CHILTERN

Designed Around You®

Applying SDTM and ADaM to the Construction of Datamarts in Support of Cross-indication Regulatory Requests

Data-what?

- This presentation will discuss some of the complexities involved when integrating study data that spans long periods of time as well as multiple sponsors and indications.
- The SDTM and ADaM models can be useful when integrating study data from multiple phase III and IV clinical trials into a single data warehouse, commonly referred to as a datamart.
- Datamarts provide relatively homogeneous and readily available data for responding to regulatory requests.
- Is it more efficient for you to build and maintain a central datamart than to create an ad hoc datamart on the fly each time a request comes in?
- External requests, either directly from a regulatory agency or via a partner, usually feature a tight timeline.
- Even if you have a datamart. There will still be some request-specific data set programming to do.

Why Should Your Datamart Follow CDISC (more or less)?

- The deliverables for a regulatory request are usually tables, figure, and listings. If the datasets are not to be submitted, then why bother with creating the datamart using a CDISC-ish approach?
- Ease of maintenance: Whoever works on the datamart adding new studies, additional derived variables, tweaking, and debugging will find a CDISC-friendly structure familiar.
- Ease of sharing: Sharing a common data standard makes it easier to share data with co-development/co-marketing partners.
- How compliant with CDISC should your datamart be? Keep diminishing returns in mind.
- Future studies will probably be compliant with some version of CDISC.

Legacy Studies

- One of the challenges in creating a CDISC-ish datamart is the integration of legacy studies. The word legacy can take several meanings. Sometimes the word “legacy” is used to describe a study that began prior to the adoption of CDISC guidelines, but will be submitted in a CDISC-compliant fashion.

- However, here legacy means an old study that started, completed, and finished long before the adoption of CDISC standards.

- Subjects may have participated in several studies over a period years. Commonly referred to as rollover subjects. Early study data may have been acquired through an IP purchase, merger, acquisition, etc. So, the data can span multiple decades, companies, SOPs, etc.

- SDTM: You have to decide whether it is more important to 1) keep datamart SDTM data sets as true as possible to the study-level data or 2) comply with the SDTMIG. SUPPQUAL comes in handy for things you’ll need in ADaM that don’t quite fit into SDTM.

- ADaM: Will you let datamart ADaM reach back beyond datamart SDTM to the study-level analysis data sets and/or study-level raw data to derive datamart variables that were not used in the CSR, but are needed for regulatory requests?

Data from Old Studies Closed Long Ago

- Do you make “corrections” to study data in old, closed studies?
- If you do, is it considered hardcoding (bad) or derivation (not bad)?
- What do you do with an AE having start date of **31**JUN19XX? Your decision could affect whether the event is included at all. And if it is, whether it is attributed to study drug, placebo, or active comparator.
- Will want to maintain a CDISC-style spec to document derivations, updates, provide insight into your thinking at the time, etc.

Closed versus Ongoing Studies

- *Closed* studies are absolutely done (pretty much), the database is locked forever (probably).
- Even closed studies will at least need coding dictionary updates to be applied at either the datamart or request level.
- But, when and how do you add study data to the datamart for *ongoing* studies?
- What data snapshots/cutoff points/milestones are you going to use for ongoing data for a given regulatory request?
- Do you want to have a single datamart that includes both closed and ongoing studies and then refresh the whole datamart whenever you get a request?
- Or do you want to leave data for ongoing studies out of the datamart until the study database locks and just add that data on-the-fly within request-level analysis data set programs whenever you get a request?
- Or do you want to create two separate datamarts, one for closed and one for ongoing studies?

Multiple Indications – another layer of complexity

- Study drugs are sometimes approved for multiple indications.
- The investigation of the different indications is not necessarily concurrent or co-located.
- You could end up with two or more series of clinical trials that have little in common other than the study drug being investigated.

Coding Dictionary Issues

- The MedDRA dictionary helps us use the same name when referring to an adverse event. The MedDRA dictionary is updated from time-to-time.

- A request from a regulatory agency may reference specific events-of-interest expressed as a list of Adverse Event Preferred Terms (AEPT), Adverse Event System Organ Classes (AESOC), etc.

- The regulatory agency uses a specific version of the MedDRA coding dictionary to generate these lists of events-of-interest. In order to accurately respond to the request, the event data in the datamart must be using the same MedDRA version as the request.

- Applying updated MedDRA dictionary terms to a data set of adverse events at either the datamart or request evel is fairly straightforward.

- Do you have WHO Drug coding issues for conmeds that come up frequently enough to be handled at the datamart level? If such requests are rare and the list of medications-of-interest is small, then maybe it is more efficient to handle WHO Drug coding updates at the request-level.

Event Mapping

- For the clinical study report, some sort of logic was used to map adverse events to the correct treatment period.
- What about rollover subjects, subjects who participated in multiple studies?
- Mapping adverse events to the correct treatment period for subjects who participated in multiple phases (for example, double-blind and open-label) and/or treatments (for example, study drug, active comparator, and placebo) is one of the more complex issues to address when creating a CDISC-friendly datamart of study data that spans decades, studies, sponsors, and indications.

Event Mapping Cont.: Rollovers, Partial Dates, Gaps

- Say a subject participated in three studies: the first and third subject took study drug, the second study the subject took placebo/active comparator.
- Do you map events using only dates of treatments and event, or do you take into account the study in which the event was reported?
- What about an event reported in the first study that falls in the treatment window of the second study?
- What about events that fall into the gaps in-between studies? What if the gap is small? What if the gap is big? What is “small”? What is “big”?
- What do you do with a partial AE start date of “2004”? Impute 01JAN2004? Impute 31DEC2004? Both? Neither? Your decision could affect whether the event is included at all. And if it is, whether it is attributed to study drug, placebo, or active comparator.