CLINICAL TRIALS PANEL DISCUSSION Clinical Trials 2005; 2: 352–358

Panel discussion of case study 3

Marc Walton, Richard Simon, Frank Rockhold, William DuMouchel, Garry Koch, Robert O'Neill

Dr Anello: Marc Walton, I wonder if you wouldn't mind going first since you had a lot of interaction with this team during the regulatory aspects phase.

Dr Walton: Certainly. This was in actuality a Phase II study, not the seamless Phase II/III that was initially being considered. But as a Phase II study, it is a good example of an approach to addressing a very important question in the process of drug development, that is, to establish a good understanding of the performance of the drug, including the dose-response relationship of the drug. This question is broadly applicable to many different settings and many different drugs. One of my questions was, what is unique about this setting that made this approach feasible or worthwhile? My answer would be, perhaps not much.

There are certain aspects of this setting that the design really grappled with directly. But conceptually, the idea of wanting to establish the dose response, wanting to explore a broad range of relatively closely spaced doses, occurs time and again in many different drug development programs. There is a lot in this example that will be broadly applicable and that is not unique to this particular setting.

The model of the early clinical evaluation, predicting the 90-day clinical status of a patient, was an important feature of this study. It sped the adaptive random allocation of patients to the dose that would be most informative. I want to note that this model had relatively little regulatory impact. From a regulatory point of view, when we look at what we have learned at the end, we rely upon the 90-day data for all patients. The importance of the one-week and the four-week data was only in selecting which doses to allocate to the next patient. The meaning of those data evaporated as the patients' time in study matured. By the end, it has relatively little importance in interpreting the final result of the study and low regulatory impact. You can view that as a nice way of utilizing these approaches without impinging upon our ability to use the data.

In estimating the dose response during the study, the Bayesian approach allowed you to include multiple kinds of uncertainties that have relatively little final concern for us, for the same reasons.

A feature of the study design that was not emphasized in your talk was our concern about the adaptive algorithm allocating patients to doses that would not teach us most about the dose response curve. We had a concern that, at some point, the utility of adding patients to the placebo group, if this was an effective drug, could become much less than focusing upon doses in the area where the upper bend in the curve lies. This was of concern because in studies like this there is the potential for a shift in the patient population as the different sites enroll and leave. There can be site-related differences related to subjective physician assessments as well.

We had concern that there might not be full comparability, and we might be misled by doses only within a narrow range. This study did include a lower bound on the probability of assigning a patient to placebo, so we could be assured that there would be a reasonably good distribution of placebo patients throughout the study, providing us with a good comparison benchmark. We can think of that as purchasing insurance. It may have slightly decreased the efficiency of learning about the dose response curve, but it ensured that we would still be able to reliably interpret the study in the event of site- or time-related effects.

The final analysis was not heavily influenced by the prior. The expectation was, and proved to be the case, that the prior would be really swamped out the data obtained during the study. From a regulatory point of view, the validity of prior knowledge did not become a concern, because we relied mainly upon the study data.

Coming back to the idea of the seamless approach, although that was the initial plan, it was wisely not the final plan. It is the rare development program where the Phase II study needs to answer only the question of what dose to use in the Phase III. We need the opportunity to

© Society for Clinical Trials 2005

10.1191/1740774505cn106oa

explore the data, and ask questions like what is the right population to employ in the Phase III study to either improve the safety or the efficacy. In this case, critical Phase II design parameters might be the range of severity in the stroke patients, or the right time window from time of onset for the stroke patients, or the concomitant treatments and medications needed or to be avoided. Other important questions are whether the protocol is being followed consistently, or are revisions needed. All of these are very important lessons from the Phase II, and I feel that we need to learn about these things from Phase II before initiating Phase III.

Dr Simon: The first point is that in many cases, Bayesian statisticians are too enamored with computing or philosophical issues, and the action really is in the priors. Bayesian methods, in my experience, really have a major contribution to make when there is prior information or prior assumptions that need to be incorporated into the analysis. Not some subjective prior of the investigator that nobody really cares about anyway, but priors based on data, or on assumptions shared by most stakeholders.

For example, I have studied a lot of Bayesian approaches to problems of multiplicity where there is a community assumption that treatment by subset interactions, qualitative interactions are unlikely, and you can use that to specify a prior. For those kind of problems, you can show that the Bayesian methods give really much more stronger inferences than the frequentist methods, and they work where the frequentist methods tend not to work. But for situations where you really don't have assumptions or data external to the trial that need to be brought into the analysis, my feeling is that the Bayesian methods actually don't have much advantage over frequentist methods.

In this particular example, there was very little focus on the prior in terms of what was presented. I couldn't even understand really what the prior on the dose response curve was. If it's true that it was sort of a noninformative prior, I would question whether that was appropriate. If it was used, then I would think that a similar kind of results could have been obtained with an adaptive frequentist method based on interim analysis.

I also had a question about the posterior distribution. Based on the definition of the ED95, given that there was no effectiveness of any dose, I would have thought the ED95 would be defined as zero. The smallest dose that is at least 95% effective is the best response, and yet that isn't what the posterior distribution of the ED95 looked like.

The second point is, does the Bayesian paradigm make the use of a non-fully randomized design with adaptive allocation more valid than would a frequentist model? My answer is no. Now, it really doesn't matter if it is purely an exploratory Phase II type of thing, but if you want to draw inference, then that issue would come up and the question is, would Bayesian analysis make the absence of a really pure randomization among dose levels more palatable than a frequentist model would?

I would say no; any frequentist can write down a similar model and do a model-dependent analysis. Both the Bayesian and the frequentist approaches depend on model assumptions that the treatment assignment is ignorable, using the phrase of Paul Rosenbaum and Don Rubin. Once you are doing a model-based analysis, whether it's Bayesian or frequentist, the Bayesian paradigm really has no advantage in terms of the adaptiveness of the treatment allocation.

The third question is, are more dose levels better than fewer dose levels? I would say the answer to that question depends on being very specific about what your objectives are. It was said that the objective was determining a dose with sufficient efficacy to take into a confirmatory trial. That is a very different objective than trying to determine the ED95. If your objective is to know if there is a dose that is sufficiently effective to take into a confirmatory trial, you could answer that question with many fewer patients by studying fewer dose levels. You could do that if you are in a situation where you have a prior, and where you put a lot of prior weight on the belief that there is a monotone dose response curve, not some multimodal dose response curve.

The fourth point was, "Do Bayesian methods permit more adaptiveness than frequentist methods?" There are many ways you can formulate adaptive frequentist methods. Many times with Bayesian methods, it is more intuitive in terms of how you are doing the adaptiveness, and that is attractive, but you actually can do adaptiveness either way. If you want to make inference, the issues of nonrandomization come up, and you have the same problems either way.

My final point is: are seamless Phase II/III protocols a good idea? Generally, no. One of the biggest problems in drug development in terms of defining effective studies and doing things quickly is making sure you have crystal clear objectives for each study you do. You can waste a lot of time in drug development without really clear objectives. I think Phase II/III kinds of protocols encourage blurred objectives. In this case, it was not a Phase II/III protocol, but there were still some blurred objectives, similar to what I have alluded to, like the difference between an ED95 and whether there is an effective dose. In this study, I wonder whether you could have gotten to the final answer more efficiently by doing a pilot with fewer dose levels to

try to answer the question of an effective dose, and only once you have answered that affirmatively, to get into what is the ED95.

Dr Rockhold: To state my prior, we have implemented many designs like this, mostly in early development, looking at tolerability. We have run a few adaptive designs for efficacy in Phase II, but these were in areas where we didn't have the advantage of having even a remote surrogate like they did in this trial.

I can definitely see the benefit in the use of decision-making. One of the important aspects of this particular trial was that it dealt with stroke and looked at benefit to risk comparisons. In a disease like this, tolerability becomes less of an issue because of the severity of the disease. If you were looking at a less severe disease, you would have to build tolerability into this model to understand how to pick a dose for Phase III. I assume they were looking at that, although it wasn't mentioned in the discussion. Looking at benefit to risk and tolerability has to be built into the model because the vast majority of studies we do are in diseases that are fortunately nowhere near this severe.

The question about exploratory versus confirmatory has certainly come up. I agree that when you run trials that are trying to serve multiple masters, you run into problems. Making that clear upfront is critical.

A question that I would like to pose is how they arrived at the number of doses and whether or not that number was critical. I try to press people in our institution to run more doses. People tend to work to fewer doses, but most of the medi cations we work on are oral medications. Running trials like this at 15, 20 different oral medication sizes is logistically problematic.

The objective at the beginning was to go into Phase III with one dose. It would have been very difficult to go into Phase III and really understand the benefit to risk at a given dose from this trial. In regard to the futility, and I use Dr Grieve's term here, they would have liked to have reached that decision sooner. They attributed it to an enrollment that was too swift. That may have been part of it, but when we have tried to run futility designs, whether it is an adaptive design or not, they are very conservative.

If you didn't have an independent data monitoring committee, and trial sponsors were looking, I believe they would have come to the conclusion that futility was reached a long time ago, regardless of enrollment.

In order to roll out these designs in the future, and – this is not a criticism of the model – it is important to make people understand the characteristics of futility and whether or not these models are useful.

You are dealing with a surrogate of a surrogate in order to fit this longitudinal model. I was interested to hear Dr Walton say that if they were going to review the data, they would have looked at the 90-day data. In an organization like mine, as Dr Grieve very properly indicated upfront, you are trying to sell this type of design to decision makers. They want to understand, if we make a decision, is it going to stick. If we are making a decision internally, and I can characterize the probability of a decision to them, they will (within the limits of how much they trust me), use that to make a decision.

If the regulators' view of the risk of the decision is another factor to consider, that makes it all the harder to sell. If you are only going to look at the endpoint data, that is going to make the design even more difficult to sell internally.

The recruitment speed issue was a good one. It is very difficult to tell people in an organization that they really ought to be recruiting slower.

From an ethical standpoint, using electronic data capture or the system they have, which is a good surrogate, is absolutely essential. It is going to be very difficult to convince people they need to recruit slower.

Dr DuMouchel: This was a very interesting study that raises a lot of interesting points about differences between being Bayesian or not. Although in some respects, I do agree with some of the other panelists that you don't really have to be Bayesian to utilize the techniques.

The flexibility was a hallmark of the way in which this study was done; there was a flexible notion of which doses to use and flexible models for the dose response curve, and flexible methodology for deciding when to stop. In the textbook studies of clinical trials, that flexibility seems to get suppressed. But I don't think that inflexibility is a fundamental feature of frequentist approaches; it's maybe a fundamental feature of regulation.

There needs to be more interest in how you can be flexible but still be strict enough. It's like parenting; you lay down a lot of inflexible rules because it just makes your life a lot easier. There are several different reasons that the idea of adaptive dose allocation has occurred in the literature. One is to allow this nonparametric dose response estimation to be better. When the adaptivity uses the past responses, that may throw a monkey wrench into certain frequentist calculations. Even there, you could compute frequentist numbers with simulation.

I was involved in a trial where the adaptivity was based on the use of the covariates. It was a surgical treatment, a fetal tissue implant for Parkinson's disease. We had a very small sample size, only a total of 40 patients for the whole study, and we wanted to

get as much balance as possible with things like age, sex and the disease severity. As each patient was presented and being randomized, I used a computation of the amount of information that you would be getting on the slopes and curves of the treatment effect in order to get the best balance in the design. The probability of allocation to treatment or control was based on that.

The idea of a stopping rule being based on posterior probabilities works well. It is again an example where the frequentist approach is a little clumsy and a clear win for being Bayesian. While is not certain whether using the posterior probabilities of interesting events is better or worse than some fancy optimal stopping decision theory, it is probably easier to explain to someone.

The idea of the simulations was very valuable here because of this ability to be both Bayesian and also to be able to predict certain frequentist properties. That is a good win.

The notion of validating the computer code is clearly one that several people have brought up over the course of our two-day conference. It points out the need for more standardized procedures in the Bayesian world, and that will probably happen over time.

There is not necessarily such a need for one-off designs every time. We can hopefully decide on a few Bayesian modifications of practice that are the biggest wins and then standardize them.

This bimodal posterior distribution was very interesting and perhaps an example of being a little too flexible in dose response modeling. Some simpler version of the dose response curve would give a better answer there. The whole idea of an ED95 is called into question if you don't have a monotone dose response curve.

The two greatest defects of the frequentist approach need more emphasis. The first is the excessive focus on *P*-values. The other big defect is the very clumsy approach that the frequentist theory has with respect to multiple comparisons. I have had quite a bit of involvement in a data mining approach on FDA postmarketing surveillance data. We are evaluating millions of parameters, trying to estimate them, and that we need a Bonferroni estimate of every parameter getting estimated seems kind of an impossible situation.

The Bayesian approach is that you don't need to adjust for multiplicity, but you need to use your prior distribution, and the prior distribution automatically shrinks the most extreme values in a way that looks like an adjustment for multiplicity. The use of shrinkage estimation and hierarchical modeling did not receive much focus here in the clinical trial examples. I would encourage people to be a little bolder with the use of historical data. That is an example where some hierarchical modeling could work.

In the past, I have been involved with the combining of experiments in many different environments, which is like using historical data. There was one example where we actually had modeled dose response curves simultaneously with data from multiple species, not to mention multiple designs. That was a most demanding problem.

Finally, there was one discussion earlier about if you have these Bayesian models and they give complex results and how you put all that on the label. One thing that will eventually develop is that the label will just have the URL for a Website where you can go and get all your questions answered. You would put in your own patient characteristics and get a prediction for what this drug will do for you.

Dr Koch: First of all, I want to express to the speakers that I very much appreciated learning about their experience. I thought they provided outstanding documentation of all of the processes that they went through in order to carry out this study. I did have a few questions for people to contemplate. One was, are we dealing with a paradigm in which there actually is a correct dose that works best for everybody, or are we in a scenario where some doses work better for some people and other doses work better for other people? It could be that the bimodal posteriors suggest that high doses work for some patients and low doses work for others. I am not sure how one can proceed to implement the design in a manner that would be potentially sensitive to that.

Another consideration that might apply in other indications is that one would not only be interested in dose response for a primary endpoint; one might also be interested in dose response for some key secondary endpoints. One might not only be interested in dose response for all randomized patients, but might also be interested in wanting to know what the dose response was for males, for females, for younger people or older people. There are ways the Bayesian or other adaptive procedures can be applied which are as informative to dose response across multiple endpoints and dose response across multiple subgroups as if you essentially assumed that the dose response is homogeneous across all subgroups or across all potential endpoints.

My second comment is that you can indeed do a Phase II/III design, but the Phase III part of it should be independent and self-standing. It should achieve certain goals in and of itself, although in terms of ultimately having convincing evidence, it is fine to integrate the two studies. I think Phase II/III is

feasible when there are clearly go signals, but that the Phase III should indeed be self-standing.

One of the curious things I was wondering in terms of implementation was whether or not the study could have recruited a relatively small number of centers. That smaller number of centers may have enrolled about 150 to 250 patients. They could have gotten as much learning as possible from that first stage of this process. If everything looked fully informative, they would then proceed to have the rapid recruitment that would have brought in the additional 300 to 600 patients. This is a way in which you don't have to have an interruption in the recruitment process, but it does allow you to have time to catch your breath before the process runs away from you.

Dr O'Neill: A lot of good things have been said. Drs Grieve, Krams and Pfizer should be commended for doing this study. One can say what else might you have done, but this is an area where we really haven't figured out things, so the dose ranging in the Phase II area in stroke is difficult.

I would hate to think that this was a design-induced negative result. At the end of the day, this actually would have been a good product, but it suffered from a design that actually said it was negative. It is probably not the case, but I have often wondered why outcome-dependent allocation designs have not been used. They have been around for 30 years in the statistical literature, but they have never been used. One of my thoughts is that for modest effects it is not clear whether the effect is due to the covariate or whether it is due to the treatment.

As you accrue, if you do have a situation where different people respond differently, then covariates are as important as the treatment itself. They made sure that you had a fixed allocation of the placebo throughout the entire design. That doesn't necessarily mean that you had the same covariates for that placebo patient as you do throughout the whole trial. If there was a dominant covariate that might have been responsible for the bimodality, then that would possibly explain what was going on.

There is a role for allocation on outcome-dependent allocations, but it is usually for big effects, and not for modest effects. If you are in the modest effect game, it is much harder to tease out dose response.

There is another issue here. Suppose this had turned out to be a positive study, everyone was happy, and you had dose response. In terms of the package, is it a confirmatory study, is it one of the two studies? Would it support approval? I would tend to think yes. This is a large trial, 400 patients. 17 doses is huge. We never see trials with 17 doses. It might have been that they could have gotten by with six doses or five doses, but that is after the fact. One could have gone into this with an inverse U shape. Too much of a good thing is a bad thing; it is not unusual that if you go a little too far, you go off the deep end.

At the end of the day, what I like is that they tried something that folks haven't been trying. All the planning that went into this, the amount of documentation that they provided, and all the simulations, is really a good example of the kind of upfront thinking that needs to go into something that has this level of complexity and needs to get the team involved. They had 18 centers involved, everybody was on-board, and there was enthusiastic recruitment. That is amazing. A lot of positives resulted from this experience.

Response

Dr Grieve: I accept entirely that almost every aspect of this clinical trial could have been done from a frequentist perspective. There are methods to do everything, to do the prediction, to do the stopping, to do all of it. It would probably be more difficult in some aspects, but it could have been done. However, I take the view, as Dr Berry did yesterday, that within the Bayesian approach you have a coherent system within which you can do all of these, so why should I decide to do it in an non-Bayesian way?

Some questions have been made about the definition of the ED95. What is the ED95? Of course, we could have had an inverted U dose response curve. The ED95 was defined as being with respect to the maximum response within the dose interval that we were investigating. Had you had an inverse U dose response, we would have been looking for the minimum dose that would give you 95% of that maximum, because we wouldn't have been wanting to go too high, where you tip back down again.

ED95 was defined in that particular way because of an early study, a patient safety study, in which there was some indication that at the top end of the dose, it was starting to come down again.

Dr Krams: Unfortunately, we recruited much, much quicker than would have been appropriate, and Dr Grieve has done analyses showing that with half the patients over the same time, we would have come to very similar conclusions.

It is important to remember that whilst we did a quite good job in implementing this trial, this was a major flaw. The other issue is that the protocol was written in a very conservative way, and it stated that we would only allow stopping the trial after 500 patients worth of evaluable data would be available. This meant that long after the independent Data

Monitoring Committee had realized this was not going to be a winner, there were still patients coming in.

While we did quite a good job, we could have done a much better one. I want to make a couple of general points and then answer some of the questions. The first point is we had the Copenhagen Stroke Study, and I am really grateful to Tom Skewer and Henry Jorgenson to allow us to work with it, but that is a study from one center in Copenhagen 10 years ago. I tried very hard to get much more realistic clinical trial data. There are 80 000 patients worth of data lying in vaults somewhere. Can we all work together to make that data available to each other?

It was asked why we have so many doses. The answer was that we went to our PK/PD (pharmacokinetic/pharmacodynamic) modeling people and asked them what a biologically sensible step would be. The answer is a combination of PK/PD modeling-driven reasoning, and what can you do without getting into an arena where dilution errors make it impossible to distinguish one dose from the other. We have run a lot of simulations since then, looking at what the advantage of having a very high number of doses is versus a smaller number of doses. All in all, if we have to do it again, we would probably run it with nine treatment arms.

There was a very strong impression that this might be a drug where we would have an inverse U shape, and that is why it was important that we use something like the normal dynamic linear model, which is not dependent on the assumptions of monotonicity. We didn't discuss all the aspects of safety observations. The independent Data Monitoring Committee looked very carefully at all the safety data, and the antibody data coming in, and was very much better suited to make judgments on a fuller understanding of the dose response.

If I was asked whether I wanted to have a full understanding of the dose response in the future drug development versus just understanding what the safety was on a particular dose, I clearly would go for the full characterization of the dose response again.

Audience questions

Dr Gould: I am Larry Gould from Merck. First of all, I would like to commend Pfizer for supporting and sticking with a fairly complicated and elaborate technical development scheme over a long period of time. That is really quite a commitment.

Second, I think there were two comments made in the discussion that everyone ought to tattoo on their hands. One of them is Richard Simon's comment, which I will paraphrase, that before you undertake each stage, you ought to have a pretty clear idea of the statements you want to be able to make after you have completed the stages and done your analysis. In other words, knowing what you were looking to get at the outset is a good thing. The other thing is Marc Walton's comment that Phase II tells you more than just what your dosage should be for Phase III, and it has a very real value.

My question is what would have happened if, given the data you had, you had actually pursued something like a more or less conventional frequentist type approach, maybe a little imaginative, using group sequential stopping rules and such? Would that actually have led you to the same kind of conclusion with about the same economy of patients?

Dr Grieve: I haven't yet done that piece of work, and when I have some more time, I will do it and I will tell you what happened.

Dr Gillespie: Bill Gillespie, Pharsight. The main question I had is early on when you had first presented this work, back in 1999 and in some subsequent publications, you described the decision analytic strategy for the stopping rule. It was certainly an intriguing approach and one that I am interested in pursuing. I was curious as to your reasons for going to the posterior probability based approach.

Dr Grieve: The reason that we ended up going for that approach was because the objectives of the trial changed during the development of the algorithm. When we first developed the decision theoretic approach, we were looking to detect bigger effects than the ones that we ended up trying to detect, and the decision theoretic algorithm didn't seem to be as sensitive to very small effects as did the one based on the bounds of posterior probability.

Dr Krams: One of the reasons why we didn't go the decision analytic route–and Dr Grieve and I had big battles about this–was that we thought it would be very difficult to explain. That is a shame. There is a huge need for an information sharing, just as we do it here, so that in future research programs, the difficulty with explaining doesn't, in itself, become a stumbling block.

Dr Rubin: Some experience suggests that when you get bimodal posterior distributions, it is a mistake, a mistake in one of possibly three senses. The most common is, when using these really expensive Markov chain Monte Carlo runs, you haven't got convergence yet, and in order to detect that, you have to run multiple chains, perhaps millions of iterations, and wait for the multiple chains to converge to each other.

A lot of experience suggests that that can happen. Even if that step is correct, even if you have the correct posterior distribution for the model being fit, bimodality of a posterior distribution often reflects the fact that you have chosen a model that is incorrect. It actually suggests that you should do something to fix the model. It is an important diagnostic.

The third sense in which it indicates a mistake is that there could be a hidden covariate that you are not modeling. It could be two groups of people, and there is a mixture of two types of people, and that is not really a mistake in the sense that the model is wrong, but it is a mistake in the sense that you haven't included something in the model that you should have.

Dr Grieve: I certainly agree that if are trying to model a dose response curve in which you anticipate either that it is monotonically increasing, or at least it's an inverted U, and it turns out that you don't have that, and it is just random variation. Then, using an ED95 might not be the most appropriate model to use.

In terms of the covariates, clearly when we specify in a protocol what covariate we are going to use to do our primary analysis, that's it, we have chosen it. After the event, we put in just about every covariate that my colleague here could think of, and it didn't change it. We still had a bimodal result. There may be some parameters out there, which even the greatest brains, clinical brains in neurology can't think of as being influential, but everything that we tried, it still stayed the same.

Dr Simon: Why isn't your ED95 zero? You have no treatment effect. You defined your ED95 as the smallest dose that gives at least 95% of the full response. The full response is zero. The smallest dose that gives 95% of the full response is zero.

Dr Grieve: The full response is not zero. If you saw that the placebo response was about 17 points, the maximum response was about 18.5, so there was a maximum response, and we are defining the ED95 with respect for that maximum difference.

Dr Simon: The way you are looking at that posterior distribution is very inconsistent with the conclusion for futility.

Dr Krams: The decision problem that we had was whether there was a three point or greater additional benefit over and above placebo.

Dr DeBrota: Dave DeBrota. I work at Eli Lilly. It occurs to me that Pfizer has done a wonderful job of showing a way that new things can be done, and I am very curious about the future of the source code that has been created. Will this continue to be a proprietary asset of Pfizer that no one will see, or will this source code be subject to greater scrutiny, or perhaps even released into a domain in which it can contributed to by many different institutions?

Dr Grieve: Since we ran the trial, we have developed the same approach on our other software platform, so we could do what we did with this home grown system using WinBugs now. In theory at least, it will at some point in the not too distant future become available more generally.

Dr Louis: Tom Louis, Hopkins. Dr. Rubin provided a long list of reasons that you might get a bimodal posterior. It is important to add that whether you are talking about the likelihood, the posterior sequentially monitored studies have the ability to come up with a bimodal posterior even in the most vanilla case with no covariates. That needs to be added to the list.

Dr Anello: I would like to thank the speakers and the discussants for really excellent presentations.

Participants

Marc Walton: Director, Division of Therapeutic Biological Internal Medicine Products, FDA
Richard Simon Chief, Biometric Research Branch, NCI
Frank Rockhold: Senior Vice President, Biomedical Data Sciences, GlaxoSmithKline
William DuMouchel Chief Statistical Scientist, Lincoln Technologies, Inc.
Gary Koch: Professor and Director, Biometric Consulting Laboratory, Department of Biostatistics, School of Public Health, University of North Carolina at Chapel Hill
Robert O'Neill: Director, Office of Biostatistics, Center for Devices and Radiological Health, FDA