nature medicine

# Critical comparisons

**The US Food and Drug Administration (FDA) is increasingly conservative in its assessment of data in new drug applications derived from noninferiority trials. Such caution might be similarly warranted in approving generic biologics.**



It's not easy being the FDA. The public demands that this regulatory agency deliver safe and effective drugs quickly and efficiently. Some additional consistency in its decision-making might also be in order.

Take, for example, the FDA's stance on the use of noninferiority trials versus that on the approval process for generic biologics. Noninferiority clinical trials are designed to compare an investigational new drug against an approved drug for the same indication. The rationale behind such comparisons is that a new drug might be safer, cheaper, more conveniently formulated or with fewer side effects than an existing drug. And, provided that, in appropriately designed trials, the new drug is not worse than the existing drug—within a certain margin—the FDA will use this information in its evaluation of a New Drug Application (NDA), in conjunction with data from other types of clinical trials. In some instances, however, data from noninferiority trials have provided the crucial evidence resulting in a new drug's approval.

But showing that a new drug is 'not inferior'—that is, comparable—to an approved drug is quite challenging. To do so, a margin for noninferiority must first be determined for each trial. Similar to a margin of error, this margin is the maximum amount by which a new drug can be shown to be *less* effective than the active comparator drug yet still be deemed clinically effective. Such margins have ranged from 5% to 20%, and the FDA considers their selection the most difficult task in designing a noninferiority trial. If the margin is too small, a useful drug may seem ineffective, whereas if the margin is too large, a poor drug may falsely seem effective. Over time, noninferiority trials may also lead to 'biocreep', in which the comparison of progressively inferior drugs could result in the approval of treatments with little to no efficacy.

With these issues in mind, the FDA issued in March new guidelines to pharmaceutical companies on the use and design of noninferiority trials. And, in July, the US Government Accountability Office (GAO) published a review of the FDA's use of such trials in approving NDAs submitted between 2002 and 2009. Interestingly, the GAO found that the FDA has become more cautious over time in granting approval on the basis of data from noninferiority trials. Since 2007, the FDA has restricted the use of noninferiority trial data to drugs for severe indications, eliminating its use in the context of certain less severe microbial infections for which there is evidence of resolution in the absence of intervention (for example, acute bacterial ear infections). The GAO also found that the FDA has become stricter in assessing data derived from noninferiority trials. Between 2002 and 2009, of 43 NDAs containing data from at least one noninferiority trial, the FDA approved 18 on the basis of pivotal evidence from noninferiority trials but found nine NDAs had poorly designed noninferiority trials that could not be used to establish the efficacy of the new drug. And, reassuringly, the GAO found no evidence of biocreep during this period. Overall, the GAO's report indicates that the FDA has an exacting stance in its assessment of the contribution of noninferiority trials to the approval of NDAs.

But then there is the case of approvals of generic biologics. In July, under the Abbreviated New Drug Application (ANDA) program, the FDA approved enoxaparin sodium injection as a generic version of the anticoagulant Lovenox (Sanofi-Aventis) without a clinical trial to support its safety or efficacy in patients. Enoxaparin, made by Sandoz, is the seventh generic biologic approved in this manner. Generic biologics approved by an ANDA can be substituted clinically for the reference-listed drug. But, unlike a generic small molecule, a generic biologic is not necessarily chemically equivalent to an approved brand. Whereas both enoxaparin and Lovenox are derived from heparin, neither is chemically defined.

The FDA required that Sandoz show that enoxaparin has the same active ingredient as Lovenox as defined by a series of *in vitro* and *ex vivo* tests, as well as by its pharmacodynamic profile in healthy individuals. But it did not require a direct comparison of safety and efficacy to Lovenox in clinical trials. In 2003, Sanofi-Aventis filed a citizen's petition requesting that the FDA withhold its approval pending clinical evidence of equivalent safety and efficacy of enoxaparin with Lovenox, but the FDA denied this request.

Although the FDA is convinced that the criteria it applied to the approval of enoxaparin provide sufficient evidence of its safety and efficacy, the use of untested mixtures of compounds in patients is inherently risky. The classic example of the unpredictable effects of a mixture of molecules is thalidomide, used by pregnant women in the 1950s for morning sickness. Although thalidomide is chemically defined, it is a racemic mixture, and the *S* enantiomer is teratogenic. More recently, a group at Genentech reported that changing the site of conjugation of a cytotoxic drug to a therapeutic antibody altered its safety profile *in vivo* (*Nat. Biotechnol.* **26**, 925–932, 2008), underlining the biological heterogeneity inherent in the generation of complex biomolecules and the need to consider its potential clinical consequences.

The enoxaparin approval establishes a precedent that might one day be applied to complex agents, such as antibodies. It therefore raises the question of how generic biologics will be assessed in the future and how to determine the cutoff for the biological complexity or heterogeneity of an agent beyond which a clinical comparison—and not only an *in vitro* comparison—will be required to establish efficacy and safety.

Although generic drugs warrant an accelerated approval process, in view of the inherent heterogeneity of biologics, a clinical demonstration of efficacy, or of noninferiority, seems a reasonable regulatory requirement to impose.