

**IECRN National Leadership Forum**  
**June 1, 2006**  
**Report of Breakout Sessions: Data Management**

**Presenters:**

**Meredith Nahm, MA, CQA, CCDM**

Director of Clinical Data Integration, Duke Clinical Research Institute

**Elsa Villarino, MD, MPH**

Chief, TB Trials Consortium Team, Clinical and Health Systems Research Branch, CDC

DR. SACHS: Thank you. I'm going to introduce the next session, which is data management and I'd like the facilitators, the moderator and the two content experts to come on up.

Elsa Villarino, Meredith Nahm, and Therese Brown Gibson. Where are they? They're coming. Great. Thank you.

MS. NAHM: We're going to jump right into the data management session.

There were several overarching observations from the session about the survey and about the questions in the session regarding the survey.

One was that the absence of clear definitions really limited the value of the survey in the session and that the participants wanted to delve down into the more detailed questions and detailed aspects about data management and how it was done in different networks.

The participants were also concerned that the findings would be utilized to mandate requirements for future RFAs and

RFPs and that that would limit an organization's ability to be creative or innovative in how they approach operational aspects of trials.

Another overarching comment was that there wasn't really an evidence base with quantitative information for data management operations and that in lieu of an evidence base, we can't really label things best practices.

We have a survey with a lot of good input and a lot of good information there about overall practices, but a quantitative evidence base would really need to exist before we could call something a best practice.

A limitation to the numbers that I'm going to present today from the session was that we had an average of 30 percent non-respondents to the voting in the session and that limits our ability to interpret the information.

In the session, we culled out several key areas of the data management descriptive survey. One was the data center placement, then management and oversight of the data center, policies and procedures, methods used to transfer data, and then utilization adoption and implementation of data standards within networks.

We had about 90 people in the room. An average of 52 persons responded to the polling questions and the range was

from 41 to 72 respondents in the room.

Seventy-three percent of the individuals in the room declared that they were associated with a clinical research network. Our first question was is centralization a best practice. The centralization theme had been brought out in the large report from the survey and we wanted to delve deeper into that.

In our session, 83 percent said that, yes, centralization of the data operations in one data center was a best practice and then there was a very good comment from a participant in the session that said, well, what about the details, what about global and international trials, is centralization still a best practice for those.

So we took an ad hoc vote. Sixty-seven percent said that, yes, centralization in global and international trials would be a best practice and, overall, 88 percent said that they would implement on their network in the absence of constraints, maybe funding, politics, existing structure, roles and responsibilities.

There were some very good session comments on the topic of centralization that I think it's important to bring out. One is that centralization can improve data quality as far as the fidelity in the data processing.

Another comment was that there also needs to be consideration of standardization of procedures that are implemented at the site. That can happen from a central center, but the important thing being that data quality begins at the site with the clinician-patient interaction and how that data is documented in the source at the site.

Other comments were that industry sponsors tended to maintain or want to maintain data ownership and that this may mean a difference between centralization or not for data centers who did industry versus government trials.

That curating the data or applying certain standard terminology sets or coding should maybe be done at the sites. And then an overarching comment was that if you integrate poor quality data at a central center, it's still poor quality data, again, placing the emphasis on the quality of the data and the quality of the documentation in the source of that data.

Later on in the session, we took an ad hoc poll about data quality and it was brought up that if individuals use data to put into a system and they get back information that helps them care for that patient immediately at the point of care, that that use of the data increases data quality because the person documenting needs the information they're getting back, and so the amount that they care about that tends to increase.

The next topic was management and oversight. Eighty-eight percent said that it was a best practice to have an individual responsible for data management on the network. The survey specified a director. The discussion in the room was more around it didn't matter what position and, in fact, people didn't want a percent of the position specified or a position title specified, because that, in itself, would limit individuals in their responses, ability to distinguish themselves or be innovative, but that there was an overwhelming 100 percent said that, yes, they would implement this, having an individual responsible for the data management operations on their network.

Fifty-nine percent said that having a standing committee that dealt just with data management issues was a best practice and then 70 percent said that they would implement that on their network.

On policies and procedures, 89 percent voted, yes, that documenting is necessary or helpful for data quality. A 100 percent said that they'd implement that.

And then there was a large discussion on who writes the SOPs. Do you leverage the infrastructure of the organization providing the data center or should the network itself write the SOPs?

The result of that discussion was that it really didn't matter who wrote the SOPs, as long as all of the stakeholders had input and that the people who were going to use and implement the SOPs had input into those SOPs.

The next discussion was around use of electronic data capture technologies in trials. Seventy-six percent said that, yes, EDC is a best practice. Eighty-eight percent said they would use it if there were no barriers to using it.

Only 48 percent said that EDC increases the cost of trials. So this is really a split in this. Between half and half thought that EDC would increase and then the other half of the room thought that EDC would decrease the total cost of trials.

Ten percent felt that use of electronic data capture increased the total time from protocol to data lock and then 90 said it decreased the time.

There were very important comments with respect to electronic data capture. One was specifically pointed out that we need studies of the cost effectiveness of electronic data capture and, also, that we need studies to evaluate the impact of electronic data capture on data quality.

We also asked about measuring site performance. Ninety-eight percent of the people that responded said, yes,

they did measure site performance and it was a best practice, in their eyes.

A 100 percent said they would implement it in the absence of barriers.

The discussion around this topic was that measuring site performance is really most applicable and most helpful when you have sites that are in trouble, either with data quality or with timeliness of submissions or adherence to procedures.

And a second comment was that in those policies and procedures, it is a critical thing to have acceptance criterion stated a priori so that there is a limit below which sites shouldn't go for their performance and if they do fall below that level, then there is something in place to address that.

The last topic that we discussed was data standards. This is a very important vote, because it really illustrates well where we are in the research industry with respect to data standards.

There were two almost mutually exclusive questions asked to the group in this session. The first was, is developing local standards specific to your network a best practice?

For this question, 81 percent indicated that, yes, it was.

The next question was is using nationally recognized standards, for example, ANSI standards or CDISC or HL7 standards, is that a best practice, and 78 percent said that, yes, this was a best practice, with 84 percent saying that they would implement nationally recognized standards in the absence of barriers.

And so we delved down into that, having very high percentages of local standards and nationally recognized standards and the discussion around that was that in the absence of nationally recognized standards, then developing local standards is good and can be helpful for a network, but that where national standards exist, they should be used and implemented.

In closing, the session participants identified a lot of future research topics and needs for data management. Unresolved issues, for example, the developing data standards and need to continue that; then, challenges in incorporating evolving data management technologies into networks, like electronic data capture.

I would like to thank the co-content expert, Elsa Villarino, and our moderator, Therese Gibson. And, also, many thanks to the participants in this session. There were a lot of great comments and we look forward to getting those into the

summary.

Thank you.

[Applause.]

DR. SACHS: Are there questions from the audience?

DR. DEAN: Mike Dean, from the University of Utah.

I thought you did a very nice job of wrapping up a session that had a lot of discussion that was in various directions.

I just wanted to add something about EDC, though. The votes and the discussion about electronic data capture were in the context that no one could define what electric data capture meant.

There was no discussion about FDA validation requirements when you did direct electronic entry. There was a lot of concern about not having source documents.

So this was not about electronic transmission of data. It was about EDC. So we couldn't even tell if that is what was meant, if electronic transmission was what was meant versus paper.

So I don't think there should be a conclusion that EDC, the way it's usually meant in the literature, which is there is no paper, should be the best practice. I don't think that's what the group agreed upon.

MS. NAHM: Thank you.

DR. FROMELL: Gregg Fromell, University of Pennsylvania. Great talk, by the way.

I wanted to make one point and then ask a question. One is on the issue of SOPs. I think SOPs are critical, but I got a little concerned when there was a mention that the group talked about SOPs specific to a protocol.

So I hope for some clarity. In my understanding, SOPs define overarching responsibilities and processes and policies and procedures or manuals of policies and procedures are specifics on how one implements SOPs.

And I would submit that specific to the protocol of the policies and procedures rather than redesigning SOPs.

The other is on site performance. I think it's great to have that metric, but did the group talk much about using site performance as a way to identify those sites that need more mentoring or infrastructure?

I think often what happens is that performance is you say "You didn't perform and I'm going to move on and use this site that did," but all of us that have been involved in research long enough know that a site that is stellar one year can be suddenly bad the next, because they don't have the right infrastructure, they lost that key coordinator.

So is it really an indicator of maybe that's a site where you devote more time and effort to bring them back up?

MS. NAHM: There was a lot in our discussion in the session about having the acceptance criterion or that bar where that helps you identify sites that are in trouble and needing help.

So I think that was brought out in the session and we'll be sure to get that emphasized in the write-up.

DR. VILLARINO: But I would also point out that it doesn't mean dismissal of the site. It's mostly to try to identify early ways where you could intervene and get them above those criteria and not lose the site.

But in reference to your first comment, I think, again, we go back to the terminology in the literature, because what somebody calls SOPs, somebody might call MOP or policies and procedures.

And that was what we tried to put in that first slide of our presentation, that unless there is common terminology, when you get asked to have an opinion about something and that something is not well defined, it becomes a problem. You don't know if you're talking about the same thing, if you're having an opinion about the same thing.

So I think that your comment is very well taken, as

well as the comment about the EDC. It was exactly the same thing and I appreciate the extra contribution from Dr. Smith. It's that we're suffering from the lack of good definitions in that area and that could help us with our current recommendations.

MS. ZACHARY: Julia Zachary, George Washington University.

I thank you for a very nice summary of the session that we had and exactly those comments that came up I would like to reinforce.

I would actually go a little further to say that there was really so much confusion about the way the question is being completed, that many of the questions in the area of data management are really invalid, I think one could say.

So I don't know that the findings from this particular part were really useful; that, in some sense, maybe it can be looked at again and redone.

And the other comment is you mentioned that 30 percent of the people in the room were non-respondents and that is exactly the reflection of the fact that the questions weren't there.

Thank you.

MS. NILAND: Joyce Niland, City of Hope. I agree it

was a very good summary of the session yesterday. Thanks.

Just one suggestion. When we talk about standards, I'm involved with the bridge modeling process in caBIG and CDISC and I would like to recommend that we indicate that we should all be moving toward adopting national and international standards as we go into global research and multi-center studies across different countries and HL7 is international. We want to be sure we keep an eye on that.

And then just moving into the next session, as long as I'm here, just a general comment that I'd like to see us begin to focus on information management rather than data management and information technology.

I think what we should really be looking at is information management and information and technologies is one of the tools that we should be looking at in that context, along with the human, organizational and other resources that we need to look at.

DR. SACHS: I think that's an excellent point, because there is so much overlap in both sections, and I appreciate that comment. Thank you.

DR. ROSE: Lynn Rose, cystic fibrosis network.

I was wondering if you talked at all about the value of the onsite monitoring from coordinating centers.

I know in our network, we often have difficulty getting enough funding to do onsite monitoring of source data against the case report forms and I think that's a very important part of data quality.

And I was wondering if people in the group that you spoke with had developed risk-benefit formulas for monitoring frequencies, both for EDC, as well as paper.

MS. NAHM: We didn't address monitoring or source document verification part of monitoring directly. The discussion went a little further upstream towards the actual quality of data in the source and getting that data documented correctly and technologies, like decision support, that can help do that.

So it was a little more upstream of the monitoring process.

DR. MORRIS: Perhaps just a quick comment. Alan Morris.

If one uses, at the bedside, an electronic tool for establishing quality and decision support, that becomes the source document and the monitoring demands diminish dramatically, because the data that are transmitted electronically are, in fact, the data that drive decisions.

So these issues are tightly linked, the quality of the

data, the quality of the interaction and monitoring.

MS. NAHM: Thank you.