Never formalized
done as an intro
to history
of Clinical
Trials

Bob Temple and Robert O'Neill Interview

TAPE 1, SIDE A

A1: So when everybody is being an idiot and doing stupid studies, he's [unclear]. And we're asking for studies and asking for really good studies, he thinks we overdo it.

SJ: We're asking too much.

A1: It's really -- he's a contrarian.

SJ: Well, it started from [unclear] with the NAS-NRC study. Those just ticked him off.

A1: He read it. He read it.

SJ: Well, but just, they weren't doing anything according to his standards.

A1: Oh, I see.

SJ: It was just a mess from his perspective.

A1: Did he think they were too easy?

SJ: Oh, yeah. He thought they weren't rigorous enough, he thought they were rushing through the whole project. It was a disaster. But . . .

A1: Well, he's right. We botched [unclear].

SJ: Yes. But our cynical look was that, you know, sort of deprived him of work in the later part of his

career, because he would have spent his whole career trying to get 10 of those processes or whatever.

But implementation; once again, implementation is important because the NAS-NRC was not obliged to take -- they could take any kind of trials in mind that they wanted to. They weren't required to look at any clinical trials at all. They could look at and make their own decisions. So they had actually plenty of flexibility, and they always went -- I mean, the only drugs we took off the market, as I understand it, were ones that were ruled ineffective. We didn't take, go after the probably effective or the possibly effective.

A1: Sure we did. That's not correct.

SJ: Later?

A1: Yeah. Their job was to rank them as effective, probably effective, possibly effective, or ineffective. We, for the most part, accepted it when they said effective. And it's perfectly true that if you look closely at the evidence, we might not have loved it so much. But if it was any of the others, they had a certain amount of time to get themselves up to effective, or they were going to go. And so we implemented the actions on the rest of those. That's what we spent decades doing.

SJ: Okay. But they still had a chance at it again, I mean . . .

A1: Yes. Even if it was ineffective, we still had to publish a notice saying that it was, and they could still challenge . . .

SJ: Right. And they could appeal it and challenge that.

A1: Right. But on the other hand . . .

SJ: But only if they'd done the clinical trials.

A1: Well . . .

SJ: Plural, as I understand it.

A1: That's, yes, it was plural. And the legislative history says they meant it to be plural. But we never said what an adequate controlled study was until 1970, and the reason we said was DESI. Okay? It wasn't [unclear] to good trials, it wasn't any of the really good reasons you'd name now. But whoever was chief counsel -- I guess it was Peter Hutt at the time -- saw hundreds of hearings coming, and that would have been ruinous. There's no possibility of doing all of these. That would have been before the ALJ. And he said, "That can't happen."

So we wrote -- one of the things about a hearing is you do not have to have a hearing if there's no issue in fact. So what we wrote was what a well-controlled study

was so that we could say, if it was true, that there is no issue in fact here. It doesn't have any of the characteristics of a well-controlled study. There's no bias reduction, no control group, blah-blah-blah-blah-blah. And if we could say that and write it in a notice, which we wrote in hundreds of notices, all of which at the time went through me -- I was the signoff under Crouch for the DESI case -- if we could say there was no substantial issue in fact -- I think it was substantial issue in fact -- but anyway, if there's no issue in fact, then we can deny the hearing. We denied hundreds of hearings. But we couldn't do it until we got into the world a rule that said what the features of [unclear] was.

SJ: And that's where Beardsley came in and said . . . I read the regs for that.

A1: No, not Beardsley.

SJ: B, it began with a B.

A1: I don't know . . . Beaver.

SJ: Beaver, thank you.

A1: I don't know who wrote the regs, but the regs resemble an attachment to Henry Simmons' speech that Bill Beaver wrote, about a 30-page -- you've probably seen that -- 30 pages on what constitutes a good study, and the regs

tracked that very well. So we put the regs out as a final or interim final or something.

The PMA, PhRMA's predecessor, challenged it and said, "You can't do that. You have to propose it." So we pulled it back, proposed it, and got it final. All of this took place in less than a year. It would take five years now. But that's how . . .

SJ: That's Peter Hutt.

A1: That's why we wrote the regs on adequate well-controlled studies. They were very [unclear] '70.

A2: Interestingly enough, they came in '70, before you and I even came to the agency.

A1: Absolutely. So Beaver knew a lot.

So whoever wrote those, those established the scientific principles for the design and analysis of clinical trials. We modified them a little bit in '85.

A2: So that was important.

A1: That wasn't around in '62. It was only till '70. And it took us another 10 years to put meat and bones on what we meant by those principles.

A2: And understand [unclear].

SJ: And that's what I wanted to . . . [unclear] when you got started, and really just reminisce for me so that I can check out some things, of who was in the division when

you came, a little bit how it was organized, just to help me a little. If you can remember it, fine. Whatever you can remember is fine. I'll be checking out everything.

A1: Well, I came in July of 1971, and I was hired by Chuck [unclear], who was my boss, who was at that time the Director of the Division of [unclear] called Biometrics. He had recently inherited a situation which was early -- there was the old Bureau of Chemistry, and there was a big reorganization that went on, and they moved from Crystal City to the Parklawn Building. And at that point in time, Henry Simmons, I believe, had recently been a Bureau Chief, Dick Crout had yet to come in probably for another year or so. But at that time, there were only six medical review divisions. As we know OND right now, there were six medical review divisions.

A2: Compared with 17 now.

A1: Compared with 17. And it was essentially metabolic and endocrine; cardiorenal and pulmonary -- they were both together -- there was an infectious disease; probably something on radiology and dental stuff; psychiatry; and that was pretty much it. There were six divisions. And it was -- I was one of about seven new PhD recruits that Nella [sp.] brought in. All of them probably hung around for another few years, but all of them left

after about five years and went back to the university or went into industry or whatever.

But at that time, it was a different place. Most of the clinical people in the divisions had come from practice, and they were not researchers or methodologists, so they were really very unprincipled in how to evaluate a clinical trial according to the 1970 rule.

A2: A huge number came from the Public Health Service. Their hospitals closed and they hired a bunch of people, embarrassingly poorly trained.

A1: And so my job was to, in early statistics, was really to understand the nature of the problem, understand just the environment of the [unclear] at the time, and essentially was in an uphill battle to educate these clinicians, because the statisticians were viewed as hired guns to turn down a drug, because at that point in time everything was really pretty atrocious. There were very few studies that were in the NDA's that had followed what we would call the 1970 Adequate Well Controlled Trial Principles. And there were like five or six of them, basic, have the protocol written before you start the study; randomization; ascertainment; blinding; analysis, and stuff like that, basic core . . .

SJ: And that's eight years after the '62 amendment.

A1: Yeah. That took eight years for that. And whoever wrote that, just so we could do this legal administrative thing, was really on the money, because that established the principles for how we've been operating for 40 years.

A2: It was probably John Jennings, you know.

A1: He could have. Whoever did it . . .

SJ: That was a lost cause. We should have talked to him.

A1: Whoever did it really had their finger on what was very visionary in terms of how it would eventually be, because we took it and ran with it.

A2: Right. But it's all laid out in this Beaver document that's attached to Henry Simmons' talk before somebody.

A1: But the other piece of this was the principles say nothing about statistics per se. I don't believe it says -- it just says there is an analysis that whatever demonstrates that there is the effect that is before you.

SJ: And that's one of my questions, is -- and the answer may be very simple. It may have been back . . .

A2: Well, where did [unclear] come from?

SJ: That's it. That's what I want to know.

A1: The answer is not simple. The answer is essentially, the answer comes out of essentially what I would call the statistical analogy to the practice of medicine at that point in time. Those were good standards, essentially, applied to clinical trials coming out of the scientific community and what was known about clinical trials.

If you look at the clinical trial itself in 1970, it was only about 20 years old at that -- it was a very young vehicle. It had come out of the U.K., The Principles: Bradford Hill Clinical Trials.

There were a couple of clinical trials that had been done in this country. You saw a polio vaccine trial in 1954. It was probably the largest randomized trial that we've ever done.

A2: You mean by '62 or by '70? By '70 there were a bunch.

A1: No, no. I'm talking about where they began, and the NIH's ability even to sort of -- there were a few that they started, you know. That was the diabetes trial . . .

A2: UGDP.

A1: UGDP.

A2: That was in the '50s.

A1: Well, no. That was probably '60. But even the Salk polio vaccine trial . . .

A2: Yeah, yeah, a good example.

A1: . . . where essentially -- that was 250,000 kids were in those studies, but only about 60,000 were in a randomized component. The others were sort of observational components. And Paul Myer has written an interesting piece about that.

But the point about where does .05 come from, .05 came from primarily the statistical and clinical trial literature as we knew it then, which was good practices and procedures, and essentially the idea of how do you control the chances of making a false-positive conclusion. And as a level of evidence by itself, .05 is weak. It was the law that we interpreted for replication that tied .05 twice that really raised the bar in terms of the standard for the level of evidence, that you could be fooled once by one study, but you're not likely to be fooled twice by the same study at the .05 level.

A2: Yeah. I mean, with only one, you have officially one chance in forty of finding, agreeing that a spurious thing is true. When you have two of them, your chance is well over one in a thousand.

A1: The other . . .

SJ: And that was known from the trial literature or experience?

A1: Yeah. The other is [unclear]. The other place it came from was, we interpreted -- and this is another interesting, I actually talked about this because I think David Feidel [sp.] at one time sort of inferred this -- but we interpreted the 1970 laws . . .

A2: Sixty-two?

A1: No, '70, Adequate and Well Controlled Trials Principles, as the framework for the design of a prospective style, a prospective study that was designed according to hypothesis testing principles as opposed to I want to estimate the effect, as opposed to I want to detect effect. And that was a huge thing.

We may not have realized what we did, but it was what the practice of clinical trials was at that time. You sized, sample sized, a clinical trial according to the principles of hypothesis testing. And that meant you posit a null hypothesis; you posit the effect size that you want to be able to detect, the alternative hypothesis; you posit what you what to control the false-positive radar; and you posit what you want, if you're right, you're correct, what's the likelihood that you will have a successful trial. Those four pieces.

A2: But the alternative hypothesis in every trial that anybody's ever done is more than zero.

A1: No, no, no, no, not -- you can't size the trial. Every trial . . .

A2: I didn't say that's [unclear].

A1: No, no, no, no, no.

But, again, that leads us to, that was the framework of where .05 came from; that's the framework from where .05 came from. So it was the literature, the books, everything that was out there at that time. That was essentially the standards, and we were just taking the community standards and implementing them. And that's an important issue because . . .

SJ: And industry never challenged that, because it seems to me like if there was a point that industry challenged, that would be [unclear].

A1: Well, they did. Actually, they tried. In fact, I have a volume that Lasagna had a meeting in the mid-'70s. They tried to revisit, let's loosen the .05 criteria. They had a whole day's meeting on that. And so he went with the wind. In fact, he was probably, in the early days, saying, this is a piece of crap, everything out there. You've got to do it more rigorously. And when the wind went the other way, he went with that crowd to try to loosen . . . And

most of that crowd didn't even understand what the basic principles were and really didn't tie it into the value to replication, the value to what .05 is doing and whether it is a . . .

But at the same point in time, Chuck Anello, the guy who hired me, in 1970 wrote an article in what was then the Bureau's current consumer magazine, whatever our magazine . . .

SJ: *FDA Papers*.

A1: The *FDA Consumer*, probably.

A2: *FDA Consumers*, something equivalent to that -- the difference between clinical and statistical significance, because way back in time, that issue was on the table, and it still is. We are revisiting that issue 40 years later when we essentially say, we don't clinically stand behind an effect that's too small or too big. It's only when there's a safety issue.

So one of the issues we're facing right now, as we're putting more stock in what the effect size is, because we have to be because now we're in the game of not being able to do placebo-controlled studies, and we have to have a pretty good fix on what the effect size is. Well, we don't have a very good fix on what effect sizes are because, like Bob says, the hypothesis testing framework essentially

says, if I can demonstrate that I have an effect that is not a chance finding, meaning that it is significant statistically at an .05 level or less, that's good enough from a regulatory perspective to assure us that there is an effect. The secondary issue . . .

A1: But before you leave that, what that means, though, is that you've shown for sure that the effect is more than zero.

A2: Right, right.

A1: And there's also a [unclear] you've learned something about effect size, but what you've shown, if you show something at .05, is that it's more than zero.

So the argument went both ways in those days. Someone would say, yeah, you've demonstrated a statistically significant finding, but it's not clinically important; it's too small. And that argument has gone 180 degrees around the other way, people saying that or the statisticians saying, this is an effect size that doesn't matter, because the only reason it's statistically significant is because you had 2,000 patients in this trial and you only needed 300, and if you had 300, you wouldn't have found this. So you've got a real effect that's so small, it's non-meaningful. And then you've got to get

that backed up by a clinical decision that, yes, it is, but we don't make that decision very often.

A2: No.

A1: And so that is very interesting as we get into the game of weighing benefit and risk, because there's also, I mean, there's a legislative historical reason why we don't make that argument.

A2: Yeah.

A1: The legislative history is crystal-clear. It says, we do not have a comparative effect in this requirement. So we don't, we're not telling you that it's better, we're not even telling this is good. We're just telling you it's worked and we've described it accurately and all that.

The thing is, we're more inclined, as Bob says, to worry about not being very good in a bad disease. In a minor disease, runny nose, we're still not going to do a relative effect on this thing.

But we just turned down a couple of drugs for antipsychotic, an antipsychotic drug and an antidepressant that were clearly inferior to everything else. We said, that's not a good deal, you know, to have an inferior, a way inferior antidepressant. That's not acceptable. But, as Bob said, we're evolving. We're thinking about it.

Only because the problem we've been dealt is that back in those days, very few effective drugs -- and the 1970 records were being used essentially to weed out and allow for a more systematic, rigorous, predictable, repeatable strategy, and we're very much faced with a fair and equitable application of that strategy to everybody: you treated me differently than you treated that guy. So that's another reason why we didn't always sort of say, well, your effect is bigger than that guy's, so I'll take you and I won't take that other guy, so . . .

A2: Let me mention one other thing.

I tried to put some of these down in the Sterling lecture that I gave on Lasagna's death. But decisions we made about what to worry about, certain things, in my view, have shaped the way everybody does things. It shapes what the journal articles ask for. And the reason is, it tells drug companies what to do, and drug companies do most of the trials in the world, so when they get something, they make everybody do it. And among those things which were absolutely not thought about are counting all patients. You know, in various ways, through various [unclear] goes down through the [unclear] trial, through other things, we established the principle that you've got to account for

all patients. Now, that's in the consort statement and stuff like that. But I think we did it by . . .

A1: Not only that, but actually, I don't know if you have it, but I actually scanned into PDF all of the administrative hearings that we had to testify at. I had Ed Nevius and two or three other of my statisticians, when we actually went to the mat, before Judge -- he's still here, what's his name . . .

SJ: Davis.

A1: Davis, Davidson.

SJ: Davidson.

A1: Davidson. He presided over the large majority of those administrative hearings, and they all came down to the interpretation of those principles, and essentially, most of them were won or lost on the statistical arguments. And I have those files, if you need them or want them. But out of those files were essentially all the stuff that we're dealing with right now. Everything that we know today that is in ICH E3 and ICH E9 came out of our experience during the '70s, when we applied the 1970s criteria to whether you won or whether you lost an administrative hearing. And all that case history essentially was digested by us at the end of the '70s, put into an internal guidance document to ourselves, because

there was nothing out there in the street, no books, no nothing about what we did, and that eventually turned into the 1986 Clinstat guidelines.

A2: The thing is, because of where we are and because we regulate such a large fraction of all the trials in the world, what we were doing to the industry affected everybody. So it isn't just trials of drugs, it's trials of everything. This has become standard knowledge to everybody, and it's because of where we were. We could make the industry do what we wanted because we could say yes or no, but it affected everybody's views of what constitutes a good trial.

SJ: And we also trained enough people to move into industry that understood the principles.

A2: And it's all these things [unclear], but it's accounting for all the patients, don't form hypotheses afterwards, don't go nosing around the data.

There are some fantastic cases that I'm sure Bob is referring to where they fail on the [unclear] study and say, "Well, if you look at the severe ones, it worked."

My favorite is one where they had two consecutive studies and they both failed, and in the first one they say, "Well, it worked in the severe," then the others say, "Well, it didn't work in the severe, but it worked in the

mild and moderate." You know, they would adapt the hypothesis to whatever was going on. [unclear].

SJ: Are you the one that talked about a placebo that worked better than the drug, so they wanted to know if they could market the placebo? [unclear] anyway.

A1: No.

SJ: I'm sure it was in jest, but . . .

A1: But it's very interesting. I found that the science -- I don't know whether I [unclear] there, but the Science Forum talk that I gave two years ago was essentially on this theme, how the FDA has influenced the practice of clinical trials, and essentially, we had no choice in the matter by the nature of what we do for a living. And because if you look at, the NIH was the major trial influencer back in the early '70s, but that morphed dramatically over the next 20 or 30 years. And most of the clinical trials, as they are right now, are done by the industry, maybe in conjunction with the NIH. But the NIH is not a major player in either methodology or standard setting anymore. It is the regulatory process that . . .

And our role in training the people who do this stuff, it's a very dramatic and dynamic thing. The people in the game essentially attrit after 10 years or 15 years, and you've got to train a whole new crowd. And I saw this very

much, and NIH was the major sort of trainer of clinical-trial methodologists back in the '70s. They had some money, they had a program, and then that stopped, and you had a whole bunch of folks, when the end of 1970, in the cancer area, that didn't want to do clinical trials. They wanted to do observational studies based upon databases that they had banked. M.D. Anderson was a standard crowd for that. And it wasn't but for the NIH statisticians standing up and saying, "Look at how you can be fooled by this," and essentially held the ground for the principles of clinical trials and un-blinding and the value of randomization and everything else, and how these databases could not give you the answer, because people wanted the answer tomorrow. I'll go in, use these databases.

About 10 or 15 years ago, years later, we revisited the exact same argument when the AIDS crisis came along. The only difference was we had a different crowd in the game. You had a different bunch of investigators -- not the oncology crowd, you had infectious disease docs, and a bunch of them that had never done a clinical trial in their lifetime. They were the neophytes and they were saying I want to do this, I want to . . . But for the FDA, which was sort of the staying power behind, you know, this is not going to satisfy their 1970 standards.

So we've been sort of the staying power through the clinical trial evolution, if you will, and the things that were only known tangentially then, which we've now spent a lot more time on, which weren't even referred to in any of the regs, probably were not dealt with in any of our early guidances, such as what do you do with dropouts, how do you handle multiplicity if you search and you retrieve, and all kinds of things that are more subtle issues that are part of the interpretation of evidence, and even evidence under conflicting evidence. What's the principles that you apply in that?

The value to the FDA as an institution was that it was both a regulator and a trainer and an educator, and essentially raised the standards for all the votes, because if the industry couldn't have done it and the academic world didn't do it, and there's no place . . . Just like the critical path today stated that no one has ownership for the development of drug-development science, it's not taught in the academic world, it's essentially FDA's development of that science and the need now for that to be codified in some way in terms of being brought into academic programs in more of a systematic program.

But the industry, the industry, for at least 20 years, was an industry that was 10 years behind the time, and you

ran the industry depending upon what you learned 10 years before. So the whole '70s was a disaster. And so FDA was in a turndown mode. They turned everything down because everything was a disaster.

And it wasn't until the '80s that the people who learned during the '70s about the turndown, what the lessons what and what you had to do to do it right, then they started to populate the industry, and the idea of, "Oh, I need a protocol in advance, I need to think of all the endpoints," that didn't take hold until the '80s.

And then those people who had learned that in the '90s sort of refined this, because a lot of these people were one-at-a-time study. They probably never worked on more than one study. So the whole idea of doing drug development, which is just not this little study that you're worried about, you've got to think about how do they all fit together. There's hardly anybody in the industry that thought that way.

A2: But they've now become extremely knowledgeable.

A1: But now they have. But what I'm saying is, the interaction, the interaction between the regulatory agency as both holding the standards and essentially influencing the practice of where the country goes and the quality of

evidence, you can't separate both of those, you can't separate both of those.

A2: And the increasing likelihood of meetings in the course of drug development just enhances all that.

A1: Yeah.

A2: It shortens the time interval between things, allows us to work to discover problems early and work on them internally.

SJ: But my perception of that has been that the meetings with FDA were almost held out as rewards. You didn't just meet with anybody who didn't know the first thing about what they were doing; that you met with selected . . . You didn't tutor them from the very beginning. You came in when they knew at a certain point what they were proposing.

A1: Well, no, I don't think so. I think we've always thought they were most helpful for the people who were worst.

SJ: Oh, okay.

A1: So they've been growing -- we're trying to control the number because it's killing us, but it's not the reward for being good. It's something we try to do at a time when it can help. In fact . . .

SJ: It's more resource-dominated, then? The issue is more resource-dominated?

A1: Well, we have, I don't know, it must be close to 3,000 meetings a year, and it just takes tremendous amounts of time, and companies recognize the value of them getting an early read on these things. So, it multiplies our influence, multiplies our work. We're trying to control it a little bit, but we're still highly committed to helping them get the trials right so that you get a good answer.

SJ: Let me ask you one question. I should have asked you before, but we were going so, whatever.

If Lasagna had won, I mean, if you had adopted a different standard than .05 for the p value, would that have been done in regulation or just practice or . . .

A1: [unclear] in regulation.

SJ: P value is not written down anywhere, so they would, you would have had to personally persuade you guys to change it?

A1: I think what . . .

SJ: That's what fascinates me about this memo, Abrogado's memo.

A2: It's like how you would be treated in a standard of care because it would be a poor standard of care. It wouldn't be a Lasagna, I mean . . .

SJ: So, it's in the context of medical practice.

A2: Yeah. It's bigger than Lasagna. It was a worldwide standard for what was held to be good evidence in a clinical trial.

A1: I think what we wanted was some intelligent flexibility, and he's not wrong.

If, for example -- people have done this -- if you had two studies, each of which was .053, that's a level of evidence really just as good, almost, as what we [unclear]. They're not stupid to think that we might have been flexible. We [unclear] bureaucrats. We're disinclined to be flexible because we're always worried about the slippery slope to oblivion and all that stuff. But what bothered him -- I knew him pretty well -- what bothered him was any sense that we were getting the wrong answer because we're slaves to some not quite fully-thought-out principle.

So, on one of these he said, he used to say, "Suppose I had nine straight studies, each of which was significant at .1." Now, that's impossible, I mean, but leaving that aside, he said, "Wouldn't that be a very powerful evidence the drug works?" Of course it would; it would be overwhelming. It could never happen, but it would be overwhelming.

And I think we're slightly inclined to think about those as possibilities, whether you call yourself being a Baysian or something like that. We're a little more inclined to think, to bring other information to bear and on our guidance on weighing evidence, so we're [unclear] about that. But, you know, there's a little more flexibility than you might think.

A2: Bob is correct about that, and the statistics of it is not as hard-and-fast and rigorous and unmovable as a lot of people have thought.

On the other hand, even the .05 standard, as we applied it, was relatively understandable and [unclear] everybody could know in advance, but for a long time we interpreted, and the industry ran their business the following way, at least from, in my mind, all of the '70s and probably through a large part of the '80s. The pharmaceutical industry would throw their sheets to the wind. They'd do lots of studies, and all they'd do is search for two that would make it at the magic .05 level. And we didn't blink an eye about the ones that didn't make it. We just searched for the ones, the two that did, and that is naïve and not a good practice either.

And it's only now, in the sort of late '90s and the 2000s, where clinical research is so costly to do, they

can't afford to throw the sheets to the wind, even though there are probably called the marketing studies and other studies done primarily to spin the product. But, in general, we now recognize that all studies that are well designed count, and we worry about inconsistencies in the results much more so than we did 20 years ago, where we essentially said, okay, we'll look for the two, particularly when we were in a situation where the Center was very much still not run by methodologists, people who doesn't satisfy the ranks, if you will. And so understanding how much variability and how much inconsistency, if you have 10 studies and they're not all at .05, some of them .05, some .1, some of them are like nothing's going on, which is nice and as clean as another study where something did show a difference, how do we interpret that?

SJ: I was going to say, how in the world . . .

A1: Well, so that's the hard part. How do you interpret what might be on its face inconsistent evidence? How do you deal with heterogeneity? We're dealing with that now. That's the year 2000 version of where we might have been when we're in a different place in the 1970s, where we were just trying to clean up a mess, where there was no standards that were really being applied.

SJ: Okay.

A1: Usually, though, in practice, like you have an antidepressant, you get three wins out of six; about half of all antidepressant trials work and half don't.

SJ: I don't understand antidepressants because they either work or they . . . Never mind.

A1: No. There's a lot of improvement in the placebo group. That's the trouble with an antidepressant.

SJ: But Prozac revolutionized the literature. I mean, *Prozac Nation*. And now people say Prozac is the one that doesn't work.

A1: No, that's bullshit. Anybody who says that is a moron.

SJ: Good, thank you.

A1: It works [unclear].

SJ: I have that on record.

A1: Well, anybody -- we can go into how do you know it works.

We don't usually see two out of 20. That would be huge. That would make you nervous. We've seen one recently, so it's [unclear], and we turned that down. We said we weren't sure that they showed that it worked.

But it's usually like three out of six or something like that. And if you actually calculate the chance of

getting two studies at .05 out of six, it's still a pretty rigorous demonstration. I don't know what it comes out to, but I'm sure it's well under 1 percent.

A2: But, I mean, part of this is, we are not inflexible. Our thinking changes with the environment we're in, and we're in a different environment right now. We're in large, you know, I mean when you look at large outcome studies and the number of large studies that we have now versus what we had 20 years ago, it's dramatic. I mean, the Vioxxes and these 8,000-, 10,000-, 15,000-, 20,000-patient trials, sure, before the Vioxx trial, no nonsteroidal anti-inflammatory drug had ever been in a trial with more than a thousand people in it. Right? It just wasn't done.

A1: That's true.

A2: We're now seeing safety trials with 5,000, 10,000 people, and outcome trials to try to show effectiveness, a size in scale and number that is completely unprecedented.

Look, until about 1975, no drug company ever did an outcome trial. The first outcome trial I know of by a drug company was the [unclear] trial . . .

SJ: Say that more slowly.

A2: Drug companies did not do outcome trials. NIH did outcome trials, the VA did outcome trials, the Medical

Research Council in England did outcome trials, Scandinavian countries did outcome trials. Drug companies did not do outcome trials. That has completely changed.

SJ: And the first one you remember was?

A2: Oh, the first one by a drug company was the Anturane re-infarction trial.

SJ: An . . .

A2: Anturane re-infarction trial. I published a paper on it.

SJ: Okay, got it.

A2: With Gordon Fletcher. That was the first one ever. It had a lot of problems, it turned out.

But now -- that was in 1975; that was the first one -- now, I don't know, we see 40 a year, 50 a year, maybe more.

SJ: And that pertains to what you're saying. In other words, industry is always claiming that the cost of these trials is outrageous, but is that because they are electing to do these large trials hoping to see a better . . .

A2: Well, the cost of those trials is very high. They're also complaining that the cost of doing a depression trial has become very large too, which is why they take [unclear].

SJ: But are they the ones choosing to have these large trials, or . . . I mean, I assume we're not requiring . . .

A2: [unclear].

SJ: They're trying to show more evidence on a larger scale?

A2: Some of them we require to evaluate safety, and some of them we require to get something out of your labeling that says, yes, I lower cholesterol, but I don't know if that does any good. They don't want that in their labeling. So they have to do a trial to show [unclear].

In the Vioxx, in the bigger class study, those Celebrex trials, for example, all they had was an indication for pain, reduction of pain in arthritis. They wanted a bigger claim. They wanted a claim that not only do I do that, but I don't give you a stomach ulcer at the same time, because all these nonsteroidals were labeled with [unclear]. They had a stomach-ulcer claim, they didn't have a major-bleed claim. But essentially what they wanted was, you know, I'm safer than the other guy, which is a big marketing advantage. They wanted to replace all of the other drugs. Okay? So, in order to do that, you had to design trials that would be able to demonstrate an

effect that might be, the background rate is 3 percent, and I'm going to reduce it to 1.5 or 1 percent.

In the course of doing that, they had to size that trial up around 5,000 to 8,000 . . .

SJ: Because it's not a common occurrence.

A2: Not a common occurrence, nor is a heart attack. But it was frequent enough that now they were able to distinguish that not only did you reduce bleeding, but you increased heart attacks by 50 percent.

SJ: And as I understand the current state, we now know that they concealed earlier trials they did [unclear].

A2: Well, that's another story, whether they did.

But the whole point of this is, as the sophistication of larger trials went on, it was primarily so that the sponsor could get an indication that they could get their money back on, if you will, because it was a public health claim and it gave them a market edge. And we've been criticized over the years for allowing trials that are just on symptoms and not on the outcomes, or on a surrogate, to allow a drug to be approved, when the big outcome is, does it allow me to live longer, or does it, you know, give me some large outcome that I would really value rather than maybe the short-term outcome? And that's a societal discussion too, because these trials are really being done

by the industry. They're not being done by NIH, because NIH themselves probably can't even afford to do the number studies.

The Women's Health Initiative, which essentially supported the removal of Prempro because of this breast cancer, that cost like $950 million to do those trials [unclear].

SJ: The VA, do they do trials? Because they're looking to save money.

A1: Oh, no, they . . .

SJ: Are they countering the trials of industry?

A1: I don't think they're doing -- well, they don't count . . . I don't think the industry outcome trials have been shown to be any less reliable than anything else.

SJ: No, but they can do the comparative trials.

A1: Comparative trials are a different question, although more and more they're overlapping.

Let's just take a trial against placebo to see if you can reduce heart attack rates. Those used to be done by the VA and NIH and stuff. Now they're increasingly, overwhelmingly, being done by industry. The VA is still doing some trials.

SJ: Because you mentioned earlier that they were much more on top of trials than we were [unclear].

A1: The VA did the first two trials ever that showed lowering blood pressure was really good for you, reduced heart attacks and strokes. And it was a complete landmark. They were wonderful trials. They changed everything.

And they've done other trials, other very important trials since. I don't keep track of how many they do, but they still do. It keeps the doctors interested and they like to do it, and they're very credible.

A2: You know, NIH does a lot of trials on behalf of companies. The Celebrex trials that essentially were behind the confirmation of whether Celebrex was implicated in the cardiovasc -- one was being done by the NIH and one was being done by Pfizer. All the Cancer Institute trials essentially -- I don't know who faults the IND, but companies are running their drugs through the cooperative programs of the Cancer Institute. So it's not that the . . .

SJ: So they pay for it and NIH runs it?

A2: Oh, yeah. Usually they [unclear] and the NIH runs it. They all have a . . . A 40,000-patient hypertension trial the NHLBI did added an arm for [unclear]. Pfizer wanted them to and Pfizer paid for it, paid for some part.

SJ: Talk to me a little about -- the literature that we have access on the AIDS era is just ghastly. I mean, it's as bad as your early drug-trial stuff.

A1: Really? I would think it would be pretty good because that's when we really started to open up.

SJ: Well, it may be in the technical literature, but what's gotten translated down to the stuff that historians would pick up first is pretty pathetic, so nobody's translated some of these. I mean, sure, we've got *Federal Register* documents [unclear].

A1: Who did you talk to?

SJ: Well, I'm just going back a couple of days.

But what I can't get straight is the development of, there were . . . Once again, like [unclear] there were ideas for -- well, the parallel track is different -- but for ways to allow . . . After laetrile, we had to start talking about ways to have patients . . .

A1: [unclear]. Nothing's ever been done under parallel track. It was a stupid idea that Tony Fauci had, and we went along with it because how could you not, but we've never done anything with it.

A2: It was already covered by accelerated approval. It didn't add anything.

SJ: But wouldn't we have been accelerating approvals long before it was labeled?

A1: We developed an antiviral division, breaking it off from the other divisions, in about -- you have to check -- something like 1982, before we even knew what AIDS was. Okay?

A2: No, no, it wasn't, no, no.

A1: It was '84.

A2: Actually, the reason why I'm asking you who did this is because myself and Larry Bowman went over to NIH when the AZT placebo control trial was finishing up. We had heard that there was this trial that NIH was doing with Burroughs Wellcome and that there were nine deaths on placebo and no deaths on AZT. And this was a time when we had gone through a merger of Biologics and Drugs. We were merged and unmerged. This was a time where Hank Myer was a Division Director. There was the Center Director, and Paul Hartman was his deputy, two guys that came from the Biologics side of the house.

And these guys, interestingly enough, because they came from the Biologics side, were really not clinical trials, because Biologics never did a clinical trial outside of a vaccine in their life, and neither did the AIDS community, because they're all infectious-disease

docs. What they did most of the time were laboratory studies, small studies.

So, where I'm going on this is I'm trying to give you some history on this.

We were reviewing this study, or were asked to review it, knowing that there was no turning back if folks didn't get this right. So we went over on a Thursday, Thursday afternoon in Tony Fauci's shop, and all of maybe 10 people in the room. There was Norm Cooper, who was then the team leader, and then [unclear] that subsequently became a division. So, at that point in time . . .

A1: Is that right? [unclear] wasn't a division?

A2: It wasn't a division. It was just, it was then that they decided they'd better do something.

A1: You're really sure?

A2: No, I'm positive, I'm positive.

So, there was -- so we went over, and eventually, in a matter of four days, we were talking to the folks there, and essentially, what do you know about full ascertainment of the outcomes for everybody in this trial? And they were pretty loosey-goosey. They didn't really have their facts and whatnot. But they knew enough to call us over to say that if we get this one wrong, you know, with the AIDS crisis being what it was at that time anyway, we asked them

to do certain things, go back and call people up. And on Saturday, they had 19 deaths in one day on AZT. In a matter of like three days, they had doubled the number, and sent us the data over the weekend. We analyzed the data and they closed the trial down on Monday, Monday or Tuesday, made the announcement, and that was it. And we were faced with that . . .

At that point in time, the decision was, well, if we're going to approve this, how do we make this available to folks before we approve it? And a lot of discussion was going on then. And then there was an advisory committee that was set up, and Kessler was here at the time. This was subsequent, as we got down into the protease inhibitors and the next level of AZ . . .

But not too far after that time, things started to mushroom, and they realized they had to create a new division, and Ellen became the Division Director, and then they started to hire some more people. And even at that time, the NIH, under Fauci, they brought in Dan, whatever his name was . . .

A1: Huff?

A2: Yeah. And I went over to talk to him. I said, "We did all the work for you on this AZT. We can't do this anymore." This was their drug.

But it is important. I mean, the first AZT approval, in terms of evidence -- I mean, Bob was describing how they got the evidence -- was overwhelming. Sixteen deaths to one? That's not a hard call. So it was only one study, but it was overwhelming.

A lot of people think the whole treatment IND idea came from AIDS, and that is . . .

SJ: I knew that wasn't.

A2: It's not true. It was proposed in 1982 -- not that AIDS didn't help.

SJ: Do you think it was . . . It's so remote from laetrile. Did it either -- was there any connection with laetrile?

A1: No. Why would there be a connection with laetrile?

SJ: Because that's the first time the agency was confronted with the idea of life-threatening illness as being a criteria for either accelerating or looking at things differently. I mean, there's no doubt that laetrile didn't work, but . . .

A1: Laetrile was a fraud, and the final nail in its coffin was the study that the NIH did.

SJ: Exactly.

A2: What I was trying to fill in here was my recollection of when we opened up the approval based upon a surrogate marker, because then we were getting into viral load, CD4 counts, and there was a big discussion. Kessler called [unclear] at that point in time, and I think was probably heavily influenced by Carl. But it was at that point that that philosophy in those regs were sort of developed.

And at that point in time, the key players that are still around that you might even want to talk to, David Fivel came in as the Division Director, still around. I think he's out on the West Coast. He's probably worth talking to.

SJ: He gave me a couple of names.

A1: [unclear].

A2: But he was an investigator during the first community-based that aerosolized [unclear] was approved on. He was principal investigator for that particular study. And after that, he came and we hired him, and he became the Division Director. But he could probably fill you in from that point on.

But during the period between what happened with, you know, the Antiviral Division being carved out of the Anti-Infective Division, it was just a small team at one time,

and then, boom, the whole area of antiviral drug therapy and whatnot became a whole area unto its own.

But that was probably the most proactive area of any area that we have seen since -- huge community activism in clinical trials; a whole development of a cooperative program that the NIH had put in place to do clinical trials that mimicked the cancer cooperative agreement. That's when Dan Hope and that crowd came over from Cancer. They put the cancer model into play. So you had this whole cooperative agreement program. And the companies were running a lot of their trials through NIH in this cooperative agreement, so a lot of the stuff that we were seeing were essentially run through the cooperative agreement programs, and there were big meetings and whatnot that were held all day, and we were invited to their open part of their data-monitoring committee meetings and whatnot -- a lot of formative stuff that went on that the AIDS crisis produced: new regulations, new practices, very early feedback on protocols because time was running and whatnot, the fact that there needed to be communication between the NIH and the FDA on these protocols and whatnot.

So, that was an interesting period of time in terms of, we influenced, without the FDA -- this is another example -- we influenced and brought the AIDS community

investigators up to snuff in about a year. They were a very naïve, amateur crowd, nobody in the game, no investigators that knew how to do a trial, and the advocacy groups too. [unclear] but implacably hostile to us and eventually became the strongest supporters.

A2: Because, I mean, they'd come up with the evidence if they'd follow what we were doing there.

And so we influenced the quality of those clinical trials in a matter of a year. If we weren't in place and the FDA wasn't around as we know it today, God knows where that would have gone. They probably might have very well followed the early oncology models, where there were no control groups and a lot of combination-drug stuff and throw the seeds to the wind, and that was disaster city. We learned nothing for about 15 years in oncology.

A1: And it's important because this is a highly toxic class of drugs. You really don't want to be messing around with somebody's health. You really need to know. That's true of oncology too.

SJ: Well, that's my question about combination drugs -- and this is total ignorance on my part, so you can laugh if you want, but . . .

A1: Combination drugs or [unclear] drugs?

SJ: Well, I'm just saying, what I'm trying to figure out is the protease inhibitors, they found at some point that the combination was better. Okay? But we had to contend with the fact that we don't like combination . . . Well, I could say [unclear]. In other words, we had to approve the drugs first as individual drugs. Am I correct?

A1: Yes.

SJ: And then we combined them and they showed a greater efficacy in combination.

A1: [unclear] have special rules on fixed combinations. If you want to market two drugs, gather in a fixed combination so people can't titrate them or do anything. We have fairly stiff requirements for what you have to show. It's sort of obvious what you have to show. But drugs are used in combination all the time. All kinds of drugs are added to the available therapies. So in heart failure, the drugs aren't fixed combinations, but everybody starts on a diuretic, they may or may not get digoxin, then you get an ACE inhibitor, then you get a beta blocker, then you get . . . They're not fixed combinations, but that's the way [unclear] evolved.

SJ: Oh, so the practice of medicine.

A1: In AIDS, much of that was the same. You couldn't leave people off of the standard therapy, so what you were

testing to see was whether adding something delayed development of resistance or gave you a better [unclear] or lowered the count more or something like that.

There are, in addition, fixed combinations.

SJ: But those are recent, very recent.

Al: Relatively recent. They're as much to facilitate use in Third World countries as anything. But those are cases where the standards, even for combination drugs, goes back even to 1970, and a lot of our early battles even were in the '70s on companies wanting to put a fixed combination together and not being able to satisfy the standard of essentially showing that the combination was better than each of its ingredients, because that was the way the interpretation of adding value, if you will, for . . .

But a lot of that started with DNAS&RC. That invented the term "effective but" to mean, at least in one case, ineffective as a fixed combination. "Effective but" in many cases meant yes, it works, but you haven't shown that each component contributes [unclear]. So a lot of the [unclear] drugs that got removed were combination drugs.

SJ: And Pinalbo was the case that went to the courts, and it was a fixed combination.

Al: That's correct.

SJ: And that's what Scott Podolsky is actually concentrating on. He did a lot of work on the early testing of pneumonia drugs, and so now he's moved on to that kind of, he's nailed it, but Pinalbo was the key.

A1: But also, [unclear] people combine sedatives with everything. All of those drugs are gone, all of them. They're for GI disease, they're for [unclear]. They used to put phenobarb with your nitroglycerine [unclear]. And people combine anticholinergic drugs with sedatives all the time or with Librax. Librax is actually still there, to our disgrace. But there were dozens of such drugs, all of which went away because they failed rigorous tests. They were actually tested properly and they all failed. My guess is $100 to $200 million of studies trying to show that irritable bowel disease, which is something everybody has, you know, would respond to this combination, and 100 percent of the trials failed. [unclear].

A2: Yeah. We probably won court proceedings on those. Those went to administrative hearings.

A1: Some of them did, yeah.

SJ: Tell me a little bit about the transitions you saw in the '80s. The perception is AIDS drugs were driving everything. I'm sure it wasn't true. But the '80s strikes me as the period when we really start having some more-

effective drugs filtering down to everyday practice. I'm sure there were before that for the major diseases, the heart disease -- I don't know about cancer.

A1: Well, in the '80s in cardiovascular, you start to see outcome studies. You see the post-infarction beta blocker trials -- those were in the '80s -- and there were more . . . I can see some [unclear] early trials maybe; I'm not sure. [unclear] be the late '80s. So people are getting the idea that you can actually do this thing.

Some of those were done by nations. I mean, the first [unclear] trial was as successful as the [unclear] studies done by Scandinavian governments, I think. And there was one done in Scotland; that was another government trial.

But then there began to be trials that were done, paid for by drug companies. I think [unclear] and those things were drug [unclear].

A2: Well, Bob's right, because at that point in time, to do these large trials, which took -- usually they were virtually all for, some of them might have been prevention, but I think they were all for treatment, you need a data safety monitoring board. And there were a lot of fuzzy practices going on, both, there were industry doing these studies, and then you had the NIH doing these trials. And the NIH generally, as a practice, had what was called the

Data Safety Monitoring Board, and they were there essentially to protect the patient, but also to stop the trial early if something was going in a funny way, either for or against you.

SJ: How does that differ from the IRB?

A2: IRB has nothing to do with it; they don't even do that. All they do is they look at the protocol in the beginning, they bless it, and they're gone.

SJ: Okay.

A2: They don't see results.

SJ: That's true, that's true.

A2: The people who are doing the heavy lifting are the people who are actually looking at the results, and they've got to do it in a very circumscribed, protected way. They're in a room like this. They're in some room, they're not supposed to be talking to anybody else because all this is supposed to be done blindly.

On the other hand, during the '80s, there was a practice going on in the industry that I could look up the data all the time, and it took us about eight or nine years to figure that out, and finally we said "time out, that's a no-no." But we put good practices and procedures in place essentially to say, don't ever do this again, because if

you do, then you're going to [unclear] ability to [unclear] trial.

A1: Right. But for a long time, we had no idea.

SJ: And the blinding, yeah.

A2: And we also didn't know there was a problem.

Well, I remember, this was the first of drugs called H2 blockers, was cimetidine, the first drug that ever succeeded in shortening the time for which you had an ulcer. That was in the '80s; I don't remember exactly when. And I know what they did in their trials. They had trials that went two weeks forwards in six weeks, and as a patient came in that's either cured or not cured, they reran a P value. As soon as it got below 5, below .05, they stopped the trial. That's how they did it.

I mean, we would never dream of [unclear], but we had no idea.

A1: At that time [unclear] the statistical methodology that was developed by a lot of people, primarily from the NIH, that essentially said that's a really bad practice.

A2: But we didn't tell them that; we didn't tell anybody.

A1: We probably . . .

A1: But until we started to pay attention, we didn't say a word about this. We never told them that the way they did those cimetidine trials was wrong.

A2: Well, we may not even have known. I don't think we knew.

SJ: Cimetidine is not an antibiotic.

A2: No. That was the first [unclear] stomach. It was an ulcer medicine.

SJ: Blocked the receptors?

A2: It's Smith Kline and French's product, special [unclear].

SJ: Well, we now treat ulcers with antibiotics. Right?

A1: Oh, that, yes. We've now discovered that *H. pylori* ulcers are an infection.

SJ: And we really do think that's right now.

A1: Oh, yeah, it's overwhelming.

But these drugs that lower acid, they work too.

SJ: Right.

A1: And you actually use them with the antibiotic to kill the ulcer and keep it from coming back. And now we use proton pump inhibitors, you know, [unclear], to treat

ulcers too. They're more powerful antacid-type drugs than the H2 blockers. But the H2 blockers were very, very powerful: cimetidine, ranitidine.

SJ: Now, how many -- do we have a lot of practicing physicians now? You were talking about that in the early period, that you [unclear].

A2: No. We have [unclear] people. It's a very impressive group of clinical [unclear]. We have turnover and whatnot, but I think . . .

A1: [unclear] practice, though.

A2: Yeah, but, I mean, they're little. What I mean . . .

A1: We'd all treat ourselves.

SJ: I wouldn't even want to be around you guys.

A2: [unclear] there are some clinicians that want to do it their way. A lot of them -- and vote with their heart as opposed to what was the data.

A1: Actually, they're totally intimidated by the biostatisticians.

A2: But we have a very good group.

SJ: I understand nobody can understand the statisticians, so I don't need to feel badly about them.

A1: They're very intimidating. They throw numbers around, you know; they're impossible.

But a lot of the clinicians that come in have a master's in public health, they have a background in study design. The problem is holding onto them so that they stay here long enough to have an impact. There's a fair amount of turnover in that kind of crowd.

But we're on another planet in 2008 versus where we were in 1971. It's apples to oranges.

I had a guy who was the second-in-charge of the cancer division tell me that the well-controlled study was one where everybody was [unclear]. He just had the idea that [unclear] a control group.

SJ: And when was that?

A1: In '73, [unclear]. I mean, really, you can't believe what we didn't know.

A2: Well, right now we have a system where you electronically file your reviews. There was a point in time where only the clinical review went into some file, and the statistical document could go into your drawer if you didn't want [unclear].

SJ: [unclear].

A2: Three, four, or five years until we figured out. But those were the days where you had five copies, I mean, salmon copy, green copy, yellow copy -- actually, you had

to type these things out. I don't know how the secretaries did it. The secretaries typed them out and took them . . .

A1: It took them six months to a year.

A2: Yeah, it was just horrible.

A1: Six to eight months of the delay in getting reviews done, you know, we used to have 34 months, 36 months, was getting reviews typed. The reviewers wrote them out; some of them knew how to type. And we at various times had either a typing pool or an assigned typist, and it would take six to eight months. And then, if it was in my division, I would mark it all up and tell them they had to change it, and then they'd have to type it again.

A2: But you've got to realize, this was pre-word processor . . .

A1: Yeah, and pre training people to type. Does he know how to type? No. Do I know how to type? No.

A2: Actually, I had [unclear].

SJ: But your staff does.

A1: All the reviewers know how to type.

A2: But from a statistical perspective, our only computational capability was to run downstairs to the second floor of the Parklawn Building, where all the mainframe computers were, and you would type up the program on these 80-column cards and see if it would run, and maybe

come back four hours later and find out it ran or not. I mean, the current crowd doesn't understand how good they have it in terms of the PC and the electronic world we're in and the facility, the ability to be able to do what you want to do in a more efficient way. But we've still got many places to go to improve the way this place operates. We have many -- we're dealing with guidances right now that are trying to push the field even though the field's been dealing with this for 15 years and probably not getting it right.

SJ: Now, the guidance documents are the ones that you guys started putting out that specify . . . Well, I'm just thinking of oral contraceptives that specify what you consider the minimum number of patients and [unclear].

A1: Now, these are -- the ones we're talking about are not medical-class specific. These are methodological-principle specific.

I would view this as taking the 1970 Adequte and Well Controlled Trials principles and expanding on what they meant 30 years later in terms of actually . . .

A2: [unclear] like the C4. Some of the ICH guidances are general guidances, not specific to a class. There's one on dose response, there's one on statistical principles, one on [unclear]. We're now getting into some

of the nitty-gritty, trying to tell people how to deal with multiplicity, how to do adaptive science, and a little more than the ICH E10 document did about how [unclear] study. That document is terrific in many ways, but it [unclear].

A1: This is another interesting wrinkle which you probably . . . You know, the FDA is such a lucky organization in terms of the resources it has available to it. You look outside the United States, and there isn't anything that even approaches the resources and the brainpower that the FDA throws at the problem.

Europe can't touch the United States in terms of the intellectual brainpower thrown at the regulatory process, for the most part because people don't do it for a living. You know, unless you do it for a living, you can't do it as a consultant. I mean, it's not like having an advisory committee member come in and spend one day every three months on understanding your problem, or having somebody come in and help you make a decision. Unless you sort of live and breathe the thing, you don't understand the depth of the problem. Europe doesn't have that talent base, nor does Japan. And we found this out big-time [unclear] 1990, when the ICH came about. We were trying to get common standards across the world in terms of pharmaceutical regulation.

So the other hidden benefit to the FDA is we calibrated the rest of the world, despite what other people, I think say. There was a strong feeling way back in the '80s that, why can't you do it like the U.K. and be as fast as the U.K.? Well, that was totally misunderstanding what was needed to do the job in the first place.

But I think what's interesting about modern drug development right now is that it's global, it's not domestic. We're finding out [unclear].

SJ: You mean the clinical trials, a lot of them are [unclear].

A1: Well, not only [unclear]. It's a common database that everybody [unclear].

A2: It's a common . . . You know, most clinical trials that we're seeing, in fact, 60 to 65 percent of them are done outside this country. But they have to be done to our standards. So the ICH framework sort of raises the bar for what those standards need to be.

But we're talking about when we get them in-house and the level of detail that we go through to evaluate them and deconstruct, well, we ask for the raw data. We want to look at patient-level data. Nobody else does that. And you can't get anybody who doesn't do that for a living to

understand that, because people used to say, even during the little Lasagna days, "Why can't you guys rely on the medical journals? They review this; they review the articles; they peer-review it. Why isn't that good enough?"

SJ: Because they're not peers.

A2: What do you add value to? People just didn't understand. Well, first of all, they don't see the data, they don't see what's changed. If you write the article up, they only tell you what they want. And there's a lot of, in criticism, these days, of the journal peer-review process, and rightfully so. Those are highly deserved. But the value added by the infrastructure of an FDA just can't be overstated. It's just . . . And we're lucky. I mean, the fact that Congress has -- we've always struggled for funds. But the fact that we got what we got right now in terms of the talent base in the FDA is unbelievable.

Peer review is a total crapshoot. It depends on who you get. And there are so many examples of . . .

SJ: Well, they have to define "peer" better, you know.

A2: Oh, I think the people are respectable, but they're all busy, they're only given a couple of weeks. The journals really pester you if you take more than two

weeks. And I've seen peer comments on my own work and other people's work, and it varies from very wise and scholarly to moronic. I mean, we sent a paper to *JAMA* on how to improve non-inferiority studies. All three comments were idiotic, one of them barely in English.

SJ: Define for me a non-inferiority study.

A2: Oh, that's [unclear].

SJ: I've heard the term, but it's not ringing any bells.

A2: That's a trial where you're not, where you do a study comparing two drugs. You're not [unclear] show that one drug is better than the other. You're trying to show that the one you're testing, the new one, is close enough so that you can assume that if the other one works, it works. How to do that is a major challenge, major, major challenge. There's an ICH document that touches on it to begin with, but we're pinning it down in a way that no one's ever done before [unclear].

SJ: Well, as I understood it, we're going to hire 600 new drug reviewers in the next God knows how long.

A2: A lot, a lot.

SJ: Okay. What are you going to do with that many people?

A2: Well, that's an interesting question. I don't think you can quote us on it because I'm not sure . . .

SJ: This is just for me.

A2: Well, we're running into our limits, for example, on being able to meet with companies, and yet with think that's a very good investment of time and money to help them get it right. So we like to be able to do it all we can.

We also want to be able to do the full-bore analysis of all the safety data and look at it. We definitely are going to hire a bunch of people -- I don't know how much of the 600 -- to make sure we monitor any developing safety issues better than we do now. The idea that we don't care about them is just completely wrong, but we're not on top of them as much as we are on an application, and we're going to try to bring that closer.

We desperately need more people who are competent in epidemiology and things like that. They have a small number of people [unclear] not nearly enough.

So there are plenty of needs.

I don't, we don't, I think we're thinking more like 300 within CDER.

SJ: Do you have enough room in the new building for that many?

A2: Oh, well, there's never enough room. We'll have to plan for growth.

And then I'm sure Bob wants some more people, if we can get them, who are . . .

A1: Part of it is just keeping up with the wheel. I mean, our staff has been [unclear] just to sort of review the applications and [unclear] for a long time. And part of this is, what does the modern review process take in terms of it being very much a data-based, evidence-based document, not a clinical judgment-based document? So what I've argued is, you know, what might have played out back awhile: What does it take to do an NDA review? What's the skill mix and the mix of people? Right now, up until we're probably 5:1, 4:1 medical officers to statisticians. It's just out of whack in terms of what's needed [unclear].

SJ: You need more statisticians.

A1: Yeah. So they're equal in total IQ, but there's not enough of them around.

SJ: Whatever.

A1: But, like you say, really, to allow time for think time and to sort of look at these things that have gone on for 10 and 15 years and try to bring them to closure based upon what we know and make the world better off for it, I mean, there's probably five to 10 of these

things hanging around that we need to devote time to clean up and build consensus so people don't waste their time.

A2: And you need enough people so that you can try to think a little outside the box.

For example, one of my -- I can tell you, we've already had some sort of rounds on this. I think we lose information. I think we're over-rigid in the way we approach things, and certain information about time-to-event and stuff like that, I think we overdo it. But I'm not going to convince anybody just by saying that. We've got to have some discussions of this stuff. And there's a lot of talk about it.

I mean, if it were true that we're not getting as much information out of trials as we could -- we're learning whether they show the drug works; I'm not worried about that part. But there's other information of a more descriptive nature. We've got to come to grips with that [unclear]. If you make it too hard, you just lose it; if you make it too easy, you get a go.

SJ: Are you talking about having the statisticians look at it or having the [unclear] look at it or some combination?

A2: [unclear]. I don't know yet. I don't know how to do it.

But, for example, my favorite example is, when we present the results of depression trials, we never ever allow people to say how soon the effect is seen, never, can't say it. That's because they don't have it as part of their initial hypothesis, and there's some subjectivity in how you decide about it. I'm not saying it's a trivial problem.

What I am saying is, we need to find a way to get that information usable because it's important to people. They have to know that. And it's not in any of our labeling now. So I have my ideas about how to do it, other people will have theirs, but we've got to come to grips with how to get the most information out of a trial once you know the drug works. And there's many examples.

A1: Well, I think . . .

A2: Many examples about how exactly to do it, how to protect against making mistakes, how to not get fooled. That's going to be plenty of trouble.

A1: What Bob is talking about is leaving enough airtime in the review process that you can have very broad discussions that come to closure on how you do this, and then make sure that you can oversight that within 17 medical-review divisions. So it impacts the structure, the process, how you operate. If you're in a "I've got to get

it out the door every Friday" kind of, that's stuff that gets done. And so you can use up that, those FTE's real fast if you know where the real problem is, if you understand where the problem is.

A2: Yeah. I mean, once you understand something, the complexities are magnified. I mean, you can say you're just making work for yourself because you want a lot of people, but . . .

SJ: They [unclear].

A2: [unclear] important problems -- important, interesting, they matter to how you communicate. So, the more you think about it, the more there is to do.

A1: And, for better or worse, we lead the world on this stuff. It's not that there aren't smart people all over the place in institutions, but they're not organized to do this. They comment, they rise up, write a paper, but we're the only ones who are in it day by day. So, really, I mean, I think that's the story sometimes in '62, gradually accumulating and growing, and that's [unclear] I don't think people understand. Maybe [unclear].

SJ: Well, I'm going to try, I'm going to try, but . . .

A2: Is Harperter trying to get that message?

A1: He's trying to come to grips with, well, his [unclear] he's a political scientist.

SJ: We haven't written enough to get them fodder.

A1: So he's trying to figure out how we got so much authority, and he's a little worried that we've had a little damage to our authority in the last few years, and he's right.

SJ: But there's a difference between authority and reputation. He's using that fuzzy reputation definition. What we're talking about is authority, and that's where John and I need to do more writing.

A1: Okay, but you're right. He thinks reputation is the key to authority. I mean, I don't have a view on all that. And he thinks our reputation is in some degree of trouble. But I think a lot of our reputation comes from the fact that we've been the people leading the world on stuff.

A3: I must be in the wrong place.

A1: No, no, you might be the right place.

A3: No. This is the history of the FDA?

SJ: Yes.

A3: No. I have another meeting at this time. I had to decline it. But there's another few years here. I think [unclear].

Al: There's not enough gray hair.

SJ: At least he's in the right place.

Al: Anyway, I mean, I think a lot of the -- it's reputation, that's where.

I think a lot of that comes because, you know, we've been doing this in a highly responsible, credible way. We're open-minded, we talk to people, and we apply resources to it that are vast compared to any other single party. So I think the story of the '62 amendments and how we filled them out is probably where the reputation . . . And he thinks that too, but [unclear].

SJ: Well, that's what I'm saying. He needs to publish some articles before he puts the book out because he's developed a few of the ideas, but he doesn't have a lot of varied kinds of documentation. So if you would take one or two stories and explain them perfectly, then some other things would fall into place.

Al: You know what? My concern is that the value to you doing this is we don't have anybody in the FDA with a real long view of the world that actually publicizes. We are our own worst problem. The FDA is a long haul. We have not generally publicized ourselves. We don't think that's part of the mission.

SJ: I'm being told I can't do that, so I'm not supposed to slant a story. I'm working on that.

A2: Well, the problem is the whole reason we're in the problems we're in. You know, we're lucky that we have enough people that take it on the road that the outside folks know what the hell goes on. Bob and myself and probably another 30 or 40 people in the place give the place the reputation. It's not that some of the other folks couldn't, but the FDA and the system doesn't promote those people to go out there and do something. We at the Commissioner's office, it's a disaster.

I sat on the Science Council for about eight years as the CDER rep, along with all the other Centers, and there was one chronic moaning and groaning all over the place, that the FDA doesn't promote our science in any sensible way, so the folks don't, they're not first thought of when somebody says science. When you get a bunch of outsiders like Von Aschenbach brought in to do this Science Board thing, these people are thinking, "Show me your laboratories." They're clueless, absolutely clueless about what is really going on here. They don't understand what has been the systemic changes that have influenced the world, if you will, over the last 10 or 15 or 20 years; the kinds of stuff that we're saying that, but for the FDA, the

modern clinical trial wouldn't be anything close to what it is right now.

[recorder turned off and on]

She's got her pulse on the interplay between evidence and data and . . .

SJ: And she does have a long view.

A2: And other folks don't.

But, also, she probably now recognized that we've got to promote her. And she's been the best I've ever seen since I've been here in and sort of trying to capture in the critical path document, for better or for worse, was a brilliant idea. It sort of said, this was FDA's idea, and everybody sort of got brought in. Now we're trying to do something on the safety first or whatever.

But the FDA generally isn't a proactive agency, so we give some talks and whatnot, so the FDA gets known as whatever. But, you know, it's only those folks who are close enough to the fire that actually see this. We've got a couple of folks that have stayed . . .

A1: The industry recognizes this.

A2: Oh, the industry does, but they don't . . .

A1: But nobody believes them.

A2: They're fair-weather friends, you know. They'll come to our defense. But, I mean, we've got a couple of

advisory committee members who know exactly what we're talking about because they've had longstanding associations, and they're very effective -- the Ralph d'Agostinos and the Tom Flemings, and we've got some clinical people on that side. Some of them turn, like the Nissens, you know, from, might have been useful to a self-promoter or whatever.

But when you look at how is the FDA perceived versus how is it projecting itself to be perceived, we don't have a clue. We don't have any game plan.

SJ: I can tell you, when I was talking about the AIDS drugs, I'm not kidding. The only substantial book comes from Quitney, who was -- I'm so sorry, it's one of the chief AIDS activists whose name is escaping me at the moment, which is quite embarrassing, actually. But he's talking, he was an early drug runner for isopromazine, and there's another one that was used in combination. Anyway, he was one of the ones trafficking in Tijuana every weekend to try to get them for his buyers' group, but the word was that he was [unclear] at the time he started.

Al: A reputable [unclear]?

SJ: No, it wasn't her.

But anyway, but, I mean, his comments about FDA are rather -- I mean, the overall impression is very scathing,

and I know that the reality is very different, but it's done in such explicit detail. And Ellen Cooper is portrayed as, well, there's just no way to describe exactly how she's portrayed in all those.

But I'm just saying, the Hills book, Hills' book is nothing but what we've already known.

I guess people, it's gotten such a good reception because nobody else knew about it.

A1: Well, it's not hostile.

SJ: Well, that, and the fact that they didn't know as much as he provided, I mean, and he provided something . . .

A1: I didn't read the book. I've only read four pages of it [unclear].

SJ: [unclear]. But my point is, I guess it's nothing that John and I didn't already know, so I'm sure there are other people [unclear] great scholarly piece.

A1: You sort of said the only information that is available about certain periods of time is information that . . .

SJ: From very slanted sources.

A1: From very high levels that were clueless about really what was going on, so you never really know . . .

SJ: But it could be outsiders who have very particular points of view, and there's nothing wrong with what they're saying, but it has to be counterbalanced.

A1: Well, and most of them never talk to us.

SJ: No.

A1: I talked to Hills, but most of these people . . . [unclear] ever talk to anybody, you could sort of talk to people who spent their whole life in the Commissioner's office, and they would be clueless about what goes on in the Centers. That's a whole different culture in its own right.

SJ: Yeah. The Commissioner's office culture is . . .

A1: So if they're all reflective about they don't have [unclear]. If they're reflective of what the outside thinks is going on in the FDA, we're in deep shit. I mean, so, you know . . .

A2: I mean, even the view that there's a lot of decisions that are in some sense political, which everybody believes, I mean, except for Plan B, I can't think of anything that was political.

SJ: Some of the AIDS, response to AIDS, Frank Young, whatever you think about Frank Young, you know, rescued a lot of reputation. [unclear] even do anything because his ability to interface with that group was particularly good.

A1: I'm not saying we don't respond to groups yelling at us, but we retain considerable rigor in AIDS in the face of a great sense of urgency, which we all shared. I mean, who didn't have it? Even the use of the surrogates was very rational, very sensible, and a surrogate is not exactly radical.

SJ: But that . . . Okay, wait a minute. That's something I need to asterisk because I'm ignorant. Was, really, AIDS the first time we looked at surrogates officially?

A1: No. We'd approved, we'd been approving drugs for lowering blood pressure since the beginning of time.

SJ: Blood pressure drugs.

A1: Not having had the conversation about the linkage part of it. We never talked about it. We just considered it self-evident.

SJ: But that had been shown pretty early, though, right?

A1: In 1967.

SJ: So outcome data was available on that, so it's not really a true surrogate, or it wasn't at least for a couple years.

A2: Well, first of all, we have lots of drugs that do something. The next drug that comes along, it may have to

prove the thing over again clinically. You don't get necessarily credit for surrogates just because you . . . I mean, there are lots of drugs that are antibiotics. Okay? They kill the bug, and then they're shown to work. Do we approve antibiotics because they kill bugs? We do not. We approve them because they show that they improve something. So even where there's a very plausible surrogate, we still ask for data in most cases.

But there, there, you either show that you've affected the thing that is very clearly the fundamental cause of the disease, or much of the disease that [unclear], or you make the virus go away. Well, that's a pretty plausible surrogate: no virus, no disease.

So we thought it through and we did all the analyses and did the relationships and made a rational choice. But it wasn't a radical departure from what we'd done. If we'd looked, we could have found the same evidence for blood pressure, which we've now gone and done, so we're going to change the labeling for all blood pressure [unclear].

SJ: Oh, very good.

Okay, just a quick one to finish up.

What's your favorite drug or favorite NDA that you dealt with? You don't have just one that sticks out, for whatever reason?

A1: No.

SJ: Or is that too weird a question for someone who's . . . Okay, that was too naïve.

A1: I've been involved in lots of different things. I believe that the, I think . . .

SJ: The one that taught you the most.

A1: Well, I think the COX-2's have taught me a lot. I sort of knew the answer even though, earlier than it came out.

But I guess one of the first NDA's that I cut my teeth on when I came here is a drug called Cromolyn, and it's still approved. But it was the first time that probably anybody made a presentation to an advisory committee. It was back in '72. Advisory committees just came into being. And this was the Pulmonary Advisory Committee. And this was a drug that had two studies. One didn't make it, and this was sort of another drug that another study that wouldn't make it [unclear]. And this was a large crossover study, a large crossover study done, about 300 patients, done in several sites.

Anyway, I had the company who had done this study was Syntex, no longer in existence, but they were a West Coast company that did a lot of oral contraceptives also. It was a [unclear]-Syntex sort of combo, a company deal.

Anyway, that was probably the first NDA where I cut my teeth on actually asking the company for the data tape and getting the raw data and reanalyzing the whole study, because the company didn't do it right; essentially presenting that to an advisory committee, and essentially having the advisory committee vote the drug up, because it would have voted the drug down on the basis of what the company did.

So, in my mind, that was the first example I had of the FDA being proactive, helped the company out when they actually had the evidence but they were screwing it up themselves. And I think that we do a lot of that in general as part of what we do for a living, having these meetings and whatnot, because I just think that's . . .

And to a certain extent, you have to, every reviewer has to have examples like that. Otherwise, they don't get it. They have to go to some aspect of that experience. Otherwise, they don't understand the culture, they don't understand what's needed, they don't understand how the data might be collected, they don't understand a lot of aspects of it.

SJ: Yeah. Bruce Stidell told me something similar. He had to learn the hard way about -- he came from the NIH environment, and he liked observational studies a lot, and

he had to be convinced how they could fool you before he could quite accept it and became a convert.

Okay, Bob, one of yours.

A2: [unclear] does make an important point. We're not here to just destroy them, and it's easy to forget that [unclear] . . .

Interview with Bob Temple & Robert O'Neill #2

TAPE 1, SIDE A

A2: . . . computer application that taught me a lot is when we went to [unclear] the NIH-generated data from what's called the Research Clinic Study [unclear], when we went down to -- several of us traveled to North Carolina to look at actual cases. And again, it taught you how subjective what seemed to be objective endpoints are, and it's very important to know that.

What we found was that the classification of events as either heart attacks or not heart attacks was obviously a totally random event, and we needed to know, the good people had designed the thing, they hadn't designed it very well. So I [unclear].

TAPE 2, SIDE B (unmarked tape, side 2)

SJ: And what was the other one you said when I was changing tapes?

A: The Interim Re-Infarction Trial, all written up by [unclear].

SJ: Good, okay.

Now, that's my final question. Do you guys have -- I can't go to the library and say, "Give me all FDA papers published by anyone in CDER for the past 20 years." Do you have your own database for that kind of thing? Or if I want to trace things in the published literature, how would someone like me do that?

A1: Unfortunately, a lot of [unclear] has his publications and I have mine.

A2: The problem is that a lot of the stuff . . .

SJ: That doesn't help if you're [unclear].

A2: . . . that we're talking about are not necessarily in published literature. There are probably a lot of PowerPoint slides all over the place that are not in any great, organized catalog way. The published stuff . . .

A1: [unclear].

A2: Huh?

A1: Angie Smith keeps telling me she's working on them.

SJ: Well, I mean, of published data?

A2: Actually, there is -- talk to Jan Johansen here. And we have recently put in place some software package that goes out and scours all of the journals, and if anybody's been published [unclear].

A1: Somebody sent me that site.

SJ: It probably [unclear].

A1: And the only, the issue is, as long as the list of journals that it searches is rich enough and broad enough, but it's a pretty fascinating thing, so why don't you . . .

SJ: I could at least talk to him about publications.

A1: Talk to him about getting access to some . . .

SJ: And what was the culture for publication? Did you often have to . . . I mean, I assume that most recently, especially after the User Fee Act or whatever, you probably didn't have time to publish as much as . . . Or you're just barely getting reviews out.

A2: [unclear] for 30 years.

A1: The biostats people published . . .

A2: Yeah. We have a culture of publications . . .

SJ: Wait, the ones I can't understand.

A1: Well, that's right.

A2: The other groups, it's much, much, much, much less time. You know, you'll find publications that John has done. I've got a hundred, you know, and we can find you those. But it's not too common to have other people write articles of a general nature, not as many as you might have thought. And I don't know what's -- but maybe that will help you look.

Anything you want to find, ask April. She can . . .

SJ: Well, I can run it by names. I mean, I can run literature searches by names. The key is maybe, may well be for getting me names.

A1: [unclear] the Science Forum talk that I gave, right, which didn't have all the background stuff.

A2: Do you have that?

A1: I'll send you some, but . . .

SJ: Good.

A2: No, I think we should write a history of [unclear], and maybe an intellectual history. The trouble is, it's so easy to make mistakes and forget things. I mean, it's all [unclear].

SJ: Well, but that's what we're here to do, is help you get some things organized that will bring up things . . . I mean, if I give you a list of all the publications,

you'll recognize somebody's name, and that'll bring something to mind, and you . . .

A2: Well, the way I've tracked or organized my thoughts . . .

SJ: We have to jog your memory as best we can.

A2: The way I've tracked or organized my thinking is sort of the decade of the '70s, the decade of the '80s.

SJ: Exactly.

A2: The decade of the '90s, 2000's, because they're actually, interestingly, demarcated by major things that have sort of changed, whether it be evolution or laws or . . . But I [unclear] everything that happened in the '70s.

And one of the things you notice is that if you don't pay attention, you forget that Thomas Jefferson and [unclear] were contemporaries. I'm just thinking it out; I'm sure that's true. And I think we're good at stovepiping stuff, but I don't know if we're so good at noticing what went on, well, biostats was doing this, what was [unclear] doing.

SJ: That's what I'm having an almost impossible time doing.

A2: It's very hard.

SJ: You've distilled all the information very clearly in your own mind, you've learned the lessons, but you have

nothing to show me on how the process worked, other than . . . That's what I'm saying, the [unclear] doesn't have, is the . . .

A2: Well, we could probably reconstruct the process. I mean, I probably have a fairly good idea how, you know, and what were the events that caused organizational changes. I mean, I've been through what was called the crowd era, which Bob was through, and at that point in time, then, we were essentially merged and unmerged. There was Biologics, and we were merged for a while. And my program had to support Biologics for about four or five years. And then we were unmerged.

And during that period of time, the AIDS crisis came about. During that time, we went from a six-unit outfit to probably an eight or nine. You held one row, Jim Philsted held another. That went on for a while. And then we became a larger organization.

Janet came into the picture, created 15 review divisions, essentially decentralized the decision-making. A lot of other stuff was going on there. Generics was coming up on one side. Then you had some law changes that sort of came in and impacted the way we operated.

So, I mean, there's an overlay on the organiza -- and the process of how we did things and why we did things for

three years, when, if you understand that, you know, you had a secretary typing a six-page review that had a piece of carbon in there and that, you know . . .

SJ: It takes a while.

A2: It takes a while.

SJ: You can't make errors because you have to change it on the top sheet.

A2: And not only that, then you had what the level of analysis was then. The level of analysis might have been the eyeball test versus "I'll redo it." So you've got to understand the fact that where we are right now, the environment we live in and the tools we have available to us are totally different.

A1: I mean, we used to like to say, "Oh no, we don't necessarily redo everybody's calculation." Well, that's because it was so hard to do. You had data, you had blah-blah-blah. Now, there's nothing to it. We just do it all, they just do it all the time.

A2: If we can get access to data. But, I mean, our problems are bigger now.

Right now, let's say somebody says, do a review of all the Trasylol studies or do that stuff in, let's say, the Avandia situation, 45, 50 studies, not all necessarily the same design, or you've got to get them in shape to do

whatever. You know, the nature of the problems that are being thrown at us now are much different than going across NDA's. Even companies at that time didn't even have their own act together. Companies were such that the people who worked on a drug are long gone by the time you maybe go back and ask that question two years later. So, companies are worse in terms of their historical memory than we are. We actually had a lot more staying power than companies do.

A1: But when we tried to go back and get the actual data on drugs that were being withdrawn because of, say, liver injuries so we could see if we missed something, we can't do it.

A2: They can't find anybody who was originally responsible, who archived it.

A1: [unclear] totally simple and obvious [unclear] depo.

A2: You talk about your [unclear] knowledge management. You can't do any knowledge management in the absence of having a structure for how you access past knowledge, and that's the name of the game we've been in until more recently.

SJ: I've got to go.

END OF DISCUSSION