**U:WATKINS (discussion)**

DR. WATKINS: So anyway, on that note, I will stop and I think, John, you wanted to have questions?

DR. SENIOR: I do, but I also wanted to say that Bob Temple said: “Don't talk about Temple's Corollary, which I think is grossly unfair since he proposed Hy's Law and poor Hy is deceased and can't defend himself.” (Laughter.)

DR. WATKINS: I don't think Bob's here anymore, is he?

DR. SENIOR: No, he had to go.

DR. WATKINS: So those are the issues that I think we discussed in the first session, and now we are open for comment on any aspect of this.

DR. AVIGAN: We may not have seen that effect. So we would not have ascertained it. There are antibiotics that are known to be hepatotoxic where the DILI signal wasn't seen in clinical trials, but only in the post-market.

DR. WATKINS: Right. And let me just make the comment that something very profound was said here today or maybe it was yesterday, when Bob Temple said, “They don't see drugs with maybe one exception, nefazodone, that show Hy's Law cases that also don't have increased incidence of transaminase elevations without bilirubin rise.”

DR. AVIGAN: I think what he meant, and I wish he was

here, is that those drugs that are used for a longer duration trials. I think he's talking about minimally three-month trials with chronic treatment drugs. I don't think he's talking about short-usage drugs associated with very early idiosyncratic liver injuries after initiation of treatment.

DR. SENIOR: This can't apply to everything. Obviously what's being looked at here is in the context of a controlled clinical trial where you have good reporting. It does not apply to these sporadic cases that occur after marketing, where you have no data. You don't know whether there was a Temple's Corollary group that never was recognized, never was diagnosed or not. You only know, you find the patient with jaundice. So you really can't compare the controlled clinical trial data with what you see post-marketing which is very little.

So we're going to have to cope with this in the future if we ever get around to trying to address the difficult issue of a Guidance for post-marketing. It's a whole other can of worms, as we say.

DR. VIERLING: I'm John Vierling. I'm very sensitive to our desire to be accurate while avoiding the dilemma of throwing the baby out with the bath water. But I want to amplify what Herb Bonkovsky said a moment ago about gender-specific normal range for ALT, which still need to be further investigated but are increasingly accepted by hepatologists.

If you put that slide back, I would like to call people's attention to the horizontal width of the bottom left quadrant containing normal values of ALT, and to speculate that the lowest values may be attributed disproportionately to women.

I believe that we should be gathering data regarding the change of ALT from an individual person's baseline. If ALT has an upper limit of normal of 40, then an 8-fold increase would be 320. However, a level of 320 would represent a 64-fold increase for a woman who had an baseline ALT of 5. I would submit that a fold-increase based on the individual person’s baseline would be a more accurate signal of concern. In defense of gender-specific norms, are Dr. Kim’s data that Dr. Chalasani mentioned yesterday correct, showing that the inflection point of women’s ALT as a risk 10 years later for all-cause mortality was approximately 19 to 20 units, while for men, it was closer to 30? These data support a gender-specific normal range for ALT and argue against a signal of concern based on ALT increasing 8-fold above a specific value. Gender-specific norms and analysis of an individual’s fold-increase of ALT should be considered when conducting a drug trial. As you can see, coming from the midpoint of the graph, you could easily get to the right lower quadrant represent the area of concern if your ALT value were in the mid-range, but you would not move out of the left lower quadrant if your initial ALT were in the far left of the quadrant. I think this issue needs further study,

which would be my request.

DR. WATKINS: Yes, I think it's a very good point. This is the maximum ALT, and still I'm personally a fan of fold times baseline, if the data were available, but they're not. Something else to consider is that in addition to finding a percentage of patients in terms of upper limits of normal, it would also give information that would allow you to go back retrospectively and look to see if we had looked at fold-increase over baseline ALT, would that have predicted something that just upper limits of normal do not. Yes.

DR. MAYNE: Just to support the fold-baseline, that really helps us with the other problem of comorbidities and how to deal with enrolling patients with preexisting liver disease. So it's certainly something to explore further.

I wanted to comment though, if you could go back a couple of slides to the slide in the stopping rules.

DR. WATKINS: Yes.

DR. MAYNE: I'm Jim Mayne, Pfizer. We've heard a lot of great discussion, and had a lot of great data presented on our understanding of where we get into danger zones with regard to ALT levels, but I wanted to just bring up the notion of putting ALT and AST on equal footing for stopping rules at least as it regards to DILI. Can we revisit that? I don't think we have the evidence --

DR. WATKINS: That's a very good question. Why

measure AST at all or measure, why have Hy's Law and stopping criteria based on AST? And unfortunately, only John and Mark are here to defend that.

DR. SENIOR: Well, you know, AST was first in the discovery of the value of serum ransaminase activities. ALT measurements came along just a little later, and it took a while before it became generally believed that ALT was somewhat more specific to the liver than AST. Now neither one of these enzymes is completely specific to the liver only, and it's been found recently hat ALT has mitochondrial activity as well as cytoplasmic activity; formerly they said only AST had both.

In addition, some work done recently by Ian Cotgreave and colleagues has shown isoforms of ALT that may give us a little more information on where the ALT is coming from that's in the serum. Is it coming from muscle? Is it coming from the turnover of the intestinal cells, kidney, heart, or liver? Right now we're measuring a combination of all those things, total serum ctivity. So we aren't really able to measure the origin of the ALT elevation at the moment, but we hope with further development of immunological identification of these isoforms, we may be able to do so.

DR. WATKINS: Mark, do you have any insight? No.

DR. MAYNE: Just as a follow-up, I agree with your points, John, on what we know and what we don't know about the two enzymes, but in terms of the stopping rules, in terms of

applying Hy's Law, we seem to have a good experience base, a database telling us what to do with ALT. We don't seem to have that or at least we haven't discussed it here when it comes to AST. So we may want to consider them in different ways.

DR. WATKINS: And the point is you could have a rock normal ALT and an AST 10-fold for another reason.

DR. MAYNE: Yes.

DR. WATKINS: Bilirubin elevation.

DR. MAYNE: Bilirubin elevation.

DR. WATKINS: The Guidance doesn't take that into consideration.

DR. MAYNE: Yes.

DR. WATKINS: Mark.

DR. AVIGAN: I think we can consider the serum ALT level by itself because I do agree that it's probably a better sensor with few exceptions for DILI than AST, and I can't really think of any exceptions. There are certain liver injuries which we've heard about, particularly alcohol-related, where the AST elevation is more pronounced but we wouldn’t consider that a DILI-like reaction specifically, so that would be something to think about. Again, I don't have a firm idea about this, John.

DR. WATKINS: Jay Barth from Merck.

DR. BARTH: I want to follow up to clarify what you had said, Paul, regarding the stopping rules. Were you saying

that as written in the Guidance they apply to liver phase trials, Phase III trials and applying not to earlier trials because that is a question that I have whether the stopping rules should be different for shorter treatment duration than small trials as in Phase I?

DR. WATKINS: Right. Well, that I think is correct and maybe John or Mark can clarify that but, in other words, this Guidance in general has been written from the perspective of looking at a NDA database. When you're first starting drug development to see somebody to go with an ALT over eight times, would probably be a pretty bold thing to do to continue it. But once you have experience with the drug then you might go up that high with continued treatment in Phase III. The Guidance doesn't say that but I think that's what changed the setting I think. John, will you agree?

DR. SENIOR: Yes, I think that a very reasonable thought. Obviously the guidelines pertain to consideration of approval, which is based on the NDA data. Now a company who's investigating a new agent is usually going to start out by treating normal healthy subjects, young healthy subjects, men and women. And they may want to set different criteria for, for further investigation in people who presumably have a healthy, normal liver to start with.

So these guidelines cannot deal with all possible eventualities but they I think are focused on data that are

being evaluated in consideration of approving the drug for use in the population.

DR. BARTH: Thanks for clarifying that. I would say then that it's not anticipated that the final Guidance would address earlier phase criteria. Is that right?

DR. SENIOR: I don't know.

DR. BARTH: Okay. Thanks.

DR. SENIOR: Let me ask Mark. Do you think the Guidance should be written only to apply for late-stage drug development data?

DR. AVIGAN: I agree with that because the early phases of clinical drug testing occur in a small exposure population. There you're looking for direct toxicity effects that may be dose related, or that might be an extension of a pharmacologic toxicology signal. It's a completely different question than the question at hand which has to do with idiosyncratic toxicity that may require large numbers of people. In the latter case we're not really talking about it in the early phases of clinical testing.

The way I think about this is that we're really confined to those large Phase III studies. We're talking about relatively unusual events, in the whole population. If you start with a Phase I study in 10 patients, and you start seeing funny liver test results, then what you're thinking about is quite different I think. There you might consider direct

toxicity effects which might be dose-related and which might be a reason to stop the study or to readjust the protocol, et cetera.

DR. WATKINS: Now the argument could be made that that Guidance only in late stage development is not very helpful to the pharmaceutical industry because obviously enormous commitment has already been made, and if there were guidelines for earlier development, that were science based, they would be more helpful.

DR. AVIGAN: Schematically, I think that before you go from Phase I to Phase III, you seek a green light to proceed to larger studies by ruling out direct toxicity effects. So now you've gotten that green light, you're going into larger trials and only then is it possible to see a range of idiosyncratic reactions in a larger exposure population. And with that frame, we feel relatively comfortable that by and large most of the patients with isolated transaminase elevations within a certain range will have adaptation and self-correct. The 8X upper limit of normal boundary for serum ALT levels to indicate the need to discontinue treatment is somewhat arbitrary. I think we had a discussion about the fact that it is arbitrary and that if we accept it, it is with that understanding.

DR. WATKINS: Right, that's pretty much it.

DR. REGEV: Just to comment on this same point, if we

do decide that eight times the upper limit of normal is not the appropriate level for Phase I studies, then we should decide which is the appropriate level for Phase I, and which is the appropriate level for Phase II. This may creates a logical problem because if we discontinued two patients at five times the upper limit of normal at Phase I, what do we do at Phase III ? Do we then continue to eight times the upper limit of normal?

DR. WATKINS: That's my interpretation, but as you get progressively more comfortable, and --

DR. REGEV: Well, how can we get comfortable if we stopped patients at Phase I with ALT of five times the upper limit of normal? We didn't let those patients continue going up. So we'll get probably less comfortable in Phase III.

DR. WATKINS: But presumably you have satisfactory PK. You've got more enthusiasm or you wouldn't have done Phase I. So the risk benefit ratio changes as you progress the compound I'd say.

DR. PEARS: Can I just, rather than dealing with a specific issue, just make a plea please to say as guidelines for people to think about because what we've got to be able to do is not box it into a corner, but we must deal with this stuff, and we -- for us as the pharmaceutical industry as making our own decisions to be able to apply out own criteria or what is to set a Guidance, and there are opportunities to

discuss what we do at the various points through the development program with people like the FDA, the EMEA and with external experts. So we can have a sort of -- discussion. I certainly don't want to be in a position where it will end up where this is what we must do that it doesn't allow us any room for leeway or logical thought or -- discussion. So it's not to say we don't necessarily know all these things but I don't want to be so hamstrung by too much detail. We just can't --

DR. WATKINS: Gerry, you've been standing there a long time.

DR. KENNA: Gerry Kenna, AstraZeneca. A comment about causes of bilirubin elevation. We've heard about UGT inhibition as being an important cause and that will lead to an elevation of unconjugated bilirubin. I haven't heard anybody -- well, I don't remember anybody yet talking about inhibition of the uptake and efflux transporters of hepatocytes. So if we inhibit OATP1B1, we'll get quite marked accumulation in plasma of unconjugated bilirubin. If we inhibit MRP2, we'll get marked accumulation of conjugated bilirubin. So it's possible to get quite big changes in bilirubin without overt liver injury, and I think that needs to be recognized and dealt wit

DR. WATKINS: Can you clarify that? Obviously if you inhibit uptake of unconjugated bilirubin and then fractioning bilirubin would help show that, but do you know of clinical examples of inhibiting MRP2 that lead to substantially twofold

increases?

DR. KENNA: We have preclinical examples. I'm not aware of clinical examples.

DR. WATKINS: Because that would obviously be a very important finding if you could raise total bilirubin largely conjugated from just inhibiting a transporter.

DR. KENNA: Yes.

DR. WATKINS: That would be a major --

DR. KENNA: Yes. So I would like to see the guideline acknowledge that as a possibility and what we would clearly need to do, if we have a drug that we want to take forward that we believe has the potential to elevate bilirubin in a fashion that's uncoupled from liver injury, is have a plausible rationale for what's going on.

DR. WATKINS: See I would think that that would be a case-by-case basis.

DR. KENNA: Yes.

DR. WATKINS: You probably wouldn’t want --

UNIDENTIFIED SPEAKER: There's a rare instance of -- context --

DR. WATKINS: Dubin and Johnson syndrome, right. But would it be worsened by a drug? Maybe it would.

DR. KENNA: I agree it would need to be case-by-case, but I think the Guidance needs to address it.

DR. WATKINS: That there's a potential for it.

DR. KENNA: Yes.

DR. WATKINS: Because right now it doesn't even really address the fractionalization issue.

DR. KENNA: That's correct.

DR. AVIGAN: Two years ago, I talked a little bit about the transporter and there are actually genetic -- there's a Swiss group which has actually described these kinds of drug reactions especially in certain genetic variance where the transporter is already somewhat compromised and as I recall it was estrogen, other drugs as well, where they get a bump in bilirubin and the present with cholestatis actually but it's I think it's a conjugated variant because of this defect.

DR. WATKINS: Right. Well, that would be a cholestatic reaction that wouldn't be categorized as Hy's Law presumably. Dickens.

DR. THEODORE: I'm Dickens Theodore from Glaxo-SmithKline. Just to reiterate in terms of the guidelines' flexibility, certainly for us the stopping rule of eight times the upper limit of normal would actually be a hindrance in some populations, for example, in hepatitis C patients. If they have just minimal elevations from their baseline, you may be in a situation where you have to stop drug. In the Roche and Schering registration trials for peginterferons where patients entered the study with ALTs five times or seven times upper limit of normal, a slight increase in the ALT from baseline

would cause you to stop the drug. For certain patient populations I think it is reasonable to define the stopping criteria based on ALT increases from baseline instead of ALT increases above the upper limit of normal.

So I think it is just worth emphasizing that these are guidelines and sponsors have to have some ability to make rational recommendations.

DR. WATKINS: Right. So I think it does say in general something to that effect and maybe if we clarify it further, it doesn't apply to preexisting liver disease.

DR. SENIOR: When we were trying to write the Guidance, this was for people with presumably normal livers. Now you were talking about a drug that's known to cause inhibition of glucuronal transferase. It's a known effect that it reduces conjugation of bilirubin. Okay. That's a known effect. That's not the sort of serious toxicity that we have to worry about. You have to factor that in.

Now if you are entering people with hepatitis B, you may have fluctuations, as Arie pointed out to us yesterday, fluctuating levels. Then what is the baseline? Is the baseline just one measurement that you take on a given day? How much does it fluctuate? Is one measurement enough to establish a baseline in a person like that? Well, maybe not necessarily. In fact, one point does not determine a line. A line requires at least two measurements and preferably more if

it's a fluctuating measurement.

So common sense has to be put into this rather than just blind adherence to some arbitrary number. We are not wise enough to prescribe a number that should be followed based on no data, just based on some opinion. This is a guideline. It's not a regulation. It's not that you must do this. Just think about it and give some explanation as to why you're doing what you do.

Phase I is obviously different than Phase III; Well, of course, everybody knows that. So you may have different rules with baseline elevations. We don't have to define that for everybody because you're going to apply your scientific judgment to those issues. Just bear these things in mind and explain what you're doing.

DR. WATKINS: Yes.

MS. LYNDLY: Is the Agency going to address using elevated group means as a measure of liver toxicity?

DR. WATKINS: So like mean shift tables?

MS. LYNDLY: Yeah, ALT compared to controls.

DR. WATKINS: Yeah, I think that's out. It's not in the Guidance and now with the paradigm shift is they don't care about any of that. It's the individuals. It's the outliers. Mean shifts, for instance, in the Rezulin clinical trials, the mean ALT went down across all players. It turned out it treated NASH and nobody knew it. The focus of this Guidance is

to find the outliers and there's no mention of any means or any group data really. Is that wrong? What do you think? Mark, do you think --

DR. AVIGAN: I think that the NASH example is a great example where things may be going in the opposite direction, so that the mean would not actually reveal the problem. You might want to know what the mean is as a kind of measure for discussion, but I think what we're really going for, I agree with what you said, is really the outliers and whether you have a phenomenon of outliers based upon those three criteria.

DR. SENIOR: Jenna, this is a problem of the difference in thinking between statisticians and physicians. The physician is interested in the individual patient. The statistician is interested in the group, and you can't compare them. They're not the same. One way to hide a big change in 1 individual out of 100 is to take the mean of the whole group. It simply hides the data. It hides the people of interest. The patient of interest is obscured by taking the mean or shift or whatever you want to call it. So we have become more and more interested in the individual person rather than just the group, and a physician's approach rather than a statistician's. Okay?

DR. WATKINS: I hope nobody lost a job due to that.