Journal of Pharmaceutical and Biomedical Analysis 49 (2009) 613–618

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

Experimental and statistical approaches in method cross-validation to support pharmacokinetic decisions

Theingi M. Thway[a,*], Mark Ma[a], Jean Lee[a], Bethlyn Sloey[a], Steven Yu[a], Yow-Ming C. Wang[a], Binodh DeSilva[a], Tom Graves[b]

[a] Department of Pharmacokinetics and Drug Metabolism, Amgen Inc., One Amgen Center Drive, Thousand Oaks, CA 91320, United States
[b] Department of Biostatistics- Medical Sciences, Amgen Inc., One Amgen Center Drive, Thousand Oaks, CA 91320, United States

ARTICLE INFO

Article history:
Received 9 September 2008
Received in revised form 6 November 2008
Accepted 5 December 2008
Available online 13 December 2008

Keywords:
Cross-validation
LBA
Sample size selection
PK profiles (AUC
C_{max})
Statistical analysis

ABSTRACT

A case study of experimental and statistical approaches for cross-validating and examining the equivalence of two ligand binding assay (LBA) methods that were employed in pharmacokinetic (PK) studies is presented. The impact of changes in methodology based on the intended use of the methods was assessed. The cross-validation processes included an experimental plan, sample size selection, and statistical analysis with a predefined criterion of method equivalence. The two methods were deemed equivalent if the ratio of mean concentration fell within the 90% confidence interval (0.80–1.25). Statistical consideration of method imprecision was used to choose the number of incurred samples (collected from study animals) and conformance samples (spiked controls) for equivalence tests. The difference of log-transformed mean concentration and the 90% confidence interval for two methods were computed using analysis of variance. The mean concentration ratios of the two methods for the incurred and spiked conformance samples were 1.63 and 1.57, respectively. The 90% confidence limit was 1.55–1.72 for the incurred samples and 1.54–1.60 for the spiked conformance samples; therefore, the 90% confidence interval was not contained within the (0.80–1.25) equivalence interval. When the PK parameters of two studies using each of these two methods were compared, we determined that the therapeutic exposure, $AUC_{(0\text{-}168)}$ and C_{max}, from Study A/Method 1 was approximately twice that of Study B/Method 2. We concluded that the two methods were not statistically equivalent and that the magnitude of the difference was reflected in the PK parameters in the studies using each method. This paper demonstrates the need for method cross-validation whenever there is a switch in bioanalytical methods, statistical approaches in designing the cross-validation experiments and assessing results, or interpretation of the impact of PK data.

1. Introduction

Few pharmacology studies are typically conducted to investigate the pharmacokinetic (PK) parameters during the lead optimization period in the development of a therapeutic biologic. Limited time and availability of reagents during this phase often constrain the initial method so that only readily available reagents are used. Later, methods may be refined after the availability of anti-idiotypic antibodies. Thus, it is common to apply multiple ligand binding assay (LBA) methods to quantify therapeutic biologic concentrations in various preclinical studies. During pilot studies, it is not uncommon to use methods that are not fully validated (i.e., they are qualified with less rigorous validation experiments [mini-validation]) based on levels of accuracy and precision, and minimal stability [1–4]. In contrast, a refined method would be fully validated [1–4]. Method comparison experiments are usually performed during the transition from pilot studies. Before conducting these experiments, goals of cross-validation should be determined based on the need to compare the results generated from two methods, each from a mini-validation and a full validation.

Cross-validation is a comparison of two or more methods that are used to generate data within the same study or across different studies [5,6]. The Conference Report on Bioanalytical Method Validation – A Revisit with a Decade of Progress was published in 2000 and provided a guideline for performing cross-validation when two or more bioanalytical methods are used to generate data within the same study. An original validated bioanalytical method is considered as "reference" and the revised method is the "comparator." On the regulatory side of therapeutic development, the Guidance for Industry on Bioanalytical Method Validation issued by the FDA clearly states the requirement for cross-validation in the following scenarios: (1) when sample analyses within a single study are conducted at more than one site (cross-validation is required in addition to partial-validation) and (2) when data

* Corresponding author. Tel.: +1 805 313 6389; fax: +1 805 499 9027.
E-mail address: tthway@amgen.com (T.M. Thway).

0731-7085/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2008.12.008

generated using different analytical techniques (e.g., LC/MS–MS vs. LBA) in different studies are included in regulatory submissions [5,6]. Nonetheless, there has been no further guidance on cross-validation of methods or on data interpretation. It is not clear to many bioanalytical laboratories whether the changes in capture reagents and/or detection systems represent "major" changes that would require cross-validation, and these issues have not been adequately addressed in the literature [6]. Some statistical approaches on method comparison between original and new methods have been discussed for daily clinical/diagnostic use; however, there have been no publications providing experimental or statistical discussion about pharmaceutical use intended to support PK and/or toxicokinetic (TK) assessments [6,8]. In this paper, we present a case study on cross-validation of two similar LBA differing in the capture reagents and the detection systems. The cross-validation concept was extended to evaluate the equivalence of two LBA methods with different platforms rather than the different analytical techniques using a variance analysis statistical approach.

2. Methods

2.1. Validation plan and documentation

Cross-validation of two LBA methods is often conducted during method validation of the more recent method. A standard operating procedure would typically be followed if there were one in place. Alternatively, a cross-validation plan can be written before the initiation of the cross-validation experiments. We prepared an *a priori* cross-validation plan, which included a detailed background of the methods, experimental design, and selection of test sample sizes for method comparison, as well as restrictive assay conditions to minimize the random error, *a priori* acceptance criteria for bioanalytical equivalence, and a description of the statistical analysis to be performed.

2.2. Bioanalytical methods

Two chronological methods were developed to support PK studies at different stages of drug development. Method 1, applied to an early study, was a chemiluminescence-based LBA in which the therapeutic target protein was used as the capture reagent. Method 2 was a colorimetric enzyme linked immunosorbent assay (ELISA) in which a monoclonal anti-therapeutic antibody was used as a capture reagent. The same detector reagent, horseradish peroxidase (HRP)-conjugated mouse anti-human antibody specific to Fc, was used in both methods. Method 2 was fully validated to support a regulated study according to the FDA guidelines [5] while method 1 was qualified (i.e., mini-validation).

Method 1 Procedure: Microplate wells were coated with biotinylated therapeutic target protein (Amgen Inc., CA). After blocking with 1X PBS with 1% BSA (blocking buffer), standards (STD), quality controls (QC), conformance-spiked samples (made by spiking therapeutic antibody into 100% Cynomolgus monkey serum), blank, and incurred samples were loaded into the wells after pre-treatment at 1:5 with blocking buffer. The therapeutic antibody present in the STD, QC, and samples bound to the immobilized mouse therapeutic target protein. After a wash step, a horseradish peroxidase (HRP conjugated mouse anti-human Fc antibody specific to Fc (Amgen Inc., CA) was added to the wells. The detector antibody was bound to the therapeutic antibody captured in the previous step. After a final wash step, a Pico peroxide substrate (Pierce, Inc. IL) was added to the well. The Pico substrate reacted with HRP and produced a light signal that was proportional to the amount of therapeutic antibody bound by the capture reagent. The intensity of the light output (relative light unit, RLU) was measured using the Molecular Devices Spectra LMAXII 384 Luminometer equipped with SOFTmax Pro software. The conversion of RLU for QC and study samples to concentration was achieved through a computer software mediated comparison to a standard curve assayed on the same plate, which was regressed according to a logistic (Auto-Estimate) regression model with a weighting factor of $1/Y$ using the Watson data reduction package.

Method 2 Procedure: Microplate wells were coated with mouse a monoclonal anti-therapeutic antibody (Clone No. Ab 1.9.1, Amgen Inc., CA). After blocking with 1X PBS with 1 M NaCl, 1% BSA and 0.5% Tween 20 (blocking buffer), STD, QC, conformance-spiked samples (made by spiking therapeutic antibody into 100% Cynomolgus monkey serum), blank, and incurred samples were loaded into the wells after pre-treatment at 1:25 with blocking buffer. The therapeutic antibody present in the STD, QC and samples bound to the immobilized anti-therapeutic antibody. After a wash step, an HRP-conjugated mouse anti-human Fc antibody specific to Fc (Amgen Inc., CA) was added to the wells. The detector antibody bound to the therapeutic antibody captured during the previous step. After a final wash step, a tetramethylbenzidine (TMB) peroxide substrate solution (KPL Inc., MD) was added to the wells. TMB in the substrate solution reacted with the peroxide and, in the presence of HRP, produced a colorimetric signal that was proportional to the amount of therapeutic antibody bound by the capture reagent in the initial step. The color development was stopped by acidification and the intensity of the color (optical density, OD) was measured at 450 nm minus 650 nm using a Molecular Devices Spectra max 340PC microtiter plate reader equipped with SOFTmax Pro. The conversion of OD units for the validation samples and the QC to concentration was achieved through a computer software mediated comparison to a standard curve assayed on the same plate, which was regressed according to a logistic (Auto-Estimate) regression model with a weighting factor of $1/Y^2$ using the Watson data reduction package.

A summary of the method formats is described in Table 1, and assay acceptance criteria for each method are listed in Table 2.

2.3. Statistical methods/approaches

2.3.1. Sample size selection

The statistical estimation for sample size described in Table 3 was generated using nQuery Advisor 5.0 sample size determination software from Statistical Solution Inc. (Saugus, MA). Sample sizes were generated for different precision (computed as %CV) and equivalence intervals assuming no initial bias between the methods. Table 3 was used to determine the test sample size in cross-validation of two methods based on the inter-assay precision in relation to the desired equivalence interval. A common practice of equivalence interval of 80–125 was used to determine if two methods were equivalent within 90% confidence interval. If the imprecision of the 2 methods is different, the higher precision of the two methods was used to determine the number of samples to be analyzed. Based on the inter-assay precision for Method 1 and desired equivalence interval of 80% to 125%, 30 samples were tested.

2.3.2. Equivalence analysis for methods 1 and 2

All data analysis was performed using SAS V9.1.3 on a Windows Professional operating system. An analysis of variance with terms for sample and assay was performed on the log responses. A 90% equivalence confidence interval was computed for the ratio of the different estimated concentrations using an estimate statement in the log-transformed value analysis, computing a confidence interval of the differences, and back-transforming (exponentiation) the difference and confidence interval to get the ratio and its confi-