Predicting and Preventing Drug Name Confusion Errors:
A Summary of Findings

Bruce L. Lambert, Ph.D.[a, b]

Swu-Jane Lin, B. Pharm., MS[a]

Sanjay K. Gandhi, B. Pharm., Ph.D.[c]

Ken-Yu Chang, B. Pharm., MPH[a]

[a]Department of Pharmacy Administration

[b]Department of Pharmacy Practice

University of Illinois at Chicago

[c]Searle, Skokie, IL

Address: Department of Pharmacy Administration,
                833 S. Wood Street (M/C 871), Chicago, IL 60612-7231
Phone: 312-996-2411
Fax: 312-996-0868
Email: lambertb@uic.edu

Introduction

This presentation will focus on look-alike and sound-alike medication errors, a category of errors that accounts for one out of every four medication errors reported nationally.[3] Look-alike and sound-alike errors occur when one drug name is confused with another due to similarities in spelling and/or pronunciation. We have developed a systematic approach to quality control and quality improvement in the drug naming process that, if widely adopted, could substantially reduce the rate of look-alike and sound-alike (LASA) medication errors.[4] This approach includes: (a) a model of the psychological mechanisms that underlie LASA errors, (b) automated measures of lexical (i.e., word to word) similarity, (c) case-control evaluation of the similarity measures in relation to the risk of error, (d) dose-response evaluation of the relationship between similarity and the probability of error, (e) tests to predict the likelihood of LASA errors between any given pair of names, (f) experimental evidence relating similarity to short-term memory errors made by pharmacists, and (g) automated search algorithms that scan medication databases for existing names that are similar to proposed new names. The theoretically-grounded, data-based, objective, systematic, and scientific nature of this approach differentiates it from previous work on LASA errors, which is comprised mainly of case reports and observational studies.[2]

Theoretical Background

Due to the fundamental importance of reading, psychologists have extensively studied word memory and word perception. What has emerged is a picture of how words are represented in memory, how they are perceived, and how they are retrieved from memory. Short-term memory provides a useful illustration. Most skilled performances, including the selection, dispensing, and administration of drugs, rely on the ability to retain verbal information in short-term, working memory.[5] A pharmacist reads a prescription, puts the prescription down, and goes back to the shelf to retrieve the drug. To retrieve the correct drug, the pharmacist must store the name in working memory, compare it to each name on the shelf, and select the shelved bottle that matches the remembered name. Working memory for verbal information consists of a system known as the phonological (or articulatory) loop.[5] The phonological loop is comprised of a phonological store, where the sound patterns of words are temporarily held, and a silent mental rehearsal system that uses 'inner speech' to continuously update the rapidly decaying representation being held in the phonological store. (Think of how one repeats an unfamiliar phone number over and over to

oneself in order to remember it.) This phonological store is vulnerable to decay and interference effects.

The more similar two words are, the fewer phonological features can be used to discriminate between them.

If two words share all but one phonological feature, and that feature is lost due to decay, then accurate

discrimination will not be possible, and errors will occur. This model of working memory has been used to

predict and explain a wide variety of error phenomena in both recall and recognition memory.[5, 6] Before

now, however, no one had applied this theoretical framework to guide research and prevention efforts in

relation to LASA errors. This general understanding of short-term memory for verbal information

motivates all of the empirical studies described below.

Methods and Results

Automated Measures of Lexical Similarity

Since the theoretical model treated similarity as a root cause of LASA errors, our first task was to

develop measures of similarity. The measures had to be objective because similarity was to be the main

theoretical variable in a series of subsequent experiments on LASA errors. The measures had to be automated

because they would eventually be necessary to compute similarity scores for large databases of name-pairs.

Automated measures could also be integrated into computerized prescribing and dispensing systems.[7]

To measure orthographic (i.e., look-alike or spelling) similarity, we used n-gram and edit distance

methods.[4, 8] The n-gram method for computing lexical similarity views each word as a sequence of

symbols. To compute the similarity between two words, simply break each word into its *n*-letter

subsequences and define similarity to be a function of the number of common subsequences. Consider the

names *atarax* and *marax*. The bigram similarity between these names is computed as follows. The bigrams

(i.e., two-letter subsequences) for *atarax* are {*at, ta, ar, ra, ax*}. The bigrams for *marax* are {*ma*, *ar*, *ra*, *ax*}.

The words share three out of a total of nine bigrams. The bigram similarity (using the Dice coefficient) is

$(2*3)/9 = 0.67$.[4] An analogous procedure can be used for letter trigrams. Edit distance, the other main

measure, is defined as the total number of edit operations (i.e., insertions, deletions, and/or transpositions)

needed to transform one word into another.[8] To change *atarax* to *marax*, one must change the initial *a* to an

m and delete the *t*. Hence, the edit distance between *atarax* and *marax* equals 2. Pronunciation guides

provided in the *USP Dictionary of USAN and International Drug Names* formed the basis for measures of

phonological (i.e., sound-alike) similarity.[U. S. Pharmacopeia, 1998 #424]r For example, pronunciation of

the drug lincomycin is given as (lin koe mye' sin), and tobramycin is given as (toe bra mye' sin). Previous

research has identified the phonological features that have the greatest effect on memory.[9] Based on this

work, we defined phonological similarity in terms of initial syllable, terminal syllable, accented syllable,

accent position, number of syllables, and number of common syllables.[10]

Case-Control Analysis of LASA Errors

To investigate the association between lexical similarity and the probability of LASA errors and to

develop a prognostic test for LASA errors, we did a series of studies based on a case-control design.[4] Cases

(*N*=969) were drawn from published lists of names that had been involved in LASA errors. Controls

(*N*=969) were drawn at random from an electronic version of the general index to the USP-DI, Volume I.[11]

The first step in the study was to examine the distribution of similarity (or distance) scores between error

pairs and control pairs of names. To be useful, error pairs and control pairs had to be distributed differently

with respect to the automated measures of similarity. The distribution of similarities for error pairs was

significantly different than the distribution of similarities for control pairs.[4] Similarity scores for error pairs

were skewed to the high end of the similarity scale (and the low end of the distance scale); whereas, the

distribution of similarity scores for control pairs followed exactly the opposite pattern (i.e., low similarities,

high distances).

The relative risk of error associated with a given level of similarity was then estimated. As

expected, similarity/distance was a significant risk factor leading to the occurrence of LASA errors.[4] The

next step was to develop a prognostic test that would allow one to predict in advance whether a pair of

names was likely to be involved in a LASA error. Development of such a test consisted primarily of finding

a threshold or cutoff point that, when used as a guide to prediction, yielded the fewest false positives and

false negatives. We succeeded in finding a cutoff that yielded a false positive rate of only 5% and a false

negative rate of only 9%.[4] The overall accuracy of the test was 94%. Finally, the dose-response relationship

between similarity and error probability was examined to see whether the likelihood of error increased as

similarity increased. As expected, a significant dose-response relationship was documented. The odds of a

pair being involved in an error increased dramatically as similarity increased.

Psychological Experiments: Similarity and Short-Term Memory Errors

Although the results reported above provide evidence of an association between similarity and LASA errors, the case-control design suffered from several weaknesses. The design was retrospective, and a variety of biases were operating in the selection of error and control pairs.[4] These factors prevented strong inferences about causality from being made. To make stronger inferences, we designed and carried out a series of prospective experiments on licensed, practicing pharmacists.[12] These experiments were designed to assess the effect similarity on recognition memory errors.[13] Recognition memory refers to the ability to distinguish between words that have recently been presented and 'new' (not previously presented) words. Two different experiments were conducted. The first examined the effect of *orthographic* similarity on recognition memory. The second examined the effect of *phonological* similarity on recognition memory. Fifteen licensed pharmacists participated in each experiment. Names were drawn from the combined 1992-1994 National Ambulatory Medical Care Survey (NAMCS) data set.[14] NAMCS prescribing frequency data were used to control for frequency in the construction of stimulus lists. Stimulus materials in the first experiment were forty pairs of drug names, eight pairs at each of five similarity levels. These names were separated into 16 study lists and 16 test lists. Each study list consisted of 5 words, and each test list consisted of ten words (e.g., the five words from the study list plus five 'new' words, each similar to the test words to a slightly different extent). For example, pharmacists were presented first with a 'study list' (e.g., chloroform, felodipine, cisapride, benzoin, etc.) then with a 'test list' (e.g., cisapride, urea, nifedipine, chloroform, filgrastim, chloroquine, felodipine, benzoin, etc.). The words were presented to the pharmacists at a one word per second rate. When presented with the test list, the pharmacist's task was to say whether the current word was old (i.e., from the study list) or 'new' (not from the study list). We hypothesized that recognition errors would increase in frequency as similarity between test and study words increased. This is precisely what was observed (see Figure 3). The results were analyzed by way of a single-factor, repeated measures ANOVA, and similarity had a statistically reliable effect on error rate, $\underline{F}(4, 14) = 9.14$, $\underline{MSE} = 0.04$, $\underline{p} < 0.0001$.

The effect of phonological similarity was assessed in a parallel experiment. The stimulus materials for this experiment were 32 pairs of drug names, 8 each at 4 levels of similarity. The names were also drawn from the combined 1992-1994 NAMCS data. Names were again matched for frequency of

prescribing. The 32 pairs of names were arranged into 8 study lists and 8 test lists. Each study list had 8 words, and each test list had 16 words. The task, analysis plan, and hypotheses were in all other respects the same as those reported above. As predicted, similarity increased the recognition error rate significantly, $\underline{F}(3,11) = 5.39$, $\underline{MSE} = 0.02$, $\underline{p} < 0.004$. The error rate nearly doubled from low to high levels of similarity.

Discussion

This section describes how the experimental results are being translated into practical prevention and quality control/quality improvement techniques. The main prevention strategy is to 'error-proof' future drug names. To achieve this goal, the similarity measures described above can be used as the basis for computerized phonetic similarity searches of existing drug name databases. Proposed new names would be screened against existing databases. If a proposed name were too similar to an existing name, the proposed name would be changed or dropped. Both the FDA and the Institute for Safe Medication Practices are currently using search and retrieval algorithms designed by the author to screen proposed new names for confusion potential.[15] This pre-screening is a workable solution for new names, but it does not address the problems created by pairs of confusing names that are already in use. We have identified two strategies for reducing LASA errors among existing names. First, automated similarity measures could be built into computerized order entry and prescribing systems. Warnings could be issued when users enter one of a suspect pair of names. Second, additional theoretical insights from cognitive psychology could be used to modify routine dispensing practices. For example, it is well known that the phonological store in short-term memory is disrupted by one's own and by others' speech. An obvious implication of this finding is that talking should be kept to a minimum in areas where drugs are dispensed (just as unnecessary talk is minimized in cockpit situations). Similarly, pharmacists and nurses should not talk when they are dispensing or administering drugs. There are social obstacles to implementing these strategies, but they at least give the flavor of the type of prevention strategies that follow from examination of psychological theories of word recognition and memory.[5, 16, 17]

Conclusions

In a logical sequence of investigations, we have determined that orthographic and phonological similarity are identifiable causes of variation in the LASA error rate. By developing methods to identify and minimize this harmful variation, the we have begun to make long-lasting improvements in the stability,

predictability, and reliability of the entire drug use system.[19] Past discussions of drug safety have focused primarily on the safety of drug products in isolated organisms. By treating similarity as a kind of 'cognitive toxicity,' our work places drug safety in a larger social context.[18] Drugs must not only be safe within the confines of an isolated biological system, but also within the larger (social) drug use system. Among other things, this means assuring the safety of drug labeling, nomenclature, and packaging.

Acknowledgements

References

1.   Cohen MR. Drug product characteristics that foster drug-use-system errors. *American Journal of Health System Pharmacy*. 1995;52:395-399.
2.   Allan EL, Barker KN. Fundamentals of medication error research. *American Journal of Hospital Pharmacy*. 1990;47:555-571.
3.   U. S. Pharmacopeia. USP Quality Review. *Stop, look, and listen*. Rockville, MD: U. S. Pharmacopeia; 1995.
4.   Lambert B. Predicting look- and sound-alike medication errors. *American Journal of Health-System Pharmacy*. 1997;54:1161-1171.
5.   Baddeley AD. *Working memory*. Oxford: Oxford University Press; 1986.
6.   Shulman HG. Similarity effects in short-term memory. *Psychological Bulletin*. 1971;75:399-415.
7.   Schiff GD, Rucker TD. Computerized prescribing: Building the electronic infrastructure for better medication usage. *Journal of the American Medical Association*. 1998;279:1024-1029.
8.   Stephen GA. *String searching algorithms*. River Edge, NJ: World Scientific; 1994.
9.   Drewnowski A, Murdock BB, Jr. The role of auditory features in memory span for words. *Journal of Experimental Psychology: Human Learning and Memory*. 1980;6:319-332.
10.   Lambert BL, Lin S-J, Gandhi SK, Chang K-Y. Effect of similarity on pharmacists' recall and recognition of drug names. *Proceedings of the Annual Meeting of the American Association of Pharmaceutical Scientists*. San Francisco, CA: AAPS; 1998:(in press).
11.   U. S. Pharmacopeia. *USP DI, Vol. I: Drug information for the health care professional*. Rockville, MD: U. S. Pharmacopeia; 1995.
12.   Lambert BL, Gandhi SK, Lin S-W, Chang K-Y. Effect of orthographic and phonological similarity on pharmacists' recall and recognition of drug names. . Chicago, IL: University of Illinois at Chicago; 1998.
13.   Wallace WP, Stewart MT, Malone CP. Recognition memory errors produced by implicit activation of word candidates during processing of spoken words. *Journal of Memory and Language*. 1995;34:417-439.
14.   Woodwell DA, Schappert SM. National ambulatory medical care survey: 1993 summary. : National Center for Health Statistics; 1995.
15.   Lambert BL. Automated screening for look- and sound-alike medication errors. *FDA Workshop on Minimizing Medical Product Errors: A Systems Approach*. Bethesda, MD: U. S. Food and Drug Administration; 1998.
16.   Luce PA, Pisoni DB, Goldinger SD. Similarity neighborhoods of spoken words. In: Altmann GTM, ed. *Cognitive models of speech processing: Psycholinguistic and computational perspectives*. Cambridge, MA: MIT Press; 1990:122-147.
17.   Grainger J, Dijkstra T. Visual word recognition: Models and experiments. In: Dijkstra T, de Smedt K, eds. *Computational psycholinguistics*. Bristol, PA: Taylor & Francis; 1996:139-165.
18.   Lambert BL. Similarity as toxicity: Hazards associated with look-alike and sound-alike names. *Strategies for error-proofing drug labeling, packaging, and nomenclature: Industry's role in preventing product-related medication errors*. Princeton, NJ: Institute for Safe Medication Practices; 1998.
19.   Berwick DM. Continuous improvement as an ideal in health care. *The New England Journal of Medicine*. 1989;320:53-56.