JOHNS HOPKINS
U N I V E R S I T Y

Center for Clinical Trials

Department of Biostatistics
Department of Epidemiology
Department of International Health

Department of Medicine
Department of Ophthalmology
Oncology Center

(Sunday) 14 January 2001

Memorandum

To: Center faculty, staff, and friends

Fr: Curt Meinert

Re: Public access to finished datasets of trials

Definitions

close of trial *n* - 1. The **point** at which **treatment** (as dictated under the **treatment protocol**), **scheduled followup**, and **data collection** end — typically marked by completion of the **close-out stage** of the **trial**. 2. The point at which treatment is stopped or suspended in a trial (not recommended usage); **treatment cessation** (defn 2); **treatment termination** (defn 2). 3. The point at which all activities related to the trial, including **data analysis**, end — typically marked by completion of the **termination stage** of the trial. 4. Termination of the **enrollment phase**. 5. Termination of funding for the trial. 6. **end of trial** ***Usage note***: Phrase subject to ambiguities, especially in settings in which **study treatments** are administered only once on **enrollment** or shortly thereafter and in which those so enrolled and treated are then simply followed for **outcomes** of interest. In one sense, in this setting, the trial is ended as soon as the last person enrolled has been treated, even though **followup** may continue for years thereafter. Usage should be restricted to settings in which all patient-related activities are terminated, including **scheduled followup visits** and related data collection. Avoid use in settings in which only specific activities or functions are terminated, such as enrollment but not followup; be specific in such settings about what has ended and what continues.

end of trial *n* - 1. The **point** at which **treatment** is stopped or changed for a person. 2. **close of trial**

public use dataset *n* - 1. A **dataset** residing in a **public repository**. 2. A dataset residing at a Federal, State, County, or City agency that is available for use by the public.

primacy, right of *n* - The **right** of being first. In regard to **trials**, the right of those who collect the **data** and carry out the trial to be first to **present** or **publish** (prior to making or being required to make data available to others for interpretation or **analysis**).

Room 5010 • 615 North Wolfe Street • Baltimore, Maryland 21205 • 410 955-8175 • Fax: 410 955-0932
E-mail: cctmail@jhsph.edu • Web site: http://www.jhsph.edu/Research/Centers/CCT

SOCA data access policy: Example

Access limited to study investigators during conduct of trial; use for presentation at public meetings proscribed except as approved by Study Officers; no public presentation of results by treatment group until trial is completed or stopped

Performance data: Distributed to study investigators on regular basis; presented at meetings of the research group

Baseline data: Distributed and presented at meetings of the research group

Treatment effects data: Access limited to members of the treatment effects monitoring committee; blackout for study investigators (except for study officers serving as members of the treatment effects monitoring committee)

Datasets used in producing publications containing results placed in public repository on publication (National Technical Information Service, Arlington, Virginia)

Issues in drafting policy on deposit

- Operational definitions of "final dataset" and "end of trial"
- Rights of investigators vs rights of others
- Reasonable limits to investigator right of primacy
- Limits on access to interim treatment results
- When and where to make data deposits
- Method of reviewing and servicing requests for access to study data

Recommended policy and practice

- Preserve patient confidentiality; do not release or deposit listings or datasets where patients can be identified or identified on a probabilistic basis
- Limit access to interim treatment results to persons or group responsible for treatment effects monitoring
- Provide unfettered access to supplementary tables on publication of manuscripts
- Deposit dataset supporting a manuscript on publication; deposit for unfettered use, preferably in a public archive
- Deposit a composite, finished, dataset just prior to cessation of all activities in the trial; deposit for unfettered use, preferably in a public archive
- Impose "statue of limitation" on investigator right of primacy (see previous memo on limits to investigator right of primacy)

This is the fifth in a series of memos concerning issues in the presentation and publication of results from trials. Previous memos have dealt with the obligation to publish, investigator right of primacy, limits to that right, and type and place of publication.

This one deals with access to datasets of trials. It is written from the position that finished datasets of trials should repose in public archives. The position derives from the fact that trials

© 2001 Curtis L Meinert

represent a form of public trust, that trialists have an obligation to maintain that trust, and that one of the ways to ensure continuation of that trust is by repose of finished datasets in public archives. Indeed, if trials are undertaken to yield *fruitful results for the good of society* (2nd item in the Nüremberg Code), is not deposit a logical expression of that requirement?

This said, it is also true that investigators are instinctively wary of public access to their finished data. The wariness comes from concern that such access will compromise their right of primacy, that it will open them to energy consuming challenges and criticisms, and that the costs and efforts spend on preparation of data files for deposit could be more productively spent on other activities of more direct and immediate relevance to the trial.

To be sure, one is not helped, when arguing for policies of deposit and open access, by the absence of norms and standards favoring deposit and open access. The reality is that such norms or standards, if they ever come about, lie somewhere in the future.

Sponsors of trials have no obligation to make data available to the public nor is the Federal Government obliged to do so in regard to NIH-funded trials (through, in some NIH contract-funded trials, the work scope includes provisions for deposit of a finished dataset with the sponsor). In regard to the FDA, datasets underlying NDA approvals are regarded as proprietary. They are not available to the public.

There is evidence of an emerging Federal standard for deposit and public access. A provision included in the Omnibus 1999 Appropriation bill (Public Law 105-277; 21 Oct 1998) "*directs the Office of Management and Budget to amend circular A-110 (Uniform Administrative Requirements for Grants and Agreements with Institutions of Higher Education, Hospitals, and Other Non-Profit Organizations) to require Federal awarding agencies to ensure that all data produced under an award will be made available to the public through the procedures established under the Freedom of Information Act (FOIA)*" (Epidemiology Monitor; Jan 1999; vol 20; No 1; see also Science vol 282; 6 Nov 1998). The impetus for the action arose from frustration in Congress with the EPA and its refusal to make data used in setting limits on air pollution available to the public under the FOIA. Technically, the provisions of the Act, at present, relate to access to data used by the Federal government in establishing regulations. But, that being the case, it is not difficult to envision provisions of the Act extending, ultimately, to the FDA in regard to making data available supporting NDAs.

© 2001 Curtis L Meinert

Policy on access varies depending on whether the request for access is internal or external to the trial and whether the trial is ongoing or completed.[1]

Access can be passive or active. *Passive access* is that made possible by providing a person or group with summary data contained in a report or document prepared by the coordinating center or some other body in the study. *Active access* is that made possible by providing a person or group with actual study data.

Generally, only passive access is allowed during a trial and that is limited to members of the investigator group. As a rule, there is no access to interim treatment results, except by those responsible for performing interim analyses or reviewing such results (eg, as represented by persons in the coordinating center, study officers, and members of the treatment effects monitoring committee). Usually, personnel responsible for enrolling, treating, and following patients do not have access to interim results. (They have the means, in the case of unmasked trials, to summarize their own results by treatment assignment, but are expected to refrain from doing so.)

The preferred approach is to deposit in relation to publications, especially in long-term trials likely to yield a series of publications. Under this approach, the dataset supporting a publication is deposited when the publication appears. The "deposit as you go" approach ensures timely deposits and avoids the problems and difficulties likely to be encountered if deposit is put off to the "end". (Personnel, when a trial is winding down, are more interested in finding new employment than in preparing data for deposit.)

The "deposit as you go" approach does not, however, ensure deposit of a compete composite dataset. Variables not represented in any of the datasets will not be covered in the "deposit as you" system of deposit. And, even if all variables are represented, the cutoff dates for the datasets will differ, therefore, piecing datasets together will not produce a complete composite finished dataset. If a composite finished dataset is desired, it must be produced after all data have been collected, entered, edited, and checked. Policies and procedures calling for deposit of a "final" finished dataset at the "end" of the trial should include operational definitions of

[1] For purposes here, an ongoing trial is one in which investigators are still enrolling, treating, or following patients and, therefore, have yet to publish treatment results. A completed trial in one where all treatment and followup has ceased, patients have been separated from the trial, and study investigators have published results, or, in the absence of publication, where their right of primacy has expired. A partially completed trial is one involving a complete or partial stop of treatment in the trial and where investigators have published treatment results for the component of the trail stopped or where their right of primacy has expired.

© 2001 Curtis L Meinert

"final" and "end" and should include funding for preparation of the finished dataset and the related documentation.

© 2001 Curtis L Meinert