(Wednesday) 15 December 2010

Memorandum

To: CLM

Fr: Curt Meinert

Re: Sanity and data sharing

As to your question: I hope I have not lost my mind, but who knows? Nobody would tell me if I have.

Maybe I need to see one of the shrinks I work with. I suspect they would come up with a diagnosis that includes "syndrome" in the name and some drug of unproven value that I would have to take the rest of my life for the condition.

I think I will skip the shrink.

I could ask you the same question. If the only uses of data in a study would be those specified in consents, we would not be doing many analyses or much paper writing, or, alternatively, consents would be as long and dense as homeowner's insurance policies.

The issue in data sharing is whether the sharing is within the limits of consents. Do the uses meet the "reasonable person" test? If such a person, on learning of sharing outside the study, would say:

I'll be damned. Why didn't they tell me that before I enrolled. I wouldn't have enrolled if I had known.

If that is the sentiment, then you have exceeded the limits of consent.

Most people enrolled in studies expect data to be shared with fellow investigators and, perhaps, even beyond study investigators, if the sharing advances the purpose of the research. Things go south when the sharing is without regard to purpose and when people given data are not constrained by consents.

By your dictum, there would not be any sharing, even that relevant to the aims of the research. For example, there would not be any sharing of SOCA data with NA-ACORD (North American AIDS Cohort Collaboration on Research and Design). NA-ACORD did not exist when most people were enrolled in SOCA and re-consenting for the sharing is logistically impossible.

The norm within a study is to share data with others outside the study so long as the sharing is relevant to the aims of the study. What is new is mandated data sharing where investigators are removed from any role in how data are shared or used.

The problems with the NIH policy are that:

(1) Sharing is a requirement absent convincingly arguments to the contrary,
(2) Uses of shared data are beyond the control of those who collected them and, hence, may be beyond uses indicated in consents,

and

(3) Data are provided free of charge.

Study investigators are accountable to IRBs. That connection is severed with mandated data sharing where uses are not subject to investigator review or control.

The most serious problem with the NIH mandated data sharing policy is that it is unfunded. The study, usually the coordinating center in multicenter trials, is expected to bear the costs of mandated data sharing. The costs of preparing de-identified datasets, data dictionaries, and for quality control and editing to ensure the de-identification is adequate and that the data are correctly mapped and transcribed are significant.

The costs should be covered by the NIH sponsor demanding the sharing. If budgets do not include funds for data sharing, the study budget should be supplemented by the NIH to cover the costs. The amount will vary depending on the size and complexity of the dataset but, conservatively, can be expected to be $50,000 or more.

In addition, persons requesting data should be charged processing fees for data requests. The fees should be large enough to filter out the faint of heart and frivolous requests. Likewise, once data are supplied, the user should be required to pay the data center a flat dollar sum (eg, $250) per query regarding use of the data supplied.