

Why No Label Comprehension Standards?

As companies prepare to discuss switching products from prescription (Rx) to nonprescription (OTC) status, the Food and Drug Administration (FDA) asks them to (at least) demonstrate that the label is well comprehended. However, there is no objective way to make this determination? Why is that?

I have been involved in label comprehension testing for switches for 10 years, and long have been puzzled by the lack of any objective means to clearly say “this label is well comprehended”, or “this label is not well comprehended”. Creating some sort of guideline, rule or standard should not be that difficult a task, but, apparently it is.

When first examining this area of research, a danger became evident. Suppose, in some fashion, one were able to determine that 85% of all consumers who may, at one time or another, have occasion to use an OTC product understand the label. Is that acceptable? It legitimately could be argued that, since 15% of potential users do not understand it, the label needs to be improved. However, based on marketing research experience, it is certain that there will never be a label that 100% of consumers understand.

The challenge, then, is to find a way to interpret a number, which is likely to be meaningless. Without a point of comparison, *any* result obtained in a label comprehension study can be called inadequate (or satisfactory, for that matter). If a situation arises in which the FDA looks at the results and says the label needs improvement and retesting, while the sponsor claims the results indicate adequate comprehension, it seems the research does not contribute much to the development of a well understood label.

Upon reflection, it seems an objective point of reference can be established by using an existing, similar label as a control, or “comparator”, to provide an objective measure to compare the test label results. The conclusion, based upon statistical testing, would be that the proposed label is comprehended as well as, better than, or not as well as the control label. Testing becomes objective.

This “standard” is supported, at least in theory, by the Nonprescription Drug Manufacturers Association which is now the Consumer Healthcare Products Association, (CHPA, which has long argued that existing OTC labels are sound, as evidenced by the very small percentage of the public that has difficulty with safe and appropriate use of OTC products. While stating this, the group agreed with the FDA that labels could be improved, and both organizations have taken steps in recent years to make certain that improvement occurs.

The use of an existing label as a comparator has a number of advantages. First, the FDA can avoid any charge of being arbitrary or capricious in interpreting label comprehension studies. There are objective data to support a judgment. Second, the sponsor need not worry about any number being viewed as “bad”, unless the data indicate consumers do not understand the proposed label as well as they understand an existing, approved label. In this worst case, sponsors are likely to improve the label and retest it before presenting the evidence to the FDA for approval.

Regulatory Corner

The major disadvantage is that it may be impossible to find a comparator. More likely, this will be a creative challenge. For example, there was once a proposed switch product that had warnings not to take it before speaking with a doctor if one had a heart condition or liver damage. Testing was constructed so that an existing cold product, the label for which included an identical heart condition warning, was used as a comparator for one group. In addition, a currently existing antidiarrheal product, the label for which included the same liver damage warning, was used as a comparator for the other group. Products were blinded, so consumers could not use knowledge of the known brands to assist them. In the end, the sponsor was able to say:

Heart condition patients comprehended the proposed warning as well as they comprehended a similar warning on an existing label.

Further, it could be stated that:

Those with a liver condition comprehended the proposed label's warning as well as they comprehended a similar label warning on an existing OTC product.

Some argue that, since the three products involved were from three different OTC categories, the testing protocol was illogical. However, since in this case the concern was with the comprehension of warnings addressed to

specific at-risk groups, the research conclusions, as stated, are entirely accurate. It was known to all that few, if any, consumers with heart conditions had heart problems as a result of using the cold product, and that few with liver conditions had health problems due to use of the antidiarrheal product. Thus, a link to what might actually occur is created. Finally, because of the comparator, the percent of consumers comprehending the proposed label does not exist in a vacuum, calling for subjective judgment.

A second, broader example comes to mind. In testing to determine whether potential consumers understood the nature of and how to use a potential

switch product in skin patch form, a search was made through the Nonprescription Physicians Desk Reference to find another OTC product in that form. The target population for the switch product was a subgroup of the adult population, while the other skin patch was labeled for use by anyone in the adult population.

A comprehension test for usage, with identical questions, was administered to the two populations, with each reading the appropriate blinded label before being tested. The test was composed of multiple choice questions, which were constructed so that the answers were *identical*, with the only difference being which answer was correct. For example, a question might read:

"This patch product is: (a) only for men, (b) only for women, (c) for both men and women."

The switch product would be for use only by men, as described in the label. The other patch product would be for both men and women, again as described. For those who read the switch product label, the correct answer is "(a) only for men". For those who saw the other label, it is "(c) for both men and women". The conclusion from this test was that:

Potential users of the proposed switch product understand the nature and use of that product from the label, as well as the potential users of the existing product understand the nature of and use of that product from the label.

While the populations differ, the conclusion is accurate, and is helpful in objectively determining the comprehensibility of the proposed label.

The notion of a comparator has been public since at least 1995, and, unfortunately, is not well accepted. The belief exists (correctly or incorrectly) that the FDA does not want existing labels to be tested. The result is that decisions about label comprehension are still made in a somewhat arbitrary fashion (e.g., witness the Flexeril conclusions). One has to wonder why.

Stephen J. Hellebusch is president of Hellebusch Research & Consulting, Inc. in Cincinnati, OH. He can be reached at 800.871.6922 or email: steve@hellrc.com