PhUSE 2009

Paper RG07

Data review: To see or not to see …

Estelle Tiemtore, Quintiles, Strasbourg, France
Damian Green, Quintiles, Strasbourg, France

ABSTRACT

Nowadays, there exist tools allowing easier and more targeted review of clinical trial data and which are adjustable to the database of each new or ongoing trial.

These tools allow data review in an interactive manner throughout the trial, focusing on the reviewers' points of interest.

To this aim, medical reviewers and biostatistics personnel (especially SAS® programmers), identify the needs of the reviewer, and develop the bridge between the CRF data and the review tool objects.
The following process allows SAS® programmers to deliver a large panel of review objects:

(1) Write specifications in order to clarify the reviewer's needs. Specifications identify how variables will be derived and displayed (by time point, by duration, by category, by study step or any other criteria)
(2) Derive data using the same derivations and flags as those for the statistical analysis and create summaries (an example using adverse events by SMQ groups will be presented).
(3) Setup study in the data review tool, import datasets and create several objects such as cross-tables, graphs, reports, listings and graphical patient profiles (a quick tour of common data review objects will be presented). To facilitate data review, some objects are cross-linked to each other and with source datasets accessible by simple point-and-click movements.
(4) Validate data review tool objects and dynamic links through the standards of the Biostatistics department QC processes.

To this end, SAS® programmers can then provide both summaries and by-patient data displays. This allows the medical reviewer to easily identify cases of interest through the summaries, visualize the individual patient's profile, and additionally drill-down into the CRF data within a controlled environment. The reviewer drives their review in a point-and-click manner in a top-down approach, based on the objects created.

Judgement, however, should be exercised as to who has access to the data. As pointed out in the FDA guidance on data monitoring committees, even aggregate data on safety and efficacy may be informative. To this end, control on who has access to the data, and furthermore which data are made visible, is important to consider from the start of the project. The data review tool is used to select data and create displays of only the data needed to be seen by the reviewer.

INTRODUCTION

Review of safety data during a clinical trial has become a common feature of many study protocols. The usual process involves the production of tables, listings and patient profiles (presenting either a selection or all of a subject's data) in text format files. Reviewers are then tasked with going through the outputs usually within a short timeframe.

Sometimes, spreadsheets are used rather than text files. This has the advantage of allowing the reviewer to sort or filter data, thereby aiding their search for important components of the data (outliers etc.). More advanced features of the spreadsheet can be used to summarize data and present it graphically.

Recently, computer-based data review tools have started to appear. These go a step further in that they dynamically link subjects' data between data panels, allowing the user to quickly browse through the data in a point-and-click environment. We will discuss a practical approach to a targeted safety review of clinical trial data, consider who should have access to the data, and furthermore, which data they should see or not see.

DYNAMIC DATA REVIEW

This section deals with interactivity in a clinical data review in order to perform a targeted data review. The example of an adverse events (AEs) review by Standardized MedDRA® Queries (SMQs) illustrates this section.

Let's define SMQs before continuing. The most common AEs coding system, MedDRA®, includes more than 18000 Preferred Terms (PTs) and it is difficult to group events that indicate the presence of a condition that typically presents itself through different PTs. SMQs are designed specifically for solving this problem [(a)].

The concept of SMQs was introduced by the Council for International Organizations of Medical Sciences' (CIOMS) working group on SMQs. The development of SMQs was intended to aid in identification and retrieval of potentially relevant reports/cases from a MedDRA® coded database. It is the result of a collaboration between industry, regulatory authorities and the MedDRA® maintenance organization (MSSO). Additionally, use of SMQs event groups is becoming more common for safety studies and safety summaries for licensing applications [(b).]

TO SEE OR NOT TO SEE: TARGETED DATA REVIEW

Safety data review during a clinical trial requires reviewers to be able to highlight the presence or lack of safety issues with respect to the treatment the subjects receive. Safety reviews should start early in the study in order that that sponsor or health authorities may continue or stop the study in view of risks to study subjects. The aim is to follow subjects as they progress through the study and build up a picture of the safety profile of the compound under study.

The Biostatistics department of Quintiles in Strasbourg has recently experienced some customers' requests illustrating a growing trend to perform this process using new data review tools. The team therefore proposed and set-up a dynamic data review to focus on specific data for internal safety review.

Everybody is aware that it is better to show less data but more specific/linked data (otherwise you can't see the wood for the trees). For example, usual tables and listings show events by SOCs and PTs. If events are shown by SMQs groups, AEs can be related to signs, symptoms, diagnoses, syndromes, physical findings, laboratory and other physiologic test data, etc., that are associated with the medical condition of interest. From a safety point of view, it is more targeted.

The first step required for a more targeted data review is to identify the exact needs of the reviewer. Depending on the goal of the review, the reviewer may have different types of needs. The programmers' role is to translate these needs into data displays.

Considering that subject-level review is time consuming, programmers explored the opposite approach, that is, start by an aggregated data review, with the following guiding points:

(1) The possibility of graphical displays to give a rapid overview of the data
(2) The capability to quickly drill down to sub-groups or subject-level data
(3) The ability to automatically link different datasets together
(4) The assurance of consistent output during the study using the same definitions as found in the statistical analysis plan.

Dynamic data review tools provide answers to these requirements by offering the possibility of detailed and summarized outputs and allowing interactivity during review.

A BRIDGE BETWEEN 2 HORIZONS: FROM DATABASE RAW DATA TO DYNAMIC OBJECTS

The first step of a dynamic data review implementation, from the programmer's point of view is to clarify the reviewers' goals and write specifications to ensure the reviewer's expectations are met. Let's consider the case where the reviewer's goal is to review AEs by SMQs groups as part of a database safety review.

The second step consists of creating specific derived datasets to be uploaded into the dynamic review tool. The programmer's expertise is very important in creating the relevant flags and the derivation of new variables. In the

present example this includes flagging treatment emergent AEs, serious AEs, AE NCI-CTC(c) grades etc. in order to produce frequency summaries.

The last step entails the creation of the data displays, once the study environment has been set-up in the dynamic data review tool. Each data display is identified by a specific icon associated with the type of display (table, listing, graph, profile). For the purpose of our example, different icons will be created for PTs listings, PTs frequencies tables and PTs frequency bar chart by SMQs group.

The reviewer can open each display with a simple click on the icon. In the present case, this will start with the PTs frequency bar chart by SMQs group. Each SMQs group is displayed on a separate page, but several pages can be opened on the same screen. If a click is made on a bar, the subjects concerned are selected and their identifiers appear automatically. The dynamic tool automatically links data across datasets at the subject level. Using the subject identifier, the tool filters and displays only the data for the selected subject in all the datasets and subject related outputs such as listings and profiles.

Other types of data (laboratory data, concomitant medication data, vital signs and previous conditions) describing the subjects medical condition are available by a few simple clicks. The reviewer has all these data in the same screenshot (see figure 1) and can easily drill down into them and make associations.

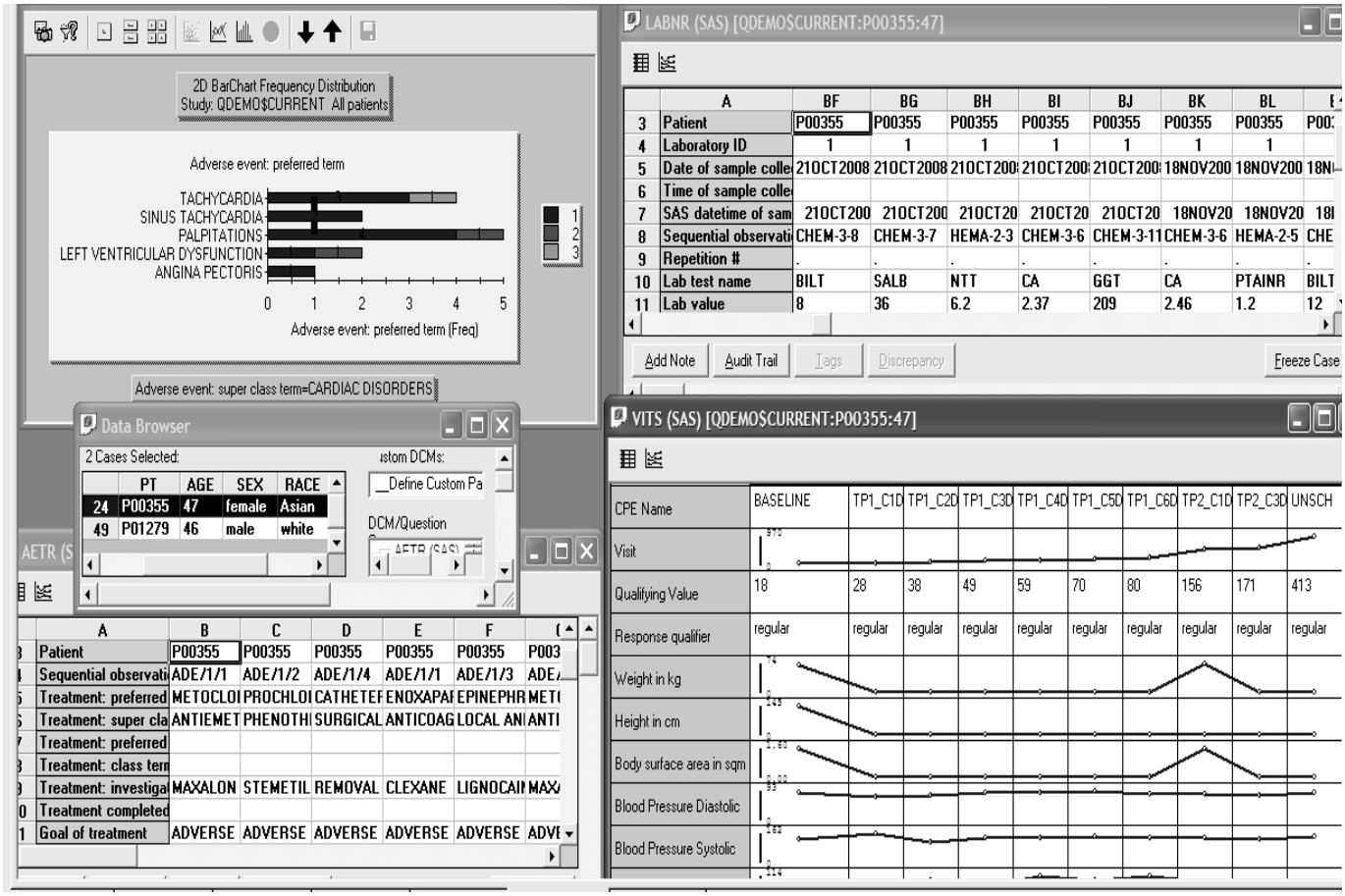

Figure 1: Drill-down and links into different data sources in the same screenshot

The data displays once implemented by the programmer are available during the life of the study. The reviewer can target a subject group by reviewing SMQs frequencies tables at any time.

For example, it's as straightforward to know how many subjects have treatment emergent kidney problems as it is to compare incidence of cardiovascular illnesses against leucopenia or hepatic failure, or if skin reactions tend to be severe or not. Subjects of interest can then be identified by a simple click.

In some cases, it is interesting to perform a direct subject-level safety review, for example, for subjects with special circumstances (discontinued due to AEs, experienced serious AEs). These criteria can represent subgroups that can

be identified using an icon created with the tool based on programmed flags. Furthermore, the reviewer can access a graphical patient profile in order to reduce line-by-line listings review. Graphical profiles present a subject's data taken from multiple datasets and plotted against a common time axis (days, weeks, months, years …). Figure 2 shows an example of a graphical patient profile.

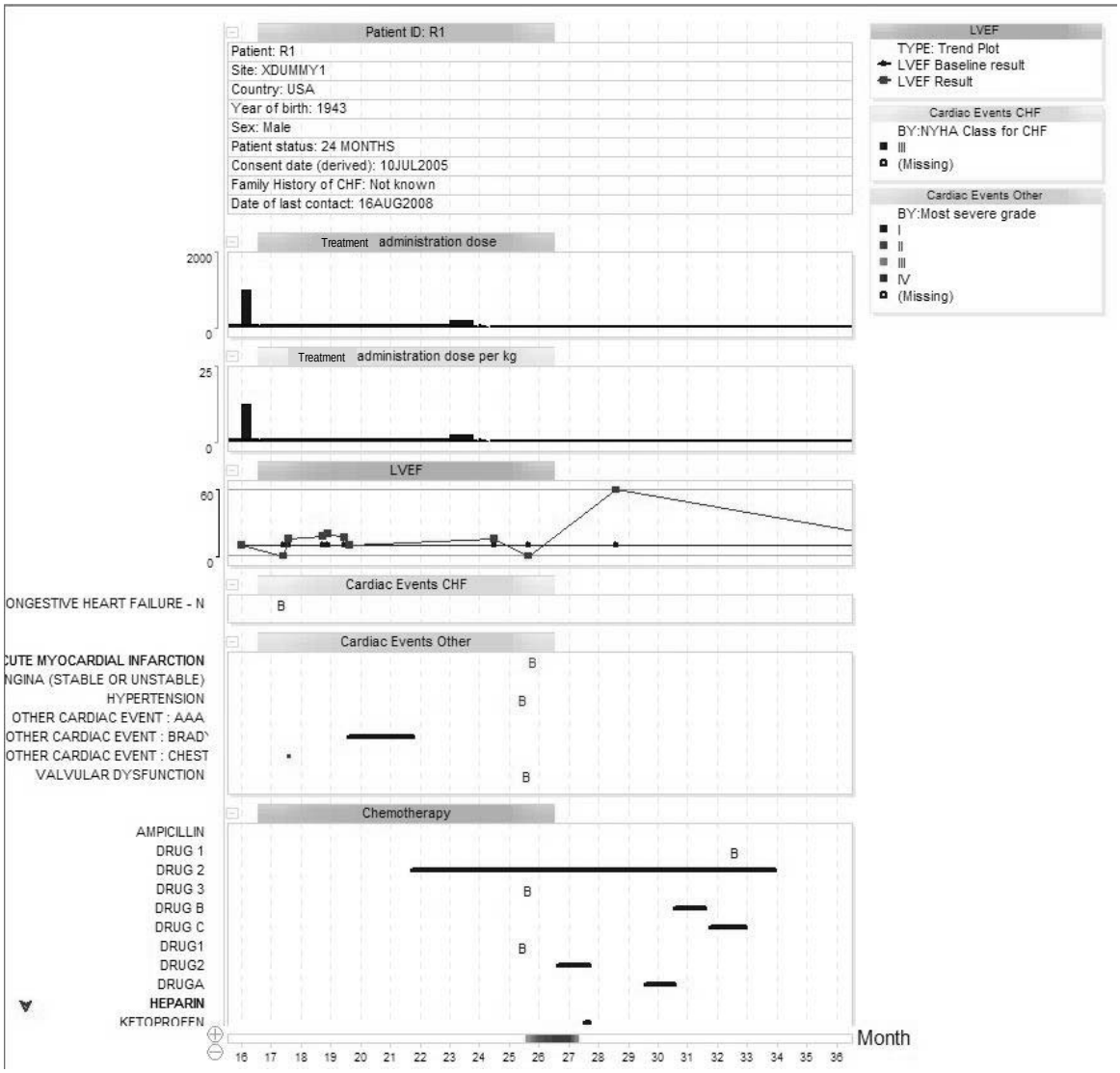

Figure 2: Graphical patient profile

In general, the reviewer has a large choice of data displays in order to conduct the review as he wants. A targeted process may start with the summarized displays (frequencies tables, frequencies plots) followed by a drill down to the detailed displays (patient profiles and listings).

The Biostatistics' department quality control process ensures the quality of data displays. This includes verification against the source data, and also cross-checking several types of outputs. The team also provides training for reviewers and technical support for optimum navigation through the data in the dynamic review tool. Reviewers are provided with a user ID and initial password, after which they have anytime access to the data displays.

A BREATH OF CHANGE

The present economical climate promotes change in the biopharmaceutical clinical research process. A sponsor may prefer to have both detailed outputs and summarized ones together within the same interface, accessible at anytime from anywhere, and available throughout the life of the study. This approach allows for an unlimited number of reviews without the consequences of large production efforts.

This is a real break with what until recently was the usual process (in our experience); providing the full CRF data dumped into a patient profile across the board for all the subjects, 'served-up' with some simple tables and listings during the study for data reviews, and then summarized outputs (tables) together with listings and graphs for the final analysis after the database lock.

Dynamic data review tools are a real contribution to changes observed in the clinical development process. They provide many concepts of interest such as:

(1) Setting up the same review approach across a set of studies (multiple protocols on same compound). The set up is done once for the set of studies and provides consistency. This approach also allows a pooled review across the whole data of the set of studies.
(2) Offering a variety of outputs (detailed or summarized) in the same tool, in the same interface.
(3) Dynamic linking between data modules and instant access to new data when connected directly to the clinical data base in a point and click environment (users do not need programming knowledge).

CONTROLLING DATA ACCESS

TO SEE OR NOT TO SEE: REGULATORY GUIDANCE

As mentioned above, the advent of data review tools opens up a wide range of possibilities for the user to review the data of an ongoing clinical trial. These tools allow users to be attributed to specific user groups which limit the options available to the user, e.g. read-access only, where a user may only view the pre-prepared data displays, or author-access, where a user may create custom tables and listings. Generally, only the programmer setting up the data displays should have author rights. This is in keeping with current practice using hard copies or electronic files with tables and listings – the reviewers see only what has been previously agreed and documented.

Furthermore, customized user groups can be created in order to further specify who sees what. For example, the needs of medical reviewers, data managers, data monitoring committees etc. are different, and hence should be taken into account. Care should be taken when granting users access to the data review tool. The FDA guidance for the Establishment and Operation of Clinical Trial Data Monitoring Committees (section 4.2.2) states "Even aggregate data on safety and efficacy may be informative; these data ...are best limited to those who cannot otherwise carry out their trial management responsibilities." A similar point is raised by Ellenberg *et al*[(d)] who write that "such data could be provided to selected individuals who 'need to know' such information to carry out their ethical or scientific responsibilities in the conduct of the trial". They give a couple of examples were aggregate data may provide 'suggestive information' as to whether the test treatment is effective or not.

Other instances where data may be informative are certain parameters e.g. a standard laboratory test that is influenced by the metabolism of the test drug, and which at the individual patient level, may potentially provide unblinding information with regards to the test treatment. Careful consideration is needed therefore not only in deciding who has access to the data, but also what data they need to see. The programmer would therefore need to ensure that the specific parameter of concern is not included in the data to be reviewed.

Hence when planning the set-up of the data review tool, the programmer has several methods at his disposal to minimize the possibility of unblinding the study team, and thus avoids introducing bias.

In the case of double-blind, randomized clinical trials, the test and reference treatments are masked, and are not known to the investigators, patients and study team. Knowledge of unblinded interim data during the study by any of these groups may introduce bias and put the integrity of the trial at risk. If the data review tool is also to be used for the purposes of providing a data monitoring committee (DMC) with unblinded data reports, then special attention should be taken to ensure that access to those data are limited to the DMC members and the unblinded biostatistics team that produced them. In our experience, this has been managed by creating a study area separate from that accessed by the study team, where the unblinded data are stored. Although the initial set-up of the study environment in the review tool may be performed by the study programmer, access is then revoked. The data display icons in the review tool are independent from the data source (blind or unblind), since they are only an interface to the data, and hence the programmer can develop or modify these without seeing the unblinded data. The unblinded study team populates their study environment with unblinded data when needed.

CONCLUSION

New tools for interactive and dynamic data review are becoming more commonplace, helping to reduce the need for repeated production runs of tables and listings which is costly and time-consuming. However, similar to the advent of the electronic case report form, the set-up phase requires planning and preparation up-front, but the any-time access for reviews is a blessing for efficiencies later.

Regardless of all the positive points noted in this paper, dynamic data review tools are not well established for independent safety review because of the traceability of each review. Data monitoring committees need printed documents (pdf) as archives. Dynamic data review tools can produce some printed outputs, but as this is not their main aim, the printing process may be cumbersome and it may be difficult to show details (such as printing graphical patient profiles).

We are entering a new age for approaches to data review, and going forward, we hope that all parties will work together to increase and improve the use of dynamic data review tools. There is no doubt that reviewing clinical data dynamically for safety issues is easier, quicker and less expensive than conventional safety review on paper.

REFERENCES

(a) Applying SMQs to Adverse Event Data, PhUSE 2008 - Paper TU05, John van Bemmelen, Schering-Plough, Oss, The Netherlands

(b) MSSO EU user group 10Mar05_SMQs.ppt on internet. Presentation of CIOMS Working Group on Standardized MedDRA Queries (SMQs) available on http://www.meddramsso.com/MSSOweb/docs/MSSO%20EU%20user%20group%2010Mar05_SMQs.ppt.

(c) NCI-CTC : National Cancer Institute – Common Toxicity Criteria (ref : http://www.cancer.gov)

(d) Data Monitoring Committees in Clinical Trials: A Practical Perspective, Susan S Ellenberg, Thomas R Fleming, David L DeMets, 2002 Wiley.

ACKNOWLEDGMENTS

I would like to thank Quintiles Strasbourg IT Helpdesk and Data Management departments for their involvement the development of dynamic data review expertise.
I would like to thank also Dr. Jorgen Seldrup for his comments that helped improve this paper.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Author Name: Estelle TIEMTORE
Company: Quintiles Benefit France
Address: Rue Jean Dominique Cassini
City / Postcode: Parc d'Innovation d'Illkirch / 67404 Illkirch
Work Phone: +33 (0)3 88 26 60 25
Email: Estelle.tiemtore@quintiles.com

Author Name: Damian GREEN
Company: Quintiles Benefit France
Address: Rue Jean Dominique Cassini
City / Postcode: Parc d'Innovation d'Illkirch / 67404 Illkirch
Work Phone: + 33 (0)3 88 77 45 09
Email: damian.green@quintiles.com