Does Tall Man lettering prevent drug name confusion errors? Incomplete and conflicting evidence suggest need for definitive study

Bruce L Lambert,[1] Scott R Schroeder,[2] William L Galanter[3]

[1]Department of Communication Studies, Center for Communication and Health, Northwestern University, Chicago, Illinois, USA
[2]Center for Communication and Health, Northwestern University, Chicago, Illinois, USA
[3]Department of Medicine, University of Illinois at Chicago, Chicago, Illinois, USA

Correspondence to
Professor Bruce L Lambert, Department of Communication Studies & Center for Communication and Health, Northwestern University, 710 Lake Shore Drive, 15th Floor, Chicago, IL 60611, USA; bruce.lambert@northwestern.edu

► http://dx.doi.org/10.1136/bmjqs-2015-004562

CrossMark

Wrong-drug errors, thought to be caused primarily by drug names that look and/or sound alike, occur at a rate of about one error per thousand dispensed prescriptions in the outpatient setting and one per thousand orders in the inpatient setting.[1 2] Most are relatively benign, but some cause severe or even fatal harm.[3–5] One of the best known attempts to reduce drug name confusion has been the use of mixed case or 'Tall Man' lettering.[6] The idea is to use capital letters to maximise the visual perceptual difference between two similar drug names. Thus, vinblastine and vincristine become vinBLAStine and vinCRIStine. If some look-alike/sound-alike (LASA) mix-ups are caused by errors in visual perception, the reasoning goes, then making the names more visually distinct should reduce the probability of confusion and error.

After being endorsed by the US Food and Drug Administration (FDA),[6] the Institute for Safe Medication Practices (ISMP),[7] The Joint Commission[8] and others, the practice has become widespread.[9] However, apart from limited evidence of effectiveness in laboratory settings, no evidence shows that this technique prevents drug name confusion errors in clinical practice. Zhong *et al*[10] attempted to assess the effect of Tall Man lettering on drug name confusion errors in a large scale, longitudinal, observational study. They conclude that this widely disseminated error-prevention strategy had no measurable effect on the rate of drug name confusions in 9 years of data from 42 children's hospitals in the USA. Below we comment on methodological issues in the Zhong *et al* study, review laboratory research on Tall Man lettering and consider policy implications.

METHODOLOGICAL ISSUES
The authors are to be commended for conducting a large-scale, empirical test of the effect of Tall Man lettering on the drug name confusion error rate in real-world clinical settings. The paper has many strengths, including its use of an interrupted time series design to test the main hypothesis. The scale of the paper is impressive, covering 1.6 million orders from 42 hospitals over 9 years (2004–2012). Still, there are limitations to be noted, beyond those raised by the authors.

The most significant limitation is that it is unknown at what point each of the 42 study hospitals implemented Tall Man (if at all), and for which pairs of names. These sites likely began the study period with paper-based ordering and moved to computerised provider order entry (CPOE), though some may have been using CPOE prior to 2004. Both the availability of Tall Man capability and the specific pairs selected for Tall Man treatment would have been decided by the hospital, the electronic medical record vendor or the medication knowledge base vendor. It is possible that none or not all of the 11 pairs of names studied by Zhong *et al* were converted to Tall Man at every site, and some pairs may have been converted earlier than 2007. The authors did not document whether Tall Man was *actually in use* in the 42 study hospitals, nor did they confirm whether or at what point the each of the 11 pairs of names selected for analysis was subject to Tall Man treatment in each hospital. Although recommended by the 2007 report of The Joint Commission as a national patient safety goal, Tall Man was never *required* by any regulatory body. It was one of several safe practices hospitals

could have adopted to meet the goal.[8] If a hospital were not using Tall Man, either at all, or for the specific pairs studied, or during a specific time period, then one would not expect a reduction in error for that site, pair or time period. Not knowing these details leaves open the interpretation that the null effect observed by Zhong *et al* was due not to the ineffectiveness of Tall Man but rather to the non-implementation of Tall Man for the study pairs during the observation period.

The authors selected confusing pairs of drug names from the ISMP list,[11] which contains confusions of all types (brand/brand, brand/generic and generic/generic). They then converted all names to generic because their database only used generic names. Some brand name pairs are similar, but the corresponding generics are much less so (eg, Prozac/Prograf vs fluoxetine/tacrolimus). If the hospitals in question used generic names for ordering, the likelihood of confusion between two relatively dissimilar generic names is smaller than if the brand names were used. This problem is further complicated by potential differences in the requirement to order by generic name only, or both brand and generic, in CPOE systems used during the study period. Hospitals that only allowed generic orders could not have made any of the brand/generic or brand/brand errors selected for analysis by Zhong *et al*. Furthermore, Tall Man is a strategy for preventing visual perception errors only. One should not expect Tall Man to affect either memory errors or auditory perception errors.

In terms of the rate of error, the authors cite two other studies that tested systems for automated detection of LASA errors,[12 13] but did not cite two studies that used direct observation to estimate the wrong-drug error rate, which both estimated to be roughly one per thousand.[1 2] Readers should be cautioned about comparing the rates reported in the current study to those reported in the direct observation studies. In the current study, the denominator is 4-day hospitalisations. In the direct observation studies, the denominator is number of prescriptions dispensed or ordered.

The authors used patterns of orders of each drug across a 4-day sliding window to define when an error occurred. This approach has a precedent,[13] but the authors used only some of the possible patterns. Certain plausible patterns were not included because they did not lend themselves to detection using the authors' method (see figure 1). Figure 1A and 1B are taken from Zhong *et al* and exhibit some typical suspicious patterns used in the study. Figure 1C and 1D were plausible but not used in the study. Figure 1C corresponds to a situation where drug B is given in error on the first day, and the error is never caught. Figure 1D corresponds to a situation when the error occurs on the last day, and there is no time to observe the error being corrected. Using the authors' methods, the first pattern in figure 1C is indistinguishable from
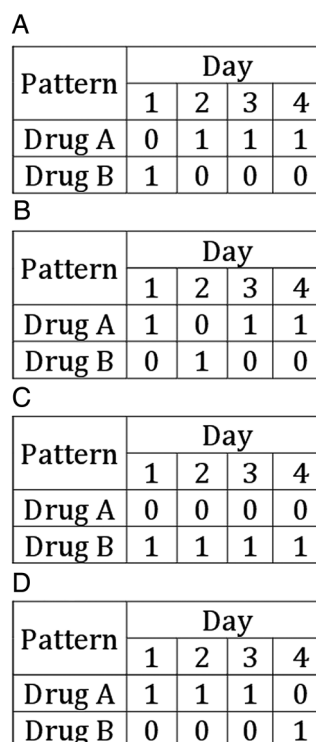
Figure 1 The figure depicts possible patterns of drug dispensing for a pair of look-alike/sound-alike (LASA) drugs. Drug A and drug B might refer to morphine and hydromorphone or vincristine and vinblastine or any other LASA pair. A and B are from Zhong *et al*.[10] C and D are plausible error patterns that were not included.

an intentional order for drug B. It could only be detected as an error if one could detect a mismatch between drug B's indications and the patient's current problems.[14]

The patterns used for detection by Zhong *et al* all required that the errors be intercepted and corrected within 4 days. This means that the study only measured the rate of *intercepted* LASA errors. Non-intercepted errors, which may account for 20% of all hospital medication errors,[2] could not be detected using these methods. While this limitation seems acceptable when trying to measure the effect of Tall Man in a quasi-experimental design, the method is not measuring the actual LASA error rate, and at best is a non-validated estimate of the rate of intercepted LASA errors. The estimate is non-validated because charts were not reviewed to confirm that an actual error occurred. For example, a patient who received 1 day of hydroxyzine and then 3 days of hydralazine might have had legitimate indications for both. Although the pattern hydroxyzine–hydralazine–hydralazine–hydralazine is suspicious, only chart review can confirm whether it is actually an error.

LABORATORY-BASED EXPERIMENTS ON TALL MAN LETTERING
Until now, the evidence base for or against the usefulness of Tall Man lettering came only from laboratory

experiments. Table 1 summarises results from all laboratory-based experiments published in peer-reviewed journals.

Table 1 illustrates that there was very little evidence, for or against Tall Man lettering, when it was recommended widely in 2007. There were only four peer-reviewed published experiments. None of them included healthcare professionals, and only two measured error rates. Other measures such as response times and eye-tracking gaze patterns are valuable, but they do not directly address errors, the crucial measure of clinical importance. Of the two experiments that measured error rates,[15 16] only one showed that Tall Man reduced drug name confusion errors.[15]

Subsequently, more laboratory-based experiments have been conducted. Of the 15 studies in table 1, 10 (67%) provided at least some evidence for the effectiveness of Tall Man—either in analyses of error rates or some other measure, such as response time. These experiments provide limited evidence to support benefit from Tall Man. Error rates constitute the crucial measure, and only eight experiments (out of 13, ie, 62%) show reduced error rates in the Tall Man condition.

Simply counting up the number of studies with positive or negative results does not account for possible differences in quality among studies. By our reading, the eight positive studies did not appear to be any better designed than those with null results. In fact, it is possible that some of the positive results are due to demand characteristics.[17] Demand characteristics describe situations in which if a participant knows the purpose of the experimental manipulation and the expected behaviour of participants, they behave in a way that conforms to expectations. In effect, the study participants adjust their behaviour (consciously or not) to support the perceived goal of the experimenter. In several Tall Man experiments,

Table 1 Published laboratory-based experiments evaluating the impact of Tall Man lettering on look-alike/sound-alike drug errors

Author(s) and year of study	Number and type of participants	Type of laboratory-based experimental task	Were participants informed of the purpose of Tall Man lettering before the task?	Was there an effect of Tall Man lettering on drug name confusion errors?	Was there an effect of Tall Man lettering on other measures?	Was there an effect of degree of drug name similarity on errors or other measures?
Filik et al (2004)[15]	20 Lay people	Visual search	Unknown	Yes	Yes	–
Filik et al (2006)[16] Experiment 1	20 Lay people	Same–different	No	–	No	–
Filik et al (2006)[16] Experiment 2	20 Lay people	Same–different	Yes	–	Yes	–
Filik et al (2006)[16] Experiment 3	28 Lay people	Recognition memory	Unknown	No	Yes	–
Schell (2009)[28] Experiment 1	102 Lay people	Same–different	Unknown	No*	–	Yes
Schell, 2009[28] Experiment 2	11 Healthcare professionals	Same–different	Unknown	No	–	–
Filik et al (2010)[29] Experiment 1	56 Lay people	Same–different	Unknown	Yes	Yes	–
Filik et al (2010)[29] Experiment 2	127 Healthcare professionals	Visual search	Unknown	Yes	No*	–
Darker et al (2011)[18]	144 Healthcare professionals	Reicher–Wheeler	Yes	Yes	–	–
Irwin et al (2013)[19] Experiment 1	60 Lay people	Visual search	No	No	No	–
Irwin et al (2013)[19] Experiment 2	28 Healthcare professionals	Visual search	No	No	No	–
Or and Chan, (2014)[23]	60 Lay people and healthcare professionals	Same–different	Unknown	Yes	No	Yes
Or and Wang (2014)[24] Experiment 1	40 Lay people	Same–different	Unknown	Yes	–	Yes
Or and Wang (2014)[24] Experiment 2	40 Healthcare students	Same–different	Unknown	Yes	–	Yes
DeHenau et al (2016)[30]	40 Healthcare professionals, 40 Lay people	Flicker change detection	Unknown	Yes	Yes	–

*There was a significant effect but in the opposite direction (worse performance for Tall Man relative to standard text).

participants were told before the task that Tall Man lettering was meant to aid discrimination between similar names, knowledge that may have influenced task performance.[16 18] In one study, when participants were not told beforehand about Tall Man's purpose, Tall Man lettering was not effective, but then in a follow-up study, when participants were told, Tall Man was effective.[16] Elsewhere, failure to find a positive effect of Tall Man lettering was attributed to participants not knowing the purpose beforehand.[19] It could be also argued that studies in which participants know the purpose beforehand better simulate real-world clinical practice because most healthcare professionals will know the purpose of Tall Man lettering. Although these studies may be more ecologically valid in a sense, they nevertheless suffer from laboratory-based demand characteristics that are unlikely to occur in the real world. In an experimental setting, the tasks are short enough for participants to keep at the forefront of their mind the purpose of Tall Man lettering and the goal of the experimenter, but in the real world, with long hours and repetitive tasks, healthcare professionals are unlikely to be able to continuously keep the purpose and goals in their minds as saliently as they would in a laboratory experiment. It is not clear how many clinicians (especially non-pharmacists) understand the purpose of Tall Man. Thus, some of the positive results in favour of Tall Man lettering may be due to experiment-specific phenomena that will not translate to real-world clinical practice.

Finally, because of submission and publication biases,[20] experiments that found no effect of Tall Man lettering may be under-represented in the published literature. If the Tall Man benefit were robust, it ought to appear in all or almost all of the experiments. Null results in more than a third of the experiments suggest that the effects of Tall Man are either too small to be detected or are dependent on certain task or participant characteristics. One would like the effect to be more robust and replicated in clinical settings prior to making policy.

POLICY IMPLICATIONS

Zhong et al's overall finding is that a widely disseminated policy, one strongly endorsed by three of the most respected medication safety organisations in the world (The Joint Commission, the US FDA and ISMP), had no apparent beneficial effect in 42 paediatric hospitals over a 9-year period. The negative result, especially if replicated in other studies, raises policy questions in the realm of patient safety. The most important of these concerns the amount and quality of evidence that should be required before new error-prevention measures are widely promulgated or required—a long-standing source of tension in the field.[21 22] In this case, as of 2007 when Tall Man went into wide use, only two published studies

on a total of 88 laypeople had been published, and only one of them showed a reduction in error rates. We should demand a higher standard of evidence. Clinical evidence of Tall Man's effectiveness ought to have been required prior to widespread implementation. The counterargument is that demanding evidence before acting might slow the implementation of safe practices that pass the test of common sense and expert opinion. And since Tall Man seems harmless, some might say the demand for evidence of effectiveness may be too stringent.

A definitive study would cluster-randomise pharmacies, hospitals or health systems to use or not use Tall Man and then measure drug name confusion error rates by chart review or direct observation. Alternatively, an interrupted time series study could be carried out in a single hospital, measuring the error rate before and after implementation (or, at this point, de-implementation of Tall Man). These designs would also permit evaluation of other methods of name differentiation, such as different colours (red) or combining boldface with Tall Man.[23 24] Such studies have not been conducted and are badly needed.

One reason to demand more evidence prior to implementation is the opportunity cost of implementing any error-prevention measure.[25] Health information technology staff represent a finite resource whose work must be carefully prioritised, especially in this era of proliferating quality measures related to 'meaningful use' of electronic health records.[26] Choosing to do one thing often means postponing something else. There are many clinical decision support measures and new technologies available that may improve patient safety, and prioritisation should be evidence based. When we implement ineffective interventions, we forego, at least temporarily, the opportunity to put in place more effective, evidence-based safety measures. In this sense even a 'harmless' mandate is not without cost. There will always be costs—to implement the mandated strategy, to monitor for its presence and the distraction from other potential activities that have greater evidence of benefit.

CONCLUSION

Zhong et al conclude that the Tall Man intervention has no benefit because the error rate for the studied pairs of names did not decrease in the postimplementation period. If Tall Man were a robust and effective intervention, one would expect to have seen some evidence of benefit in a study of this size and quality. The study makes an important contribution to understanding the role of Tall Man in preventing drug name confusion errors, but methodological limitations prevent it from offering a definitive result. If this finding is corroborated in other settings with other pairs of names, after addressing the methodological problems cited above, the continued use of Tall Man in its current form should be reconsidered. The

history of the Tall Man intervention, with its widespread implementation proceeding without evidence of its effectiveness, is an object lesson for those who make policy about patient safety.

Twitter Follow Bruce Lambert at @bruce_lambert

Competing interests BLL has ownership interests in two companies that provide software and consulting services related to preventing drug name confusions. Those companies had no role in the preparation of this manuscript.

Provenance and peer review Commissioned; internally peer reviewed.

REFERENCES

1 Flynn EA, Barker KN, Carnahan BJ. National observational study of prescription dispensing accuracy and safety in 50 pharmacies. *J Am Pharm Assoc (Wash)* 2003;43:191–200.
2 Cina JL, Gandhi TK, Churchill W, *et al*. How many hospital pharmacy medication dispensing errors go undetected? *Jt Comm J Qual Patient Saf* 2006;32:73–80.
3 Maeda K, Ueda M, Ohtaka H, *et al*. A massive dose of vincristine. *Jpn J Clin Oncol* 1987;17:247–53.
4 Chiche L, Thomas G, Canavese S, *et al*. Severe hemorrhagic syndrome due to similarity of drug names. *Eur J Intern Med* 2008;19:135–6.
5 Grissinger M. Morphine and Hydromorphone: an omnipresent risk of mix-ups. *Pharm Ther* 2008;33:194–247.
6 U. S. Food and Drug Administration. Name differentiation project. 2007. http://www.fda.gov/drugs/drugsafety/medicationerrors/ucm164587.htm
7 Institute for Safe Medication Practices. FDA and ISMP lists of look-alike drug names with recommended Tall Man letters. 2011. http://www.ismp.org/tools/tallmanletters.pdf
8 The Joint Commission. National Patient Safety Goal. 2007. http://www.pbm.va.gov/vacenterformedicationsafety/othervasafetyprojects/appendixijcaholasa.pdf
9 Grissinger M. Tall man letters are gaining wide acceptance. *Pharm Ther* 2012;37:132–48.
10 Zhong W, Feinstein JA, Patel NS, *et al*. Tall man lettering and potential prescription errors: a time series analysis of 42 children's hospitals in the USA over 9 years. *BMJ Qual Saf* 2016;25:233–40.
11 ISMP. ISMP's list of confused drug names. 2015. http://www.ismp.org/Tools/Confused-Drug-Names.aspx
12 Basco WT Jr, Ebeling M, Hulsey TC, *et al*. Using pharmacy data to screen for look-alike, sound-alike substitution errors in pediatric prescriptions. *Acad Pediatr* 2010;10:233–7.
13 Phatak HM, Cady PS, Heyneman CA, *et al*. Retrospective detection of potential medication errors involving drugs with similar names. *J Am Pharm Assoc* 2005;45:616–21.
14 Galanter WL, Bryson ML, Falck S, *et al*. Indication alerts intercept drug name confusion errors during computerized entry of medication orders. *PLoS ONE* 2014;9:e101977.
15 Filik R, Purdy K, Gale A, *et al*. Drug name confusion: evaluating the effectiveness of capital ("Tall Man") letters using eye movement data. *Soc Sci Med* 2004;59:2597–601.
16 Filik R, Purdy K, Gale A, *et al*. Labeling of medicines and patient safety: Evaluating methods of reducing drug name confusion. *Hum Factors* 2006;48:39–47.
17 Orne MT. On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *Am Psychol* 1962;17:776–83.
18 Darker IT, Gerret D, Filik R, *et al*. The influence of 'Tall Man' lettering on errors of visual perception in the recognition of written drug names. *Ergonomics* 2011;54:21–33.
19 Irwin A, Mearns K, Watson M, *et al*. The effect of proximity, Tall Man lettering, and time pressure on accurate visual perception of drug names. *Hum Factors* 2013;55:253–66.
20 Easterbrook PJ, Berlin JA, Gopalan R, *et al*. Publication bias in clinical research. *Lancet* 1991;337:867–72.
21 Auerbach AD, Landefeld CS, Shojania KG. The tension between needing to improve care and knowing how to do it. *N Engl J Med* 2007;357:608–13.
22 Shojania KG, Duncan BW, McDonald KM, *et al*. Safe but sound: patient safety meets evidence-based medicine. *JAMA* 2002;288:508–13.
23 Or CK, Chan AH. Effects of text enhancements on the differentiation performance of orthographically similar drug names. *Work* 2014;48:521–8.
24 Or CK, Wang H. A comparison of the effects of different typographical methods on the recognizability of printed drug names. *Drug Saf* 2014;37:351–9.
25 Slight SP, Berner ES, Galanter W, *et al*. Meaningful use of electronic health records: experiences from the field and future opportunities. *JMIR Med Inform* 2015;3:e30.
26 Meaningful use definition & objectives. 2015. https://www.healthit.gov/providers-professionals/meaningful-use-definition-objectives
27 Poon EG, Cina JL, Churchill W, *et al*. Medication dispensing errors and potential adverse drug events before and after implementing bar code technology in the pharmacy. *Ann Intern Med* 2006;145:426–34.
28 Schell KL. Using enhanced text to facilitate recognition of drug names: Evidence from two experimental studies. *Appl Ergon* 2009;40:82–90.
29 Filik R, Price J, Darker I, *et al*. The influence of tall man lettering on drug name confusion: a laboratory-based investigation in the UK using younger and older adults and healthcare practitioners. *Drug Saf* 2010;33:677–87.
30 DeHenau C, Becker MW, Bello NM, *et al*. Tallman lettering as a strategy for differentiation in look-alike, sound-alike drug names: the role of familiarity in differentiating drug doppelgangers. *Appl Ergon* 2016;52:77–84.