and immunogenicity profiles. For example, sponsors should consider available public information, as well as the sponsor's characterization of the reference product, in determining the potential impact of an attribute on clinical performance.

- *The degree of uncertainty around a certain quality attribute*—For example, when there is limited understanding of the clinical impact of an attribute, the FDA recommends that that attribute be ranked as having higher risk because of the uncertainty involved.

The FDA recommends that an attribute that is a high risk for any one of the performance categories (i.e., activity, PK/PD, safety, or immunogenicity) should be classified as high risk. Ideally, the risk assessment tool should result in a list of attributes ordered by the risk to the patient. The risk scores for attributes should, therefore, be proportional to patient risk. Because there may be a limited number of attributes that can be evaluated with equivalence testing (see Section 9.3.1.2), attributes that are known to be of high risk to patients (i.e., high impact attributes) should be a priority over attributes with unknown but potentially high risk (i.e., attributes with a high-risk ranking due to uncertainty).
The scoring criteria used in the risk assessment should be clearly defined and justified in the analytical similarity assessment plan, and the risk ranking for each attribute should be justified with appropriate citations to the literature and data provided.

9.3.1.2 Determination of the statistical methods to be used

The FDA's current approach to evaluating analytical similarity is to define three tiers corresponding to the use of three different methods for comparing attributes. The FDA believes that the use of these three tiers with appropriate similarity acceptance criteria should help support a demonstration that the proposed biosimilar is highly similar to the reference product. Equivalence testing (Tier 1) is typically recommended for quality attributes with the highest risk ranking and should generally include assay(s) that evaluate clinically relevant mechanism(s) of action of the product for each indication for which approval is sought. The use of quality ranges (Tier 2) is recommended for quality attributes with a lower risk ranking, and an approach that uses visual comparisons (Tier 3) is recommended for quality attributes with the lowest risk ranking. The three methods are described in Section 9.3.2.

In addition to risk ranking, however, other factors should be considered in determining which tier of statistical evaluation should be applied to a particular attribute or assay. Although many attributes may be considered high risk, subjecting all of these attributes to Tier 1 testing may result in a false negative conclusion (i.e., a determination that a product is not highly similar when it truly is). Some additional factors, besides risk, that should be considered when determining the appropriate tier include