Comments in Response to FDA's Questions about Methods and Approaches for Minimizing Drug Name Confusion (i.e., look-alike, sound-alike) Errors

Bruce L. Lambert, Ph.D.
Associate Professor
University of Illinois at Chicago
833 S. Wood St. (MC 871)
Chicago, IL 60612-7231
lambertb@uic.edu
Ph: 312-996-2411
Fx: 312-996-0868

Thank you for the opportunity to address the committee. Because I have only 9 minutes and because I addressed many of these same issues in my public comments during the June 26th meeting, I would like to direct the committee's attention to my previous testimony and PowerPoint presentation, both of which are available on the FDA's web site or from me directly. In addition, I have submitted to the committee reprints of several peer-reviewed articles published by my colleagues and me during the past 7 years. Although it is not possible to summarize the main findings of those articles in the time allotted, each article presents evidence that is directly relevant to the questions being debated today. In fact, they are, to the best of my knowledge, the *only* peer reviewed studies that provide evidence as to the validity of computer-based methods for drug name screening.

To paraphrase a cliché from the domain of real estate, when it comes to regulatory acceptance of new test methods, there are only three issues to be concerned about: validation, validation, validation. Before a new testing method can be accepted by a regulatory agency, it must be scientifically validated. Validation alone is not enough to warrant regulatory acceptance. But without validation, acceptance ought to be out of the question. As I prepared these remarks, it occurred to me that regulatory agencies must constantly need to evaluate new testing methods. I felt certain that there would be standard methods for establishing the validity of newly developed testing methods. I was both right and wrong about this. On the one hand, there are no uniform policies for validation and regulatory acceptance of new testing methods across government agencies. EPA, FDA, USDA, NIOSH, and others each have their own approaches.

On the other hand, recognizing this lack of coordination both within the US and internationally, toxicologists and regulators from around the world have worked over the last decade to develop a standard approach to the validation and regulatory acceptance of new testing methods. The ad hoc Interagency Coordinating Committee on the Validation of Alternative Methods (ICCVAM) is a US governmental body run out of the National Institute for Environmental Health Sciences. Together with a similar group in Europe and from the OECD, the ICCVAM has developed clear guidelines for validation and regulatory acceptance of new tests. These were developed in the context of traditional toxicology, with a special focus on finding new alternatives to animal testing, but the overall framework should apply more generally to all validation and regulatory acceptance situations. I strongly encourage the committee, the audience, and the Agency to study these guidelines. It is my recommendation that these guidelines be followed in validating and determining the acceptability of new tests on the

confusability of drug names. If they are not accepted, I would request that the agency spell out its own guidelines for validation and acceptance, and I would also request the Agency's rationale for not adopting an existing framework that has proved successful elsewhere.

What can be learned from the FDA's own Redbook?

Validation Criteria

I will first summarize the ICCVAM's main criteria for validation. *Validation* is a "scientific process designed to characterize the operational characteristics, advantages, and limitations of a test method, and to determine its reliability and relevance." Furthermore, "validation is a process by which the reliability and relevance of a test method are evaluated for the purpose of supporting a specific use." The ICCVAM criteria for validation are as follows:

(Refer to ICCVAM Report)

Regulatory Acceptance Criteria

(Refer to ICCVAM report)

These criteria are sensible and represent the consensus of an international group of experts. They also have some status as policy within the federal government, although individual agencies are not bound by them. Again, I recommend they be adopted in this context, and if they are not, I request that the Agency's own criteria for validation and regulatory acceptance be published.

None of the methods discussed here today meet these criteria, although mine come closest, as evidenced by the extensive validation studies published in peer-reviewed scientific journals. The methods described this morning by Dr. Dorr and currently being used by FDA are likely to be sound, but they have not been validated in peer-reviewed journals, nor have all their operational details been publicly disclosed. I recommend that no method be accepted for regulatory use until it is adequately validated in accordance with the criteria set out above.

Problems With Validation of Name Screening Methods

1. **Lack of a gold standard.** Direct observation of orders, fills, and administrations would be best in terms of realism (ecological validity). An NDC-style field trial would be second best. High fidelity simulation next, followed by traditional psychological experimentation, computer testing, and panels of experts.
2. **Problems with true positives and true negatives.** Some names appearing in reporting databases are near misses not actual errors. They thus have questionable status as true positives. Names not appearing in reporting databases may in fact have been involved in errors but not reported. Absence of evidence is not evidence of absence. The USP list is not a gold standard; it is an iron pyrite standard. Iron pyrite is also known as fool's gold. Since any test method will be validated by assessing its ability to distinguish between truly confusing and truly non-confusing names, the ambiguity around true positives and true negatives is highly problematic. Related to this is the need for the proportion of true

positives and true negatives in the test sample to be the same as the real-world proportion of true positives and negatives. But we don't know these real-world proportions.

3. **Wrong unit of analysis.** Much of the work on computer methods for name screening, including my own early work, has focused on pairs of names. But FDA must approve single names, not pairs of names. Methods are needed that use the single name as the unit of analysis, not the pair of names. Any pair based method will have poor positive predictive value because of the sheer number of pairs (N-squared minus N over 2) and the inevitable false positive rate of any predictive test.
4. **No attention to frequency.** Frequency is a fundamental mechanism in human error, but it is absent from most discussion about name confusion. There has been too much focus on similarity. Similarity measures are symmetric (A to B = B to A), but errors are not. If a rare name is similar to a common name, the common name will be seen when the rare name is presented, but the opposite will almost never occur. We must build prescribing frequency into our predictive models.
5. **Non-name attributes.** These contribute to errors but we don't know precisely how.
6. **Conflicts of interest.** A lot of money is at stake in naming decisions. We need to make sure those doing the safety screening do not have a vested interest in the outcome of the screening.
7. **Public costs vs. private benefits in the name approval decision.** Normally FDA weighs risks and benefits in drug approval decisions, but here the costs are to the public and the benefits are to the manufacturers.
8. **Harm reduction as the ultimate goal.** When evaluating a proposed name, we need to think not just about the probability of error but also about harm. Harm is a complex function of probability of error, number of opportunities for error, severity of each error, probability of not detecting the error, etc. It also depends on patient status, time of exposure, time without intended medication, concomitant medications, etc.
9. **Acceptable thresholds for sensitivity, specificity, positive and negative predictive value are not known.**
10. **Flynn, Barker and Carnahan have reported that the wrong drug error rate in outpatient pharmacy, by direct observation, is 0.13% (6/4481). That's 3.9 million wrong drug errors per year or about 65 per pharmacy annually or about one per week in every pharmacy in the US.**
11. **It's not all about names.** We are stuck with lots of confusing names that are on the market and that are not going to be changed. Only safe medication practices like putting the indication on the prescription, restricting verbal orders, and using CPOE will protect us from these names. Restricting formularies.

Thank you for your attention.