Data Transparency and Sharing: Research Benefits, Risks and the Future

Janet Stuelpner, SAS Institute
Matt Becker, SAS Institute

ABSTRACT

Whether called data transparency or data sharing, there's a movement to give more researchers greater access to patient-level clinical trial data. The goal is to create an environment for innovation in clinical research. Join this presentation to discuss what is being done, including exploring the value to the overall health care system of creating a multi-sponsor environment that gives researchers access to larger pools of data.

INTRODUCTION

In 2012, the idea of pharmaceutical companies providing their proprietary clinical trials data as a public resource would have been met with skepticism and disbelief. Fast-forward to 2013 and 2014, and pharmaceutical companies are racing to market with implemented solutions that not only enable them to share their own company trial data, but enables that data to be combined with data from other pharmaceutical companies. Whether it's to stay ahead of emerging guidance from the European Medicines Agency, or to simply document the integrity of their research programs, one thing is clear - clinical trial data transparency is a critical topic to understand in 2014.

Clinical Trial Data Transparency (CTDT) became a hot topic for the Life Sciences industry in the last two years. Companies are now positioning themselves to have something in place before the regulatory authorities' requirements are published. Many companies have jumped at the chance to be ahead of the curve when it comes to how the industry should be making their data available to researchers before they are required to do so. This paper will show the key data transparency developments that have happened within the last year. We will also talk about the timeline for getting something in place and what needs to be included or excluded. As we move forward in time, the principals of the regulatory agencies are changing and being refined. In conclusion, the future development and direction data transparency will be explained.

A LITTLE BIT OF HISTORY

The European Medicines Agency (EMA) has been releasing clinical trial reports, on request, as part of its access-to-documents policy since late 2010 and is now working toward its goal of publishing clinical trial data pro-actively for the medicines that it has assessed. The EMA policy is about access to documents related to medicinal products for human and veterinary use. According to this policy, the EMA releases documents, including clinical trial reports, once the decision-making process for the medicine in question has been completed. Between November, 2010 and April, 2013, the EMA released over 1.9 million pages of clinical trial data in response to safety-related requests. The EMA is committed to continuously extending its approach to transparency. A key goal in this process is the pro-active publication of clinical trial data for medicines once the decision making process on an application for a European Union (EU) wide marketing authorization is complete. The EMA has embarked on this process because it believes that the release of data is about establishing trust and confidence in the system. This will enable the independent re-analysis of the evidence used by the EMA's committees to determine their benefits and risks and is expected to lead to public health benefits. The EMA believes that the sharing of complex data can open up new horizons.

More than 200 people from all stakeholder groups applied to participate in one or more of the five advisory groups. The groups met between January and April, 2013, with meetings taking place via teleconference. The topics that were covered by the advisory groups are:

1. Protecting patient confidentiality: how can the EMA protect the identity of the subjects that are participating in clinical trials?

2. Clinical trial data formats: how can the EMA ensure through its policy that clinical trials data can be shared in the interests of public health in a clear and understandable format?

3. Rules of engagement: are there rules that need to be in place before any external stakeholders can download information?

4. Good analysis practice: are there good guidelines that should be used as hypotheses are being created and requests for information are made?

5. Legal aspects: are there any other legal aspects other than data protection that need to be addressed when drafting the policies?

At the end of the fourth quarter of 2012, many of the top pharmaceutical companies jumped out front to provide their data for several studies to outside investigators of the research community. GlaxoSmithKline was the first company to promise to make detailed subject level data available to independent researchers so that scientists can draw their own conclusions about the safety and effectiveness of their new compounds.

In the summer of 2013, a draft guidance was released from the EMA. That guidance stated that starting in January of 2014, all future submission would be of de-identified data and would be made available. Many companies felt that they would support it, but they were not happy with the draft guidance. The guidance was opened for review and resulted in thousands of comments. Thus, the EMA delayed the release of the official guidance with the list of regulations.

Biopharmaceutical companies are committed to enhancing public health through responsible sharing of clinical trial data in a manner that is consistent with the principles of responsible sharing of clinical trials data sharing. There are 5 guiding principles for data transparency.

1. Enhancing data sharing with researchers

2. Enhancing public access to clinical study information

3. Sharing results with patients who participate in clinical trials

4. Certifying procedures for sharing clinical trial information

5. Reaffirming commitments to publish clinical trial results

There has been a great deal of hype and buzz about the EMA guidelines. The Life Science companies completely supported the five principles, but they needed to have the sharing of information done in a way to honor patient privacy and the intellectual property that is created for a compound.

DIFFERENCES BETWEEN SHARING, TRANSPARENCY AND OTHERS

What if we could share, integrate, and analyze our collective historical research data in a single location? This is the definition of data sharing. The ability for an independent group to pull data together for use by independent researchers to gain insights into a massive amount of data. There is an increasing volume of data and the continuing engagement of a diverse global community focused on finding solutions. This is the Project Data Sphere initiative that provides easy access to Oncology study data that is shared, can be integrated, and can analyze comparator arms of historical trial data sets so we can learn from them and accelerate research. The study data for Project Data Sphere Oncology projects is **patient level data** and requires agreement between all contributors.

The concept of data transparency was originally driven from the threat of regulation by the EMA. Each life science company would create their own de-identified data to be made available to independent researchers. The need is to make the information accessible for further medical research.   Again, this is patient level data. It is data that is only available to researchers with a valid research project and is not available to the public. Companies can create an environment that is only their own data or they can load their data into a multi-sponsor environment. In the multi-sponsor environment, each company can only view their own data. Even though it is a shared environment, they cannot see the data that has been loaded from another company.

ClinicalTrials.gov is a web site that constitutes a registry for all clinical trials. It is the largest clinical trials database holding regulations from a multitude of trials. This effort served as an example of what might be done to improve public access to clinical trials, and motivated specific disease-related interest groups to push for something for all diseases. The Food and Drug Administration Modernization Act of 1997 required the NIH create and operate a public information resource, which came to be called ClinicalTrials.gov, tracking drug efficacy studies resulting from approved Investigational New Drug (IND) applications. This is **summary level data**. There is no patient level data available in this database. The primary purpose of ClinicalTrials.gov was to improve access of the public to clinical trials where individuals with serious diseases and conditions might find experimental treatments. This law required information about:

1. Federally and privately funded clinical trials
2. The purpose of each experimental drug
3. Subject eligibility criteria to participate in the clinical trial
4. The location of clinical trial sites being used for a study
5. A point of contact for patients interested in enrolling in the trial.

There are three data sharing models in the life science industry:

Black Box/Database This is where an analysis is requested and can be done by someone else. You get back only the results, no data. You can't see the data. Someone else does the analysis.

Open Access This is the first data sharing model that is being proposed by the EMA.

Controlled Access There is a control arm or an independent review entity. Anyone can access the data, but they need to go through a gatekeeper. Also, they can only access the data in a specific environment and cannot extract the data. There are two way that this is being implemented. First is a single sponsor environment where a company will keep their own data separate from everyone else. Second is the multi-sponsor environment where data is shared in one environment. Each project is completely isolated and you can only look at the data for a specific research initiative. The multi-sponsor environment is the direction that the industry is going.

SUMMARY

- What is it
    - Life Sciences companies are providing access for valid research purposes to patient-level de-identified clinical trial data
    - Single sponsor or multi-sponsor environment
- Who is providing it
    - Life Sciences companies
    - Potential for other organizations including non-profits
- Who will use it
    - Researchers from the industry and academics (NOT the sponsors)
    - Potential health care users, CROs, etc.in the future
- How will they use it
    - "Valid" research purposes as defined by an independent panel
    - Not for general exploration – must submit research proposal
    - Life Science companies will have access to results depending on the type of request (single vs multi-sponsor)
- Who is in these discussions
    - Life Sciences companies who are currently evaluating a solution or may in the future
    - Companies interested in addressing impending regulations
    - Third party companies that may add services
- What information is included
    - Patient-level clinical trial data
    - Any other supporting documentation (CRFs, SAP, etc)
- Why are they doing this
    - Enables the replication and verification of results
    - Valuable resource for the generation of new findings
    - Obligation to participants and reuse of their data
    - Improve public health, enhance patient safety, and spur drug development
    - Increase public trust
    - Proactively addressing potential regulation