Data Submission Standards and Evidence Requirements

JEFFREY ABRAMS,^a ROBERT ERWIN,^b GWEN FYFE,^c RICHARD L. SCHILSKY^d

^aNational Cancer Institute, Bethesda, Maryland, USA; ^bMarti Nelson Cancer Foundation, Vacaville, California, USA; ^cConsultant, San Francisco, California, USA; ^dUniversity of Chicago, Chicago, Illinois, USA

Disclosures: Jeffrey Abrams: None; Robert Erwin: *Consultant/advisory role:* Member CALGB DSUB (unpaid); member ASCO Research Committee (unpaid); Gwen Fyfe: None; Richard L. Schilsky: None.

The content of this article has been reviewed by independent peer reviewers to ensure that it is balanced, objective, and free from commercial bias. No financial relationships relevant to the content of this article have been disclosed by the independent peer reviewers.

IMPORTANCE OF STREAMLINING DATA COLLECTION

The goal of the U.S. Food and Drug Administration (FDA) guidance documents is to provide insight into the data necessary for FDA reviewers to reliably assess the risk–benefit ratio of an investigational agent for a particular clinical indication. The current FDA registration guidance for cancer therapy trials does not completely describe the level of detail necessary for informative data capture to support claims of safety and efficacy for supplemental indications of new cancer treatments [1]. The guidance, as currently set out, does not distinguish drugs with substantive safety information and a definite benefit to patients from drugs with limited safety data that may carry safety risks that have not yet been recognized. Data collection requirements thus become essentially the same whether for a primary indication or a supplemental application. This can result in collection of excessive and sometimes unnecessary data by investigators, particularly for trials designed to explore additional indications where substantial toxicity data about an agent already exist. Further, because there is no established standard for collection of data in support of supplemental applications, sponsors interpret the requirements variably, resulting in inconsistent quality and quantity of data. Frequently, the data collected do not result in modifications to FDA labeling or inform medical practice, yet the data collection requirements add complexity and cost to conducting

the study. Therefore, optimized standards for data collection should be developed for well-studied cancer therapies to improve the efficiency of safety evaluations without sacrificing the scientific integrity and validity of study results.

Streamlining data collection will help ensure better patient safety by improving the overall quality of data submitted in supplemental applications. Collecting essential data that will help inform patient safety, such as toxicities leading to death or dose discontinuations, is more important than collecting large amounts of data, such as cataloging all mild adverse events, that ultimately add little information to the existing safety profile of the drug. Collection of unused data may actually distract from gleaning crucial information. When faced with large amounts of safety data, it becomes difficult to prioritize safety events, distracting sites from focusing on the collection of important information, such as understanding what makes physicians or patients modify or stop treatment. Thus, large amounts of data can sometimes obfuscate knowledge of new and relevant safety data. Furthermore, streamlining data collection will greatly reduce the administrative burden on the clinical trial system and will focus finite resources on collecting key data elements. Reducing burdensome and unnecessary data collection will improve physician participation in clinical trials. Surveys to understand why patients do not participate in clinical trials reveal that doctors often do not recommend clinical trials to their patients. Among various other rea-

sons, doctors cite that they are weary of the high administrative workload and liability associated with conducting clinical trials. In an effort to understand the burden of excessive data collection on trial administrators, a working group, resulting as an outgrowth of the 2008 Conference on Clinical Cancer Research and formed under the aegis of the American Society of Clinical Oncology (ASCO), solicited input from several cooperative group and industry sites. Of 110 responses received to the poll, >85% expressed the view that data optimization (as recommended below) would moderately or significantly impact site resources, allowing collection of higher quality targeted data and greater participation in the clinical trials process [2].

POTENTIAL TRADEOFFS OF DATA OPTIMIZATION

In order to further explore the tradeoffs between complete and optimal data collection, the Data Optimization Working Group assessed the extent of safety data collection necessary and sufficient to inform clinical and regulatory decisions in a supplemental application with the basic assumptions that:

- Streamlined toxicity data collection will not be used for initial indications (or the first supplemental application following accelerated approval).
- Streamlined toxicity data collection will be used only if the prior approval process included a safety database that was acceptable for a full regulatory approval.
- The statistical analysis plan will be structured to minimize the risk of missing important safety signals.
- Data on serious adverse events (SAEs), deaths, and dose modifications and/or discontinuations (with reasons) will be collected for all patients on all study arms.
- Data on targeted adverse events (AEs) would be collected based on the known safety profile and pharmacology of the drug and the study patient population.

Streamlining data collection will ensure that the data collected will be used and that unnecessary data will not be collected. Data collection requirements will vary as necessary depending on whether a sufficiently large safety and drug interaction profile already exists, the similarity of the study population to the population for approved use, the similarity of the supplemental regimen to the regimen already approved and, finally, whether the supplemental application follows initial full or accelerated approval. By collecting data on SAEs, deaths, dose modifications and/or discontinuations, and targeted AEs of interest in all patients on all study arms, sponsors are reasonably as likely to detect important safety signals as with the current data collection process.

STUDY ORGANIZATION AND PARTICIPANTS

At the Conference on Clinical Cancer Research held in September 2008, a panel on Data Submission Standards and Evidence Requirements proposed a framework for data collection necessary to support claims of safety and efficacy for supplemental new drug applications (NDAs) and biologic license applications (BLAs) [3]. In order to further explore elements of that framework, ASCO formed the Data Optimization Working Group. The Working Group provided a forum for all interested stakeholders (the FDA, the National Cancer Institute [NCI], academia, industry, and advocacy) to retrospectively review data sets from completed phase III trials, many that were used for FDA supplemental approvals, and discuss potential revisions to data collection standards.

Four companies and one cooperative group collaborated on this project. A statistical analysis plan was developed, reviewed by the FDA, and used by all participating sponsors. The project involved a reanalysis of eight trials, in both the metastatic and adjuvant settings, studying cytotoxic chemotherapy, targeted biological therapy, and hormonal therapy, as shown in Tables 1 and 2 [4].

STUDY FINDINGS AND RECOMMENDATIONS

The purpose of this study was to determine whether important safety information would be lost by only gathering toxicity data on a subsample of patients enrolled in a supplemental NDA trial with a drug for which a substantial toxicity profile already exists. In candidate trials where subsampling is appropriate, it is assumed that SAE information, including all deaths, dose discontinuations, and dose modifications along with the associated reasons, would continue to be collected on all patients. The reanalysis demonstrated that data subsampling did not appear to omit important information about the safety profile, that is, similar conclusions regarding the safety profile would have been reached if a subsampling approach had been used.

The study identified statistical methods for determining appropriate subsampling sizes that can be scaled to fit different cutoff rates. The subsampling size range recommendations using this statistical methodology are as follows.

For determining subsampling size (assuming a 2% excess and a two-arm trial):

- In the metastatic setting, approximately 400–500 patients should be subsampled (full study size, 800–1,200 patients).
- In the adjuvant setting, a total size of approximately 400–900 patients should be subsampled (full study size, 800–6,000 patients).

Table 1. Candidate trials evaluated in the study in the metastatic setting

Company, candidate study	Patient population	Treatment	Trial size	Primary endpoint
Genentech, AVF2107g	First-line mCRC	Arm 1, bolus IFL plus placebo; arm 2, bolus IFL plus rhuMab VEGF; arm 3, 5-FU and LV plus rhuMab VEGF	813	OS
Genentech, ECOG 4599	First-line nonsquamous NSCLC	Arm 1, paclitaxel and carboplatin; arm 2, paclitaxel, carboplatin, and bevacizumab	878	OS
Genentech, AVAIL	First-line nonsquamous NSCLC	Arm 1, cisplatin and gemcitabine; arm 2, cisplatin, gemcitabine, and bevacizumab	656	PFS
GSK, EGF 30001	Metastatic breast	Arm 1, paclitaxel and placebo; arm 2, paclitaxel and lapatinib	580	TTP
Eli Lilly and Co., JMDB	First-line NSCLC	Arm 1, cisplatin plus pemetrexed; arm 2, cisplatin plus gemcitabine	1,669	OS

Abbreviations: 5-FU, 5-fluorouracil; AVAIL, Avastin in Lung; bolus-IFL, irinotecan, 5-FU, and LV; ECOG, Eastern Cooperative Oncology Group; JMDB, H3E-MC-JMDB; LV, leucovorin; mCRC, metastatic colorectal cancer; NSCLC, non-small cell lung cancer; OS, overall survival; PFS, progression-free survival; rhuMab VEGF, recombinant human monoclonal antibody vascular endothelial growth factor; TTP, time to tumor progression.

Table 2. Candidate trials evaluated in the study in the adjuvant setting

Company, candidate study	Patient population	Treatment	Trial size	Primary endpoint
Novartis, BIG 1-98	PMP women with HR ⁺ EBC	Arm 1, letrozole; arm 2, tamoxifen	8,028	DFS
CALGB, 89803	Patients with resected adenocarcinoma of the colon	Arm 1, LV and 5-FU (500); arm 2, irinotecan, LV (20 mg/m ²), and 5-FU	1,264	OS
Genentech, HERA	HER-2 ⁺ adjuvant breast cancer	Arm 1, observation; arm 2, trastuzumab	3,386	DFS

Abbreviations: 5-FU, 5-fluorouracil; BIG, Breast International Group; CALGB, Cancer and Leukemia Group B; DFS, disease-free survival; EBC, early breast cancer; HER, human epidermal growth factor receptor; HERA, Herceptin[®] Adjuvant; HR, hormone receptor; LV, leucovorin; OS, overall survival; PMP, postmenopausal.

- Subsampling may not be appropriate or advantageous for trials with <600 patients.

The study also examined various subsampling methods, such as sampling patients at random, sampling study centers at random, sampling patients at the largest centers, sampling the first patients enrolled, sampling the last patients enrolled, and sampling the first and last patients enrolled (the last patients enrolled and the first and last patients enrolled were analyzed only for comparative purposes, not as a practical methodology). Sampling by centers at random was determined to be the most logistically feasible and accurate methodology for subsampling. To ensure full representation, a stratified population of patients from small, medium, and large centers should be chosen.

A lack of consensus regarding data collection, specifically toxicity data, has led to frequent discordance between

practices in NCI cooperative groups and industry-sponsored clinical trials. The goal of this project is to recommend and justify sufficient data collection to generate safety data for drug labeling and clinical use and to reduce collection of unnecessary data elements in supplemental NDAs and BLAs. The effort and resources saved can be better channeled to focus on collecting more meaningful and accurate information that informs clinical and regulatory decisions and leads to greater participation in the clinical trial process.

FDA RESPONSE

In a guidance published in 2001 (Cancer Drug and Biologic Products—Clinical Data in Marketing Applications), the FDA provided recommendations for sponsors on data collection for cancer clinical trials submitted to the agency to

support marketing claims in NDAs, BLAs, and supplemental applications for new drug and biologic indications. The regulations (21 CFR 314.50) require that supporting data be submitted with study reports from well-controlled trials, but they do not describe the amount and type of data that should be collected.

Commercial sponsors may collect large amounts of information to ensure that they have all the data that regulatory agencies might request. Noncommercial organizations, for example, U.S. cooperative groups, frequently collect less information than commercial entities, although their trials may provide adequate data for important risk-benefit assessments supporting regulatory approvals. The FDA recognizes that extensive data collection can be expensive and time-consuming and that collection of unnecessary data is not an optimal use of clinical trial resources.

In the 2001 guidance, the FDA acknowledged that it is not possible to provide precise data collection requirements that could be applied to all trials because of the complexity and variability of clinical trial design. The FDA strongly encouraged sponsors to develop specific proposals for data collection and discuss their proposals with the agency prior to initiating clinical trials. The FDA maintained that agreement between the agency and the sponsor of the drug or biologic on prespecified data collection plans would “avoid the collection of unnecessary information, allowing resources to be directed toward studying important endpoints.”

As discussed in the 2001 guidance, the following factors should be considered when assessing what data elements are necessary to collect:

- The type of regulatory submission (e.g., new marketing application versus efficacy supplement).
- The similarity of the proposed new use of the drug to already approved uses of drugs.
- The population being studied (e.g., patients in the surgical adjuvant setting, patients receiving first-line treatments, or patients with refractory disease).

- The amount of available supplemental information from other sources on the safety of the drug, such as data from trials in a similar population.

The goal of the Data Submission Standards and Evidence Requirements project (Panel 1) was to identify the scope of data collection sufficient to generate safety data for drug labeling and clinical use and to reduce collection of unnecessary data elements in supplemental NDAs and BLAs. A study was conducted to determine whether important safety information would be omitted by collecting data on a subsample of patients enrolled in trials to support supplemental BLAs or NDAs for approved drugs with extensive safety information already available. Data sampling did not appear to omit safety information that would be needed for labeling or the benefit-risk evaluation.

Although this study focused on supplemental NDAs and BLAs for cancer drugs, the FDA believes that these findings could apply to safety data collection for supplemental NDAs and BLAs for all therapeutic drug classes. Safety data collection from all subjects would still be needed for initial marketing claims for NDAs and BLAs. However, based on the factors outlined in the guidance (i.e., type of submission, similarity of proposed use to approved use, population being studied, available additional information for other sources), it should be possible to more narrowly focus the scope of data collected without a detrimental impact on the regulatory evaluation of supplemental marketing applications of drugs or biologics.

The FDA is committed to developing a guidance applicable to all therapeutic classes. That guidance will further clarify and illustrate the principles outlined in the 2001 guidance for cancer drugs and biologics, as well as incorporate the findings from the Data Submission Standards and Evidence Requirements project.

ACKNOWLEDGMENTS

FDA Response provided by Robert Temple. Kelsey Mace and Cindi Stephens of ASCO and Jeff Allen and Rasika Kalamegham of Friends of Cancer Research assisted with drafting.

REFERENCES

- 1 U.S. Food and Drug Administration Office of Training and Communication, Division of Drug Information, HFD-240, Center for Drug Evaluation and Research. Conducting a Clinical Safety Review of a New Product Application and Preparing a Report on the Review, Food and Drug Administration. Available at <http://www.fda.gov/cder/guidance/index.htm>.
- 2 Schilsky R. Summary of data re-analysis. Presented at the American Society of Clinical Oncology Data Optimization Meeting, Alexandria, Virginia, July 30, 2009.
- 3 Schilsky R, Abrams J, Woodcock J et al. Conference on Clinical Cancer Research, Issue Brief, September 3–10, 2008.
- 4 A detailed manuscript for submission is currently in preparation under the leadership of the Working Group.