see Lyell et al.[17]

The triggering and content of CDS alerts were manipulated across the following three conditions:

- *Correct CDS* alerts were triggered by prescription of the medication with the prescribing error (true-positives). The absence of alerts always indicated true-negatives.
- *Incorrect CDS* failed to alert the prescribing error (false-negative) and instead provided one false-positive alert for a medicine unaffected by prescribing error. These CDS errors provided opportunities for one omission error and one commission error.
- *No CDS* served as the control condition in which there were no CDS checking for errors. Participants were told that CDS had been switched off for these scenarios and were advised to use the drug reference to manage any errors.

Task complexity was manipulated by varying the number of prescriptions requested and information elements in scenarios.[32,33] Low-complexity scenarios requested three prescriptions and contained three additional information elements such as medical conditions, symptoms, test results, allergies, and observations that could potentially contra-indicate those medications. High-complexity scenarios requested eight medications and contained nine additional information elements. As a result, high-complexity scenarios had five more drug references that could be viewed, had more information elements to be cross-referenced, and required more verification than low-complexity scenarios. We previously reported that participants found high-complexity scenarios significantly more cognitively demanding than low-complexity scenarios.[28]

Allocation of patient scenarios to experimental conditions was counterbalanced to ensure that scenarios were evenly presented in all conditions. The order of presentation was randomized to control for order effects.

Procedure

The experiment was presented as an evaluation of an e-prescribing system in development. No information was provided on what types of errors the system would check and alert. Participants were shown an instructional video on how to use the e-prescribing system, including demonstration of a correct CDS alert, and how to verify using the drug reference.

Participants were instructed to approach tasks as if treating a real patient, exercising all due care, and not prescribing any medication believed to contain a prescribing error.

Statistical Analyses

Chi-square test for independence and Fisher's exact probability tests were used to test whether access of drug references relevant to errors was associated with omission and commission errors. Differences in access between CDS conditions and levels of task complexity were tested using McNemar's tests.

Multilevel modelling,[34] which is not affected by missing data,[35] was used to analyze view time percentage as participants did not access drug references in all conditions. The predictors assessed for inclusion in the model were task complexity, quality of decision support, and whether the participant made an omission error and commission error. We assessed all two-way interactions. A stepwise backward elimination method was used for predictor selection, where all predictors were entered into the model, and then interactions were removed one by one in order of least significance. The process was repeated for main effects. Model fit was evaluated by comparing models using the likelihood ratio test.[36] Only predictors with a significant effect on model fit were retained. The model included a random intercept for each participant, taking into account the nested structure of the data. Models were constructed using maximum likelihood for parameter estimation.

Results

A total of 120 participants were included in the analysis. One participant completed the experiment twice (on two separate occasions), and the data from their second attempt were excluded. Participants' average age was 24 years, and 46.7% were female. The median time to perform was 2:45 minutes (interquartile range = 1:42 to 4:08) for low-complexity scenarios and 5:25 minutes (interquartile range = 3:59 to 7:21) for high-complexity scenarios. Overall, participants accessed the drug information reference at least once in 64.7% of scenarios. Thirty-four participants viewed at least one reference in all scenarios, whereas 11 participants did not view any references (accounting for 25.9% of the scenarios in which no references were viewed).

Accessing Drug References for Medicines with Prescribing Errors

Omission Errors Were Higher When Drug References for Medicines with Prescribing Errors Were Not Accessed
When prescribing without CDS (control), omission errors were higher when drug references for medicines with prescribing errors were not accessed (►**Table 1**). This was

Table 1 Percentage (number) of participants who accessed the drug reference for medicines with prescribing errors by whether an omission error was made

	Control (no CDS)			Correct CDS			Incorrect CDS			Total
	No error	Error	Total	No error	Error	Total	No error	Error	Total	
Low complexity										
Accessed	59.7%	40.3%	51.7% (62)	94.1%	5.9%	51.7% (62)	47.6%	52.4%	42.5% (51)	17.5% (21)
Not accessed	48.3%	51.7%	48.3% (58)	91.3%	8.7%	48.3% (58)	15.2%	84.8%	57.5% (69)	82.5% (99)
Total	54.2% (65)	45.8% (55)		92.5% (111)	7.5% (9)		20.8% (25)	79.2% (95)		
High complexity										
Accessed	70.6%	29.4%	42.5% (51)	90.9%	9.1%	42.5% (51)	62.5%	37.5%	45.8% (55)	13.3% (16)
Not accessed	37.7%	62.3%	57.5% (69)	90.8%	9.2%	57.5% (69)	19.2%	80.8%	54.2% (65)	86.7% (104)
Total	51.7% (62)	48.3% (58)		90.8% (109)	9.2% (11)		25% (30)	75% (90)		
Total										
Accessed	64.6%	35.4%	47.1% (113)	92.5%	7.5%	47.1% (113)	54.1%	45.9%	44.2% (106)	15.4% (37)
Not accessed	42.5%	57.5%	52.9% (127)	91.0%	9.0%	52.9% (127)	17.2%	82.8%	55.8% (134)	84.6% (203)
Total	52.9% (127)	47.1% (113)		91.7% (220)	8.3% (20)		22.9% (55)	77.1% (185)		

Abbreviation: CDS, clinical decision support.

significant for high-complexity scenarios (χ^2 (1, $n = 120$) = 12.716; $p < 0.001$; $\phi = -0.326$) but not for low-complexity scenarios (χ^2 (1, $n = 120$) = 1.569; $p = 0.210$).

A similar relationship was found with incorrect CDS, which failed to alert the prescribing error. Omission errors were significantly higher when the drug reference for the medicine with the prescribing error was not accessed in both low-complexity (Fisher's exact test; $p = 0.002$; $n = 120$) and high-complexity conditions (Fisher's exact test; $p = 0.001$; $n = 120$).

For correct CDS, there was no relationship between accessing the relevant drug reference and omission errors, as would be expected for correctly alerted prescribing errors (►Table 1; Fisher's exact tests: low complexity, $p = 0.731$, $n = 120$; high complexity, $p = 1$, $n = 120$).

Across all conditions, 35% of participants in the control and 46% of participants in the incorrect CDS conditions made omission errors despite accessing the reference necessary to identify the error.

Clinical Decision Support Reduced Participants' Access of Drug References for Medicines with Prescribing Errors

Significantly fewer participants accessed drug references for medicines containing prescribing errors with incorrect CDS compared with no CDS (control; McNemar's tests: low complexity, $p < 0.001$, $n = 120$; high complexity, $p < 0.001$, $n = 120$). However, there was no difference in access between correct and no CDS (control; McNemar's tests: low complexity, $p = 0.169$, $n = 120$; high complexity, $p = 0.665$, $n = 120$).

Commission Errors Were Higher When Drug References Relevant to False-Positive Alerts Were Not Accessed

False-positive alerts were more likely to lead to commission errors if the drug reference for the medicine triggering the alert was not accessed (►Table 2; low complexity, χ^2 (1, $n = 116$) = 16.673, $p < 0.001$, $\phi = -0.379$; high complexity, χ^2 (1, $n = 111$) = 18.690, $p < 0.001$, $\phi = -0.410$). Even when the relevant reference was consulted, 45.9% of participants across all conditions went on to make a commission error despite accessing references contradicting the alert.

Task Complexity Did Not Affect Access of Drug References Relevant to Errors

There was no difference in the proportion of participants who accessed drug references for medicines with prescribing errors (opportunities for omission errors) between the low- and high-complexity scenarios (McNemar's tests: control, $p = 0.071$, $n = 120$; correct CDS, $p = 0.665$, $n = 120$; incorrect CDS, $p = 0.405$, $n = 120$). Similarly, there was no difference in participants accessing drug references relevant to false-positive alerts (opportunities for commission errors) between the low- and high-complexity scenarios (McNemar's test: incorrect CDS, $p = 0.117$, $n = 108$).

Multilevel Analysis of View Time Percentages

The multilevel analysis focused on the 466 scenarios (64.7%) in which drug references were accessed. View time

Table 2 Percentage (number) of participants who accessed the drug reference relevant to the false-positive alert from incorrect CDS by whether a commission error was made

	No error	Commission error	Total
Low complexity			
Accessed	48.4%	51.6%	53.4% (62)
Not accessed	13%	87%	46.6% (54)
Total	31.9% (37)	68.1% (79)	
High complexity			
Accessed	61.2%	38.8%	44.1% (49)
Not accessed	21%	79%	55.9% (62)
Total	38.7% (43)	61.3% (68)	
Total			
Accessed	54.1%	45.9%	48.9% (111)
Not accessed	17.2%	82.8%	51.1% (116)
Total	35.2% (80)	64.8% (147)	

Abbreviation: CDS, clinical decision support.
Note: includes only scenarios in which false-positive alerts were displayed.

percentage could not be calculated in 100 scenarios (21.5% of these), where task time was not recorded due to a software issue ($n = 93$), outliers for task time ($n = 9$) and view time ($n = 1$) were removed, or view time data was missing ($n = 6$). Several scenarios were affected by multiple issues. View time percentage was calculated for the remaining 366 scenarios (78.5%) and included in the model. With no systematic differences detected in the missing data, they were treated as being random.

Thirteen models were evaluated (►**Supplementary Appendix A**, available in the online version), and from these, four fixed effects were found to significantly contribute to the fit of a multilevel model and were included in the final model. The significance of fixed effects (predictors in the model) is given in ►**Table 3**, and the model coefficients are presented in ►**Supplementary Appendix B** (available in the online version). The comparison of effects is reported based on the estimated marginal means computed by the model. Significance probabilities have been adjusted for multiple comparisons using the Bonferroni correction.[37] The final model was significantly better than the intercept only model ($\chi^2(7) = 132.867$; $p < 0.001$). The intraclass correlation coefficient was

0.23, indicating that 23% of the variance in verification was attributable to variation between participants, supporting the conduct of a multilevel analysis.[38,39] The model residuals were normally distributed.

Participants who made omission errors spent significantly smaller percentage of task time viewing drug references ($M = 24.7\%$; 95% confidence interval [CI] [21.1%, 28.2%]) than those who did not make errors ($M = 28.4\%$; 95% CI [25.1%, 31.6%]).

Similarly, participants who made commission errors with incorrect CDS spent significantly smaller percentage of task time viewing drug references ($p = 0.018$; $M = 23.6\%$; 95% CI [20%, 27.2%]) than those who made no errors ($M = 29.7\%$; 95% CI [25.6%, 33.6%]). This interaction occurs because only the incorrect CDS conditions displayed false-positive alerts that provided an opportunity for commission errors. There were no differences in the correct CDS ($p = 0.977$) or control ($p = 0.120$) conditions.

View time percentage was significantly reduced by the provision of decision support (►**Fig. 4**). View time percentage was highest in the control condition, which provided no decision support ($M = 34\%$; 95% CI [29.7%, 39.9%]), and this was significantly higher than correct CDS ($p < 0.001$;

Table 3 Significance of fixed effects in the multilevel model of view time percentage

	df	F	p-Value
Intercept	1, 244.483	317.245	<0.001[a]
Task complexity: low complexity, high complexity	1, 302.436	105.383	<0.001[a]
Quality of decision support: correct CDS, incorrect CDS, control (No CDS)	2, 335.743	10.443	<0.001[a]
Omission error: omission error, no omission error	1, 361.914	4.498	0.035[a]
Quality of decision support * commission error	3, 346.223	2.712	0.045[a]

Abbreviations: CDS, clinical decision support; df, degrees of freedom.
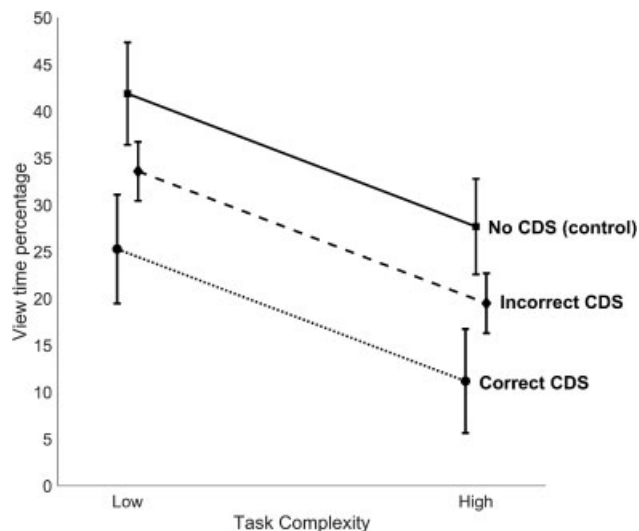[a]Indicates significant effect ($p < 0.05$).

Fig. 4 Estimated marginal means with 95% confidence interval (from the multilevel model) for view time percentage by task complexity and quality of decision support. CDS, clinical decision support.

$M = 18.2\%$; 95% CI [12.7%, 23.8%]) and incorrect CDS ($p = 0.012$; $M = 26.6\%$; 95% CI [23.7%, 29.5%]).

High task complexity significantly reduced view time percentage. Participants spent a significantly greater percentage of task time viewing drug references in low-complexity scenarios ($M = 33.6\%$; 95% CI [30.2%, 37%]) compared with high-complexity scenarios ($M = 19.5\%$; 95% CI [16.4%, 22.5%]).

Discussion

This experiment demonstrates, first, that decreased verification, manifesting as either failure to access references or reduced view times as a percentage of task time, leads to increased omission and commission errors. Second, the presence of CDS decreases verification and that decreased verification leads to increased omission and commission errors when CDS is incorrect.

We found that omission and commission errors increased when participants did not access relevant references (►Fig. 5). Troublingly, some participants went on to make omission and commission errors despite accessing references containing information necessary to avoid those errors. Prior studies have reported a similar "looking-but-not-seeing" or "inattentional blindness,"[13,15,21] which describes how people may fail to perceive something in plain sight because they are not attending to it.[40] Consequently, accessing the relevant references did not guarantee errors were detected, but failure to do so made errors more likely.

Seeking further insight into why accessing relevant references avoided some but not all errors, we analyzed view time percentages. We found that participants who avoided errors spent a significantly greater percentage of task time viewing references than those who made errors. Together the access and view time percentages results suggest the following: (1) verification should not be viewed as all-or-nothing but rather on a continuum of adequacy or vigilance and (2) greater verification can reduce both omission and commission errors.
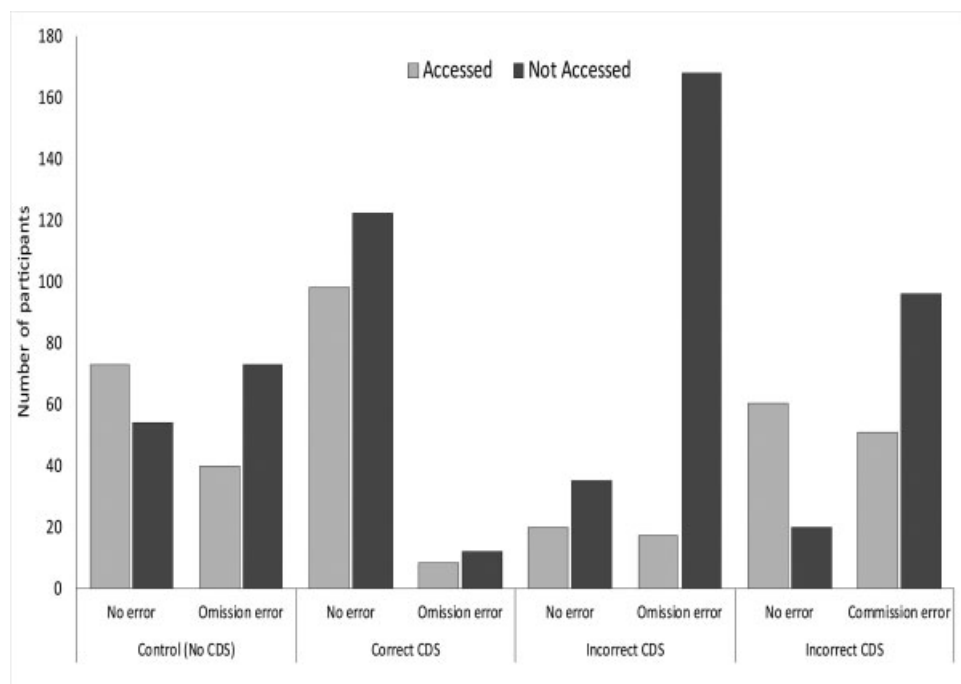

Fig. 5 The number of participants who made errors by quality of clinical decision support and whether the relevant drug reference was accessed. Summarizes the data presented in ►Tables 1 and 2, aggregating the low- and high-complexity conditions. CDS, clinical decision support.

Clinical Decision Support Reduced Verification

We reported finding evidence of automation bias in this experiment; participants made significantly more omission and commission errors when provided with incorrect CDS compared with when they had no CDS.[17] The risk posed by automation bias is that CDS becomes a replacement for, rather than a supplement to, clinicians' efforts in error detection. The analysis of verification behavior provides some support for the idea of CDS replacing participants' error detection efforts. A significantly smaller percentage of task time was spent viewing references in CDS-assisted compared with unassisted conditions (see ►Fig. 4). This reduction in verification was associated with increased errors. It is very likely that this relationship is causal, with reduced verification impeding the discovery of errors.

Furthermore, when CDS was incorrect, participants who made omission or commission errors spent a smaller percentage of task time viewing references than those who did not make errors. This is consistent with prior automation bias research, which mostly employed aviation and process control tasks.[13–15,19–22] This study confirms that this association extends to the detection of prescribing errors assisted by CDS medication alerts.

Manzey et al[13] suggest that the looking-but-not-seeing effect, whereby participants made errors despite viewing information that could have prevented them, represents an automation bias induced withdrawal of cognitive resources for processing verification information. Therefore, while the necessary information was accessed, it was not processed in a way that enabled errors to be recognized. Our analysis of participants' cognitive load, reported separately, provides support for this. Participants who made omission errors allocated fewer cognitive resources to the task than those who did not.[28] Curiously, there was no difference for commission errors. The present findings suggest that in addition to reduced processing, there may also be reduced acquisition of information.

This is consistent with a *cognitive miser* view of automation bias[16,28] that people prefer adequate, faster, and less effortful ways of thinking rather than engaging in more accurate but slower and more effortful thinking.[41] These findings also support Mosier and Skitka's description of automation bias as the use of automation as a heuristic,[16] with CDS appearing to be used as a shortcut in place of verification.

The same cognitive miser profile could also be found in participants who made errors in the control condition but to a significantly lesser extent. This may indicate the presence of other factors that trigger reduced verification in addition to automation bias.

Less Verification in High Complexity

High-complexity scenarios asked participants to prescribe five more medications, just over two and a half times the number requested in low-complexity scenarios. We expected that the time to enter prescriptions into the e-prescribing system would increase as a function of the number of medications prescribed. Likewise, drug reference view time was expected to increase with the number of prescriptions and drug references that could be viewed. While there were no differences in access of relevant drug references as complexity increased from low to high, the view time percentage was significantly lower. The reduction in the percentage of task time viewing references could represent participants' efforts to manage the increased workload created by needing to verify more information in high-complexity scenarios. Despite this, we have previously reported that high task complexity did not increase automation bias errors.[17] This is puzzling, especially in light of present findings that high task complexity reduced verification, suggesting that it may be a risk factor for automation bias. It is possible that participants' verification efforts were more sensitive to task complexity than errors, with both low- and high-complexity conditions exhibiting automation bias errors to a similar extent. If task complexity is a risk factor for automation bias, then both complexity conditions likely exceeded the threshold at which it presents. More research is needed to fully understand the relationship between task complexity and errors.

Implications

These findings highlight the importance of verification in preventing prescribing errors and may be generalizable to other forms of CDS. When prescribing is assisted by CDS medication alerts, verification provides the crucial means to differentiate between correct and incorrect CDS. However, the very presence of CDS is likely to exacerbate the problem, contributing to decreased verification, which, in turn, impedes the discovery of errors when CDS fails. This is the risk and challenge of automation bias. High task complexity further complicated matters, appearing to place downward pressure on verification, although the link between complexity and errors remains unclear. Improving the reliability and accuracy of CDS can reduce opportunities for error. However high-reliability automation is known to increase the rate of automation bias,[25] which, in turn, risks clinicians being less able to detect CDS failures when they occur.

The challenge for designers and users of CDS is to ensure appropriate verification in circumstances that may promote decreased verification. To date, automation bias has proven stubbornly resistant to attempts to mitigate its effects,[23] including interventions that prompted users to verify.[42]

While our findings describe how CDS changed the access of references and view time percentages, little is known about what factors prompt clinicians to verify, the information sought and how they go about verifying, including the assessment of information and resolution of potential conflicts between different information sources. More research is needed in this area and how to best assist clinicians with effective verification. Such efforts need to focus on how to best incorporate verification information into workflows, presenting only relevant information when, where, and in the form it is needed. The challenge is to do this in a way that minimally impacts workload, does not overwhelm clinicians with too much information, and maximizes efficiency when CDS is correct.

Ultimately, clinicians need to be mindful that CDS can and does fail,[7–9] and when it does, verification is the primary means to avoid errors. While it is impractical and undesirable to verify all prescriptions, clinicians would be well advised to verify whenever they suspect medication safety issues, even in the absence of medication alerts. It would also be prudent when prescribing unfamiliar or little-used medicines or for unfamiliar issues.

Limitations

This experiment was subject to several limitations. The use of medical students provided a necessary control for knowledge and experience of prescribing. This provides an indication of verification behavior by junior medical officers entering practice but may have limited generalizability to more experienced clinicians. Clinician knowledge is likely to play an important role in verification but exceeds the scope of this study. Likewise, the completeness of knowledge will also be an important consideration, for example, a clinician may know a medicine's contraindications for conditions but not know all its possible adverse drug interactions.

Replication of our study with other cohorts, including more experienced clinicians, and clinicians operating in different clinical contexts would need to be undertaken. The evidence for the presence of similar verification results in other nonhealth care settings[13–15,19–22] suggests, however, that these results are indeed generalizable to clinical decision-making assisted by CDS.

Other factors that are likely to impact verification include the design and accuracy of CDS and the accessibility of verification information. Further research identifying the relative contributions of such factors would be informative for developing mitigations.

Participants were not subjected to experimentally imposed time constraints or required to manage competing demands for their attention that clinicians would ordinarily experience in clinical practice.

Finally, the inclusion of conditions designed to elicit both omission and commission errors in the same condition means that we cannot fully differentiate the effects of verification for each error type.

Conclusion

This is the first study to test the relationship between verification behaviors and the detection of prescribing errors, with and without CDS medication alerts. Increased verification was associated with increased detection of errors, whereas the presence of CDS and high task complexity reduced verification.

These findings demonstrate the importance of verification in avoiding prescribing and automation bias errors. CDS can alert clinicians to errors that may have been inadvertently missed; however, they are not perfectly sensitive and specific. Clinicians should allow CDS to function as an additional layer of defense but should not rely on it if they suspect a medication safety issue as it cannot replace the clinician's own expertise and clinical judgment.

Clinical Relevance Statement

Verification of CDS provides one means to avoid prescribing errors and is especially prudent when prescribing unfamiliar or little-used medicines or for unfamiliar issues. CDS medication alerts can help prevent prescribing errors, but CDS is imperfect and can be incorrect. The presence of CDS appears to reduce verification efforts, and when CDS is incorrect, reduced verification is associated with prescribing errors.

Multiple Choice Questions

1. What strategy can be used to reduce prescribing errors when using CDS medication alerts?
   a. Improve the accuracy of CDS medication alerts.
   b. Verifying medication alerts, or their absence, with a gold standard, evidence-based drug reference.
   c. Introduce messages into CDS systems that prompt clinicians to verify prescriptions.
   d. Phase out CDS medication alerts.

   Correct Answer: The correct answer is option b. Our results found that when CDS was incorrect, greater verification was associated with reduced prescribing errors. CDS medication alerts have been shown to reduce prescribing errors (not option d), but they introduce a risk of overreliance. While improving CDS accuracy would reduce opportunities for errors from overreliance, perfectly sensitive and specific CDS is likely unattainable. Additionally, highly accurate decision support increases the rate of automation bias errors (not option a). Automation bias has proven stubbornly resistant to mitigations including prompting users to verify (not option c).

2. When is verification of CDS medication alerts, or their absence, especially prudent?
   a. When prescribing unfamiliar or little-used medicines.
   b. When prescribing for unfamiliar problems.
   c. When a medication safety issue, such as contraindication, is suspected.
   d. All of the above.

   Correct Answer: The correct answer is option d. Clinicians will be familiar with and have a good knowledge of the medicines they frequently prescribe for commonly encountered issues. However, when prescribing unfamiliar or little-used medicines or prescribing for unfamiliar issues, clinicians may have gaps in knowledge and rely more heavily on CDS. If CDS is incorrect, there is a risk of omission or commission errors occurring. In general, it is prudent for clinicians to verify computer-generated alerts, or their absence, if they suspect a risk of a prescribing error.

Authors' Contributions

David Lyell conceived this research and designed and conducted the study with guidance from and under the supervision of Enrico Coiera and Farah Magrabi. David

References

1 Thomsen LA, Winterstein AG, Søndergaard B, Haugbølle LS, Melander A. Systematic review of the incidence and characteristics of preventable adverse drug events in ambulatory care. Ann Pharmacother 2007;41(09):1411–1426

2 Tully MP, Ashcroft DM, Dornan T, Lewis PJ, Taylor D, Wass V. The causes of and factors associated with prescribing errors in hospital inpatients: a systematic review. Drug Saf 2009;32(10):819–836

3 Wolfstadt JI, Gurwitz JH, Field TS, et al. The effect of computerized physician order entry with clinical decision support on the rates of adverse drug events: a systematic review. J Gen Intern Med 2008;23(04):451–458

4 Ammenwerth E, Schnell-Inderst P, Machan C, Siebert U. The effect of electronic prescribing on medication errors and adverse drug events: a systematic review. J Am Med Inform Assoc 2008;15(05): 585–600

5 van Rosse F, Maat B, Rademaker CMA, van Vught AJ, Egberts AC, Bollen CW. The effect of computerized physician order entry on medication prescription errors and clinical outcome in pediatric and intensive care: a systematic review. Pediatrics 2009;123(04): 1184–1190

6 Sweidan M, Williamson M, Reeve JF, et al. Evaluation of features to support safety and quality in general practice clinical software. BMC Med Inform Decis Mak 2011;11(01):1–8

7 Wright A, Ai A, Ash J, et al. Clinical decision support alert malfunctions: analysis and empirically derived taxonomy. J Am Med Inform Assoc 2018;25(05):496–506

8 Wright A, Hickman TT, McEvoy D, et al. Analysis of clinical decision support system malfunctions: a case series and survey. J Am Med Inform Assoc 2016;23(06):1068–1076

9 Kassakian SZ, Yackel TR, Gorman PN, Dorr DA. Clinical decisions support malfunctions in a commercial electronic health record. Appl Clin Inform 2017;8(03):910–923

10 van der Sijs H, Aarts J, Vulto A, Berg M. Overriding of drug safety alerts in computerized physician order entry. J Am Med Inform Assoc 2006;13(02):138–147

11 Nanji KC, Slight SP, Seger DL, et al. Overrides of medication-related clinical decision support alerts in outpatients. J Am Med Inform Assoc 2014;21(03):487–491

12 Oxford English Dictionary. verification, n. Oxford University Press; 2018

13 Manzey D, Reichenbach J, Onnasch L. Human performance consequences of automated decision aids: the impact of degree of automation and system experience. J Cogn Eng Decis Mak 2012;6 (01):57–87

14 Bahner J, Huper AD, Manzey D. Misuse of automated decision aids: complacency, automation bias and the impact of training experience. Int J Hum Comput Stud 2008;66(09):688–699

15 Bahner J, Elepfandt MF, Manzey D. Misuse of diagnostic aids in process control: the effects of automation misses on complacency and automation bias. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, September 22–26, 2008, New York, NY

16 Mosier KL, Skitka LJ. Human decision makers and automated decision aids: made for each other. In: Parasuraman R, Mouloua M, eds. Automation and Human Performance: Theory and Applications. Hillsdale, NJ: Lawrence Erlbaum Associates; 1996:201–220

17 Lyell D, Magrabi F, Raban MZ, et al. Automation bias in electronic prescribing. BMC Med Inform Decis Mak 2017;17(01):28

18 Mosier KL, Skitka LJ, Heers S, Burdick M. Automation bias: decision making and performance in high-tech cockpits. Int J Aviat Psychol 1997;8(01):47–63

19 Bagheri N, Jamieson GA. The impact of context-related reliability on automation failure detection and scanning behaviour. Paper presented at the IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No.04CH37583), October 10–13, 2004

20 Bagheri N, Jamieson GA. Considering subjective trust and monitoring behavior in assessing automation-induced "complacency.". In: Vincenzi DA, Mouloua M, Hancock PA, eds. Human Performance, Situation Awareness and Automation: Current Research and Trends. Vol. 2. Mahwah, NJ: Lawrence Erlbaum Associates; 2004:54–59

21 Reichenbach J, Onnasch L, Manzey D. Misuse of automation: the impact of system experience on complacency and automation bias in interaction with automated aids. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, September 27 to October 1, 2010, San Francisco, CA

22 Reichenbach J, Onnasch L, Manzey D. Human performance consequences of automated decision aids in states of sleep loss. Hum Factors 2011;53(06):717–728

23 Lyell D, Coiera E. Automation bias and verification complexity: a systematic review. J Am Med Inform Assoc 2017;24(02): 423–431

24 Goddard K, Roudsari A, Wyatt JC. Automation bias: a systematic review of frequency, effect mediators, and mitigators. J Am Med Inform Assoc 2012;19(01):121–127

25 Bailey NR, Scerbo MW. Automation-induced complacency for monitoring highly reliable systems: the role of task complexity, system experience, and operator trust. Theor Issues Ergon Sci 2007;8(04):321–348

26 Povyakalo AA, Alberdi E, Strigini L, Ayton P. How to discriminate between computer-aided and computer-hindered decisions: a case study in mammography. Med Decis Making 2013;33(01): 98–107

27 Sweller J. Element interactivity and intrinsic, extraneous, and germane cognitive load. Educ Psychol Rev 2010;22(02):123–138

28 Lyell D, Magrabi F, Coiera E. The effect of cognitive load and task complexity on automation bias in electronic prescribing. Hum Factors 2018;60(07):1008–1021

29 De Vries TPGM, Henning RH, Hogerzeil HV, Fresle DA. Guide to Good Prescribing. Geneva: World Health Organization; 1994

30 Australian Medicines Handbook Pty Ltd. Australian Medicines Handbook 2015. Available at: http://amhonline.amh.net.au/

31 Day RO, Snowden L. Where to find information about drugs. Aust Prescr 2016;39(03):88–95

32 Sweller J. Cognitive load theory, learning difficulty, and instructional design. Learn Instr 1994;4(04):295–312

33 Sweller J, Chandler P. Why some material is difficult to learn. Cogn Instr 1994;12(03):185–233

34 Hoffman L, Rovine MJ. Multilevel models for the experimental psychologist: foundations and illustrative examples. Behav Res Methods 2007;39(01):101–117

35 Tabachnick BG. Using Multivariate Statistics. In: Tabachnick BG, Fidel LS 6th ed. Boston, MA: Pearson; 2013

36 Peugh JL. A practical guide to multilevel modeling. J Sch Psychol 2010;48(01):85–112

37 Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. BMJ 1995;310(6973):170

38 Twisk J. Applied Multilevel Analysis: A Practical Guide. Cambridge, UK: Cambridge University Press; 2006

39 Hayes AF. A primer on multilevel modeling. Hum Commun Res 2006;32(04):385–410

40 Mack A, Rock I. Inattentional Blindness. Cambridge, MA: MIT Press; 1998

41 Fiske ST, Taylor SE. Social cognition. New York, NY: Random House; 1984

42 Mosier KL, Skitka LJ, Dunbar M, McDonnell L. Aircrews and automation bias: the advantages of teamwork? Int J Aviat Psychol 2001;11(01):1–14