

Drug selection errors in relation to medication labels: a simulation study

P. Garnerin,¹ T. Perneger,² P. Chopard,² M. Arès,¹ R. Baalbaki,³ P. Bonnabry³ and F. Clergue¹

1 Anaesthesiology Service, 2 Quality of Care Service, and 3 Hospital Pharmacy, Geneva University Hospitals (HUG), rue Micheli-du-Crest, 21, 1211 Geneva 14, Switzerland

Summary

The aim of this study was to assess the impact of differences in drug label information on injectable drug selection errors. Differences in the display of drug strength information were assessed in a randomised controlled trial involving ward nurses, intensive care nurses, nurse anaesthetists, ward physicians, and anaesthetists. A set of 24 on-screen tasks were constructed. For each task, a label corresponding to an instruction consisting of two from three possible pieces of information (concentration, quantity, volume) had to be selected from a list of 10 items. The set was presented three times to participants using three different label formats. Format A provided two pieces of strength information different from those in the instruction. Format B and C provided all three pieces in a random and a fixed sequence, respectively. The frequency of errors was statistically higher with formats A and B than with format C, and greater in nurses than in anaesthetists. Regulatory bodies should therefore implement a standard requiring that the concentration (expressed in 'mg.ml⁻¹'), the amount and the volume of drug be displayed on medication labels in fixed locations.

Correspondence to: Philippe Garnerin

E-mail: philippe.garnerin@hcuge.ch

Accepted: 27 May 2007

Incorporating safety into the design of healthcare systems is essential for the prevention of medical errors [1]. In practice, however, many hospital care processes have not been designed for safety. Rather, they have developed over time and often contain intrinsic system weaknesses. One example is the design of drug labels. Labels represent the main 'drug/user' interface, and numerous case reports [2, 3] suggest that their characteristics, such as the drug name, type-face, colour coding or expression of drug strength greatly contribute to the incidence of medication errors. However, the impact of drug label design on human performance has received little attention. Studies have shown that print characteristics affect the readability of labels [4, 5], and that orthographic and phonetic similarity between drug names increases the probability of a false recognition error [6]. Printing a section of a drug name in capital letters has been shown to reduce errors in drug name recognition [7, 8].

As calculation errors are common amongst healthcare professionals [9–11], several authors have advised that the display of drug strength information on injectable medi-

cation labels be standardised [10, 12]. To understand the potential impact of this measure on human reliability, we designed a computer-based experiment to compare the effect of three different drug strength information formats, incorporating various levels of standardisation, on the frequency of drug selection errors made by groups of healthcare professionals.

Methods

Study population

Five different professional groups (ward nurses, intensive-care nurses, nurse anaesthetists, ward physicians, anaesthetists) each group consisting of 15 volunteers, were involved in the experiment. The volunteers, who were recruited in the institution by an advertising campaign, were informed of the aim and design of the study, gave their written consent, and received monetary compensation of approximately €15. The protocol was approved by the institution's ethical committee for human research.

Study design

In this randomised experiment, health professionals were asked to complete a series of 72 on-screen tasks in a dedicated room. For each task, a target drug instruction appeared at the top of the computer screen and a list of 10 drug labels were displayed below. Of the list, only one label matched the instruction. Participants were asked to select the correct label by clicking on the corresponding radio button. A radio button tagged 'no corresponding label' was also available.

Three types of target drug instruction were used: quantity and volume of the drug, volume and concentration, and quantity and concentration. Concentration, quantity and volume were expressed in $\text{mg}\cdot\text{ml}^{-1}$, mg and ml, respectively.

The series of on-screen tasks was based on an initial set of 24 distinct tasks (eight tasks per type of target drug instruction), which were presented three times to participants. For each presentation, a different format for displaying drug strength information on the 10 labels was used (Fig. 1).

Format A: incomplete information. This format was particular in that only two of three pieces of information (i.e. quantity and volume of the drug, volume and concentration, quantity and concentration) were shown on drug labels. Half of the labels were of the same type as the target drug instruction but the correct label was always of a different type. For instance, if the instruction stated the quantity and volume, five labels on the list contained a quantity and a volume, and five, including the correct label, mentioned a concentration and either a volume or a quantity. Furthermore, for this particular format, the concentration on drug labels was expressed in $\text{mg}\cdot\text{ml}^{-1}$ in half of the tasks and as a percentage (%) otherwise.

Format B: complete information (concentration expressed in $\text{mg}\cdot\text{ml}^{-1}$ only, quantity, and volume of the drug) in a random sequence that differed on each of the drug labels;

Format C: complete information (concentration expressed in $\text{mg}\cdot\text{ml}^{-1}$ only, quantity, and volume of the drug) in a fixed sequence (concentration, quantity, volume) on all drug labels.

In all three drug strength information formats, quantity and volume on drug labels were expressed in mg and ml, respectively.

Within each set of 24 tasks, the sequence of tasks was randomised once for all and was therefore identical for each participant. To neutralise the effect of learning or accumulating fatigue, each participant performed the three sets presented in a random sequence.

Each participant, who carried out 72 tasks overall, was provided with two examples before starting the three sets. Participants were given up to 60 s to complete each task.

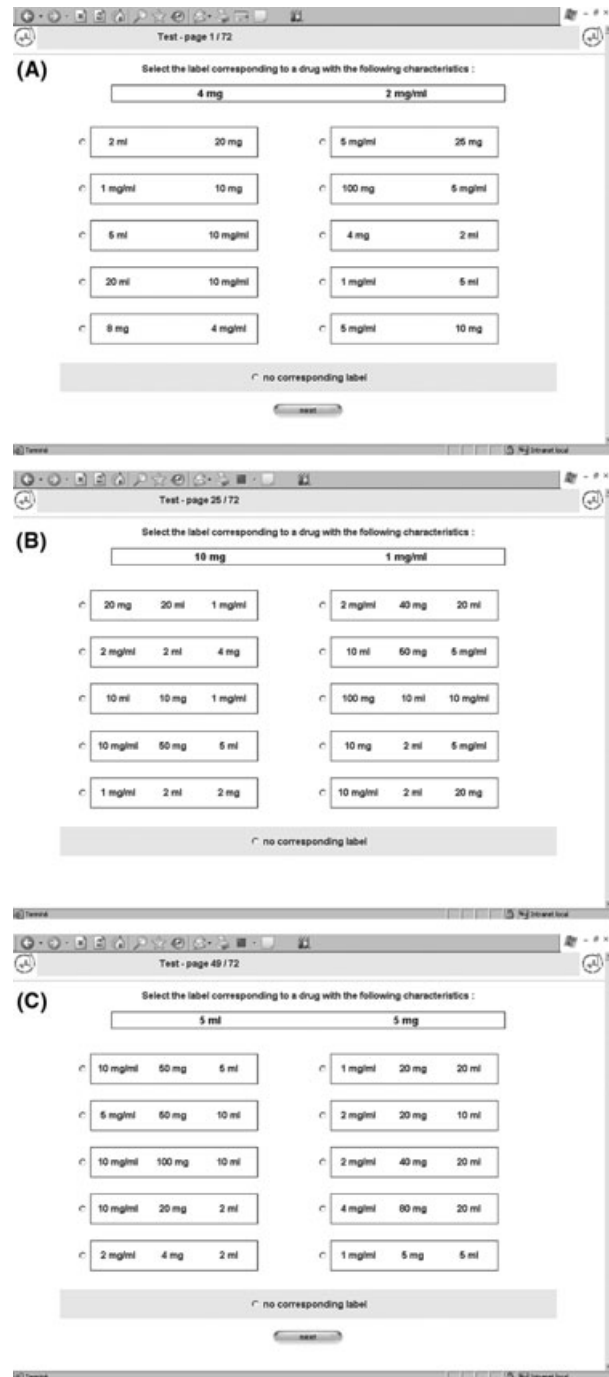


Figure 1 Examples of on-screen tasks for drug strength information formats A, B, and C.

Data collection

For each task, data collected comprised the answer, the response time, the presentation sequence of the three sets, the drug strength information format, the type of target drug instruction, the professional group, and the participant's identification number. For drug strength

information Format A (incomplete information), the concentration unit (mg.ml^{-1} or percentage) was also noted.

Statistical analysis

The main dependent variable was selection error, defined as anything other than the right answer. Secondary outcomes were 'drug label not found' selection error (defined as either 'no corresponding label' answer or no answer within 60 s), and response time. The independent variables were the presentation sequence of the three sets, type of target drug instruction, drug strength information format, concentration units (mg.ml^{-1} or percentage) and professional group. The 75 participants and 72 tasks were also identified.

The frequencies of selection errors, 'drug label not found' selection errors and response time were obtained for levels of each independent variable. Univariate and multivariate analyses were conducted using generalised linear models. To account for the dependency in an individual's answers, we performed these analyses using generalised estimation equations, with the participant's identification number used as grouping variable, and the correlation matrix set to 'exchangeable'. Comparisons of selection error frequencies across participants were done using Pearson's Chi-squared test.

A p value of <0.05 was considered statistically significant. All calculations were performed using STATA 8.2.

Results

Seventy-five participants completed 72 tasks each, producing a total of 5400 tasks. Overall, 722 (13.4%) of the answers corresponded to selection errors, among which 331 (6.1%) were 'drug label not found' selection errors. Mean response time and standard deviation were 21.7 and 13.4 s, respectively (median 17 s).

The frequency of errors was higher for drug strength information displayed in formats A (mean error rate 29.7%) and B (6.4%) than in format C (4.1%) (Table 1). In multivariate analysis as well (Table 2), the most complete and standardised format, C, led to the least number of errors. The error rate was also higher for the eight instructions that mentioned the quantity and volume of the drug, than for those that incorporated the concentration (Tables 1 and 2). This was also observed for tasks where labels included the percentage symbol as concentration unit, rather than mg.ml^{-1} . Anaesthetists produced fewer errors than other physicians and than the nursing groups. This pattern of results was similar for 'drug label not found' selection errors, except that this type of error did not depend on the type of target drug instruction, and varied less with the professional group (Table 1). Mean response times differed considerably according to drug strength information formats and increased with the use of percentage as a concentration unit, but differences across professional groups were small (Table 1).

Table 1 Univariate results.

	Frequency (%)	Drug selection error		'Drug label not found' selection error		Response time	
		%	p value*	%	p value*	Mean	p value*
Drug strength information format			$p < 0.001$		$p < 0.001$		$p < 0.001$
A: incomplete	1800 (33.3)	29.7		14.4		30.3	
B: complete, random order	1800 (33.3)	6.4		2.9		18.6	
C: complete, fixed order	1800 (33.3)	4.1		1.1		16.1	
Concentration unit			$p < 0.001$		$p < 0.001$		$p < 0.001$
mg.ml^{-1}	4500 (83.3)	9.1		4.1		19.5	
%	900 (16.7)	34.9		16.2		32.4	
Type of target drug instruction			$p < 0.001$		$p = 0.89$		$p < 0.001$
Quantity and volume	1800 (33.3)	17.3		6.4		22.3	
Concentration and volume	1800 (33.3)	10.6		6.0		20.7	
Concentration and quantity	1800 (33.3)	12.2		5.9		21.9	
Professional group			$p < 0.001$		$p = 0.09$		$p = 0.03$
Nurse	1080 (20.0)	19.0		7.7		24.5	
Nurse anaesthetist	1080 (20.0)	11.5		6.8		21.3	
Nurse intensive care	1080 (20.0)	16.9		6.9		21.9	
Ward physician	1080 (20.0)	15.5		6.7		20.6	
Anaesthetist	1080 (20.0)	4.0		2.6		20.0	

*From generalised estimation equations model.

Table 2 Odds ratios of a drug selection error adjusted for all variables in table. Values are odds ratio (95% CI) and p value.*

Drug strength information format			
A: incomplete	9.1	(6.6–12.7)	p < 0.001
B: complete, random order	1.7	(1.2–2.3)	p = 0.001
C: complete, fixed order	1.0		
Concentration unit			
(drug strength information format A only)			
mg.ml ⁻¹	1.0		
%	1.8	(1.4–2.2)	p < 0.001
Type of target drug instruction			
Quantity and volume	2.1	(1.7–2.6)	p < 0.001
Concentration and volume	1.0		
Concentration and quantity	1.2	(1.0–1.5)	p = 0.08
Professional group			
Nurse	14.1	(6.9–28.9)	p < 0.001
Nurse anaesthetist	4.8	(2.3–10.1)	p < 0.001
Nurse intensive care	12.3	(6.0–25.2)	p < 0.001
Ward physician	8.8	(4.3–17.9)	p < 0.001
Anaesthetist	1.0		

*From generalised estimation equations models.

Analysis of 75 participants

The frequency of selection errors varied between 0.0% and 40.3% across individuals (p < 0.001), with a mean of 13.4% and a standard deviation of 10.9%, as well as between participants within professional groups (p < 0.01). This frequency varied less amongst anaesthetists than in other professional groups (Fig. 2a). The average response time ranged from 13.6 s to 32.8 s, with a standard deviation of 4.4 s. The mean error rate and the average response time were moderately correlated across the 75 individuals (Pearson *r* = 0.42).

Analysis of 72 tasks

The mean frequency of selection errors per target instruction ranged from 0.0% to 53.3%, with a mean of 13.4% and a standard deviation of 13.6%. More variation in error frequencies was observed in drug strength information Format A than in Formats B and C (Fig. 2b). The average response time ranged from 11.4 s to 43.2 s, with a standard deviation of 7.1 s. The mean error rate and the average response time for the 72 prescriptions were closely correlated (Pearson *r* = 0.90). This correlation decreased somewhat if ‘drug label not found’ selection errors and corresponding response times were eliminated (Pearson *r* = 0.74).

Discussion

In our study we have demonstrated that standardising the format of drug strength information presented on labels for injectable medication, substantially improves human performance during selection tasks. In that respect, the

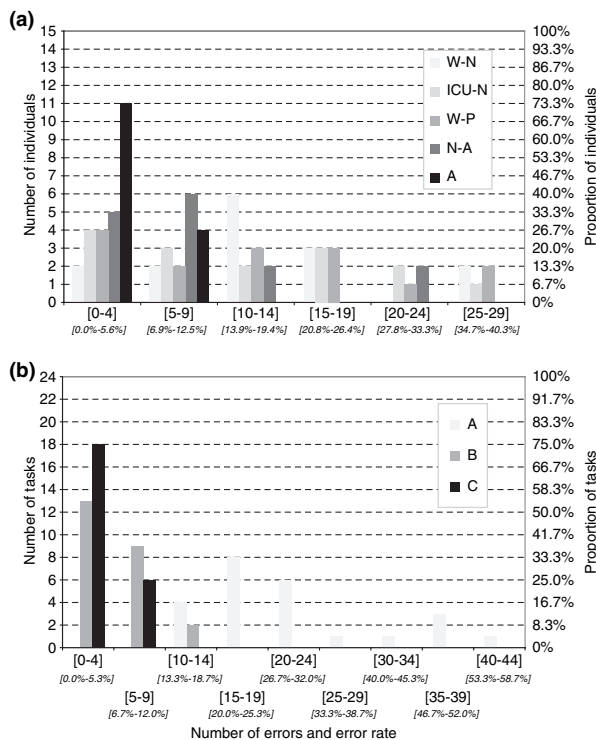


Figure 2 Drug selection errors frequency per participants within professional groups (a) and per tasks within drug strength information formats (b). a) W-N, ward nurses; ICU-N, intensive care nurses; W-P: ward physicians; N-A, nurse anaesthetists; A, anaesthetists; b) A: incomplete drug strength information format; B: complete drug strength information format, random order; C: complete drug strength information format, fixed order. Example of a reading: (a) 11 anaesthetists (73.3% of participants; *n* = 15) made 0–4 errors (error rate: 0–5.6%, *n* = 72); (b) in drug strength information format C, 18 tasks (75.0% of tasks; *n* = 24) were responsible for 0–4 errors (error rate: 0%–5.3%; *n* = 75).

study confirms prior expectations regarding drug label standardisation and provides an early insight into the potential magnitude of its impact.

Standardised labels mentioning the concentration, amount and volume, with the concentration expressed in ‘mg.ml⁻¹’, produced the lowest rate of selection errors and the most rapid selection of the correct product. Indeed, by providing full drug strength information, standardisation removes the requirement for mental calculation, an activity prone to errors, especially when participants have to calculate concentration from amount and volume or convert concentrations expressed as a percentage. However, extending standardisation by displaying the concentration, quantity and volume at fixed locations on the label further improves human performance. This is probably because this feature helps users spot details of drug strength information for the identification

of medication. Indeed, although mental calculation was no longer required, the rate of 'drug not found' selection errors was substantially reduced when the order of the information was fixed rather than variable.

However significant the error prevention mechanism produced by standardisation, it has to be borne in mind that medication errors are in essence systemic. Although standardisation reduced variations in the rate of selection errors between professional groups, anaesthetists still performed better than the other groups, a finding already noted by others [11]. However, large differences are noted between individuals. This observation emphasises that the reliability of the medication process will only significantly increase if all stakeholders contribute to this objective. Pharmaceutical companies need to improve the design of drug labels, and nursing and medical schools as well as other healthcare organisations will have to implement ways of improved selection and training of healthcare professionals with regard to arithmetic skills.

Although striking, our results were obtained using a computer-based simulation, and the question remains if they can be extrapolated to real-life healthcare settings. First of all, to achieve adequate statistical power, we put professionals in an artificial situation where they had to choose drug strength from 10 options – an unlikely clinical scenario. Most of the time, indeed, available strengths are few, and frequently only one is available. In addition, participants were not influenced by factors such as stress, workload or work environment that affect human performance in real conditions. Consequently, the error rates and response times we measured should not be seen as directly applicable to real-life human performance. However, professionals occasionally have to calculate whether or not the drug they intend to use is the correct strength: a change of supplier might mean a different type of packaging, or new staff might be unfamiliar with local medication available. Confusions over drug strengths will then occur. In such instances, and whatever the uncertainty attached to our figures, it is still likely that the benefit provided by removing the need for mental calculation by using a standardised format of drug strength information will remain valid.

Our simulation also partly mimicked the variety of situations in which drug strengths are used. Indeed, participants were asked to select a specific strength from amongst several available. However, we did not explore the alternative where professionals have to calculate a volume from drug strength information on the label. Although this should be confirmed in another similar experiment, our hypothesis is that the same error

prevention mechanism associated with standardisation will be seen.

Given the potential hazards conveyed by the administration of most drugs in solution, and although some exceptions should be considered for biological products such as vaccines [13], international and national regulatory bodies should implement a standard requiring that the concentration is expressed in $\text{mg}\cdot\text{ml}^{-1}$, with the amount and the volume displayed on medication labels at fixed locations.

References

- 1 Kohn L, Corrigan J, Donaldson M. *To Err Is Human. Building a Safer Health System*. Washington, DC: National Academy Press, 2000.
- 2 Edgar TA, Lee DS, Cousins DD. Experience with a national medication error reporting program. *American Journal of Hospital Pharmacy* 1994; **51**: 1335–8.
- 3 Cohen M. *Medication Errors: Causes, Prevention, and Risk Management*. Sudbury, MA: John and Bartlett, 1999.
- 4 Watanabe R, Gilbreath K, Sakamoto C. The ability of the geriatric population to read labels on over-the-counter medication containers. *Journal of the American Optometric Association* 1994; **65**: 32–7.
- 5 Wogalter MS, Vigilante WJ Jr. Effects of label format on knowledge acquisition and perceived readability by younger and older adults. *Ergonomics* 2003; **46**: 327–44.
- 6 Lambert B, Chang K, Lin S. Effect of orthographic and phonological similarity on false recognition of drug names. *Social Science and Medicine* 2001; **52**: 1843–57.
- 7 Filik R, Purdy K, Gale A, et al. Drug name confusion: evaluating the effectiveness of capital ('Tall Man') letters using eye movement data. *Social Science and Medicine* 2004; **59**: 2597–601.
- 8 Filik R, Purdy K, Gale A, et al. Labeling of medicines and patient safety: evaluating methods of reducing drug name confusion. *Human Factors* 2006; **48**: 39–47.
- 9 Bayne T, Bindler R. Effectiveness of medication calculation enhancement methods with nurses. *Journal of Nursing Staff Development* 1997; **13**: 293–301.
- 10 Rolfe S, Harper NJ. Ability of hospital doctors to calculate drug doses. *British Medical Journal* 1995; **310**: 1173–4.
- 11 Wheeler DW, Remoundos DD, Whittlestone KD, et al. Doctors' confusion over ratios and percentages in drug solutions: the case for standard labelling. *Journal of the Royal Society of Medicine* 2004; **97**: 380–3.
- 12 Wheeler SJ, Wheeler DW. Dose calculation and medication error – why are we still weakened by strengths? *European Journal of Anaesthesiology* 2004; **21**: 929–31.
- 13 Nunn DS. Calculating drug doses. Mass concentration can't be used universally. *British Medical Journal* 1995; **311**: 512.