The Story of eDISH
(evaluation of Drug-Induced Serious Hepatotoxicity)

John Senior and Ted Guo
For DILI Conference XV
18 March 2015

Good morning. My name is Ted Guo. I am a statistician. You might immediately think I am a wrong speaker showing up at the wrong conference. But should I talk? Actually, I talk about the medicine. And actually, I have never attended medical school. I don't know anything about medicine. And if I talk about statistics, you know, this is really the wrong place to talk. So, what should I talk about? Maybe I should just tell a story about eDISH. (Laughter.)
A historical account of eDISH

- Met John Senior by chance – 2002, then to -09
  - RSR (Regulatory Science Review) enhancement program to aid reviewers at CDER
- He had an interesting idea and I was willing to get outside of the “box”
  - What is DILI? Can a biopsy determine DILI?
  - He taught me about differential diagnosis
  - We used powers of computers to search data, and of humans to recognize patterns at a glance

About 10 or 12 years ago nobody knew what eDISH was, so let me tell you a story about how I got to know Dr. Senior. Maybe 12 years ago, we were all in the Parklawn Building. FDA was not on a big campus. It was just in one building and people were very close and nobody locked their office doors. One day when I was working, across the hallway there is a figure entering a vacant office. I didn't know who he was, and he was looking for something because nobody was in the office. The previous employee left the job. And then, there was someone knocking on my door who said, "I have some problems dealing with the data." And I saw he was Dr. Senior. He was holding four boxes of 3.25" floppy disks, if you remember them, four boxes. (Laughter.) Since the boxes were sealed, never opened, never used, apparently nobody had analyzed the data in the boxes.

So, I just said, "Let me see what I can do. I'll just use my own time to convert all the SAS dataset to an Excel file." Since then, we started to talk to each other, you know, about different things. Talked about our families, where we travel, and we got to know each other. Then, he came one day and said, "You know, I have an idea. Maybe you can do something about it."
This is the first graph of eDISH, but we didn't call it eDISH then. He talked to me about what DILI is and how to diagnose DILI. He used Excel for the data, and he is very good at Excel. He talked to me, but I was kind of ignorant. I said, what's the big deal? If the patient has a liver problem, do a biopsy of the liver and put tissue under the microscope; you will see the drugs in the liver and you will see how that affects the liver. So, what's the big deal? Why not do that?

Then he started to educate me about diagnosis, medical differential diagnosis and what that is. And I just learned. I was very interested in that. We both kind of went out of our own fields, went out of our boxes, and started to talk to each other. And I say, "Well, this is something I can do. Excel is very nice software, but it is not designed to handle a large quantity of data. I can do it in SAS and handle all of the data, 3,000, 4,000, 10,000 subjects. He said, "If a subject is located in the upper right quadrant, that may be a potential Hy's Law case. I want to know what happened to that subject."
And I said, "Well, maybe I can do something. I can do that in SAS. I can draw that plot. I can do an individual patient time course data record."

Talking about it today, there are now a lot of software-makers. They all say they can do eDISH. This is a good thing; it is welcome because, before that, nobody knew what the eDISH is. Now some software-makers and government contractors include eDISH in their software. So, I think we need to look at it closely to understand better what that is.

The Concept (Con’t)

• We wanted to see all the liver test data for a single selected subject over the time of study
That first graph I showed you, is that eDISH? Is that all? I know some statisticians try to apply some statistical methods to diagnose DILI. Can DILI be diagnosed from the numbers with statistical means? I know you need to diagnose and predict; you need to do a causality analysis assessment. You need some more information, not just ALT and bilirubin. You need to know the time course data, and then, you need to know something more. You need to read the subject's narrative, something like that.

As we worked together, and I was really educated by John; I learned a lot, even though I don't know anything about medicine. I kind of got his idea. I think I produced this in SAS, and I knew there is a nice feature. You can make each subject a hyperlink. You can move your mouse over any subject and click it, and it will immediately drill down to a time course data for that subject. I thought that's nice. I know that. Not many software people are using that feature.
Visualization of a Concept – Step 1

- More than 3900 subjects on one graph
- Drug X showed 7 times as many ALT elevations >3x ULN, and 14 times as many in RUQ with both ALT >3xULN and TBL >2xULN
- Inspection showed that many were probably drug-induced; no other explanation
- Drug X was not approved – and later it was found that certain people were especially susceptible to X-induced liver injury

But that is not ideal for a tool for everybody to use because you have to install some software on each individual, each user’s machine. And it would be nice to have the data and program on a server, so people in my organization can just go online and use it.

Fortunately, SAS has a feature called SAS Intranet that allows you to deliver this on a network. So, I had an idea how I could do it, but it took a long time to figure it out, to make it work. And I think for this, actually, John needs to explain it.
So, when we developed this tool called eDISH, working together, we spent some time to think about how to get the right data. We need to tell the sponsor what form of data we should include, what kind of variables we should include. We don't want to include many, many things that are unrelated to this software. So, this software is designed to be simple, to be small, to get the job done. We are not here to compete with other sophisticated software developers. So, we keep this in mind.

Then, we developed a spreadsheet that allowed people to download it, sent it to the sponsor, tell them exactly what we need to do. As a result, after we receive the data, usually in most cases within 30 minutes I can upload that to a server and people can start to run this tool eDISH to examine the data. No instruction, no training is required, no user manual is required. They just follow the few simple steps, and they can do it.
Now the most critical thing we need is the clinical narrative. I know that some software commercial developers claim they can replicate eDISH. I talk to them. I said, "How do you handle narratives?" They said, "Well, we generate the narratives automatically, programmatically." I ask, "from where?" They say, "from the case report form, from the sponsor's data." I thought this is not good, so I talked to John, who said this is not relative; it is not the way a narrative should be prepared.

So, we modified our data requirement. And recently, John spent some time writing a paragraph for who should write the narrative, what it should include, the purpose of the narrative. I think this is a very important piece of the puzzle.
So, we have three steps. First, the eDISH graph one: the purpose is not to predict or make diagnosis; it is to separate out from the mostly normal subjects a handful of subjects of special interest. That is all its purpose.

And we have dividing lines, two times, three times the upper limit of normal for bilirubin and ALT. And there is a reason to do that. Using statistics I found out that three times and two times, actually, are very close to the 95 percentiles. That makes sense because, then, that indicates a rare event. For a rare event, if you use 95 percentile as a cutoff point without any medical background, it is pretty close. And DILI is usually a rare event.
So, in our current version of eDISH, the threshold or the dividing line is fixed. Now we need to find a way to make it flexible. So now, probably it is time to talk a little about eDISH-2. So, what is that? Is that a second version? Is that an enhancement of eDISH-1? Is that an idea? Is that a necessity? I don't know. Before this meeting, I thought maybe we should talk only about eDISH-1, but I changed my mind in the last minute. I thought the speakers before me were talking about the change from baseline, talking about the measurement of ALT, there are a lot of issues. As a tool, eDISH should serve a purpose. We should have flexibility to change from the upper limit of normal to looking at a change from baseline. We should be able to change the thresholds. And after we change the threshold, what are we looking at? So, there are a lot of new issues. There are a lot of problems about how to get better narratives. We don't have a good solution, but we want to improve that part.

**eDISH2 : Enhancement of eDISH**

- More features, for research exploration
- More data to augment basic liver tests
- Clearer eDISH-data specifications, especially for the narratives
- The goal is to

*get it right, make the best possible diagnosis of what's causing the problem*
And this, just to show one advantage of our eDISH. I think every time we receive the data, if you run it once, the data are further standardized, installed in one single location, unlike some other software. Data are still either in the users' desktop or still scattered around in a network. But all the eDISH data is on one server. That is a good thing. We can in the future do research, a pool of data. That is very convenient.

Let me see. Now eDISH-2 should have more data than just serum chemistries. We are now talking about treating patients with hepatitis C or B, we need to get the viral load data. Sponsors submit that. We should include viral load data. If you look at the graph, the lines connecting black dots, that is viral load data. And the lower part is our data requirement. We add the viral load. That is the data on demand. You know, that is when the drug is treating hepatitis C. We probably need to ask the sponsor to submit that.

So, we continue to improve our data requirements, to enhance the tool. The tool is not only a review tool; it should be a review tool and a research tool. So, there is a lot of work to do.
How did I do? John: Are you going to talk about the narratives? Maybe you can start. I think John is going to tell you something very important; that is the narrative. That is something we encountered many, many times. Sometimes we got a lot of narratives generated automatically by machine. It is like a data dump. It doesn't serve any purpose. And a lot of times we found that people don't understand what the narrative should include. People ask, you know, "We have to generate it by hand, by writing." Yes, true, you cannot generate this by machine. You have to write down what the cause of those abnormalities. So, John is going to talk. I have already spent a lot, too much time on this. Sorry about that. John?

DR. SENIOR: That's all right, Ted. Can you all hear me? (Chorus of yeses.) Okay. I don't want to go up on the stage because I am so unsightly. I cut my lip, shaving this morning, and it kept bleeding because I'm anticoagulated on warfarin.

I think there are just a few more slides. I am not going to read them. You can read them faster than I can say them. What I will say is that Ted Guo and I have been working across these two major disciplines of statistics and medicine. Now statisticians are trained and are very skilled at analyzing data, but what they are not trained to do is to diagnose patients, because that is what medical doctors, physicians, do from day one in medical school. When they see a patient, they immediately think about what is the cause of the patient's problem. That is a diagnosis. Why is it important? Because medical doctors, as distinct from other kinds

Get a clinical narrative - Step 3

- Dr. Senior will comment briefly about narratives
of experts --- pharmacologists, toxicologists, chemists, and statisticians, --- medical doctors have a responsibility to treat patients, to write prescriptions, to order tests, to do something to treat the patient that nobody else has the responsibility or authority to do. And therefore, if they are going to treat the patient correctly, they have to make the correct diagnosis. So, they are the only ones who make diagnoses.

Trying to get statisticians to understand the process of medical differential diagnosis, finding the cause, is not so easy. And eDISH is a diagnostic tool really to help the medical doctor make a diagnosis using statistical data.

For eDISH we are using the computer and we are using the human mind. They are different. The human mind is very good at recognizing patterns, recognizing faces. I can recognize Ted Guo. I can tell Arthur Karmen in a half-second, even though I haven't seen him in years. I can immediately look at them and in one glance I say, "That's Arthur Karmen," "That's Patrick Kirby," "That's Naga Chalasani," "That's Leonard Seeff." I can do that zip, zip, zip. A computer can't do that. A computer can't recognize faces as well as humans can. We have been trained over eons of time to protect ourselves by quickly recognizing friends and enemies. So, we recognize our friends and we greet them. The enemies we stay away from.

So, computers and human minds work in different dimensions. On one hand, as Ted Guo said, I came to him with disks, with data from thousands of patients entered on them. It would take me weeks to do what his computer can do in less than a second, analyze the data. The computer is powerful in quickly analyzing data that the human mind cannot deal with.
Get a clinical narrative - Step 3.1

- eDISH is meant to be a clinical, medical diagnostic tool, not a statistical exercise of counting numbers;
- protocols cannot anticipate all possible adverse effects of a drug, especially if rare, so extra information is needed;
- physicians uniquely worry about causality, because they have responsibility to treat or prevent problems;
- ideally, narratives should be written by MD investigator at the site who can determine what extra information is needed to establish the probable cause of the problem;
- narratives should not be made by summarizing case records

On the other hand, once you display the data on all the patients and ask "which ones have both elevated bilirubin and ALT?" and you display that on a graph, a human mind can look at that graph in a blink of an eye and say, "Oh, I'm interested in the few patients up in the upper right quadrant. I don't want to worry about the patients whose data are all normal." So, I can recognize the pattern at a glance.

But, then, we say, let's go back and have the computer tell us everything that is in the data about that one person. Give us a time course. Now a time course is really adding a third dimension. One dimension is the ALT; that is the injury dimension. Another dimension is the bilirubin. That is the dysfunctional measure, a second dimension. And the third is time. How do the data change over time, day by day, week by week? Is the ALT rising? When does the bilirubin rise? Is it falling? That is a very important piece of information that helps diagnose what is going on.

So, putting the time course up, which takes the computer another half a second, allows the physician to look at the time course and interpret how the values are changing and what that might mean for making a diagnosis.
But we then need a third step. We need additional information that was not in the protocol. It was not in the case reports. And there is no use searching in the case report because the information isn’t there. It wasn’t anticipated as needed when the protocol was originally written. You never can anticipate all of the things that might happen.

Get a clinical narrative - Step 3.2

- Hy’s Law is NOT just elevated ALT and TBL, but requires a determination of probable cause; there is no such thing as “Hy’s Law chemistries”
- Hy Zimmerman said “drug-induced hepatocellular jaundice is a serious lesion”, with considerable mortality. The first requirement is that the drug caused the liver injury, and not something else. The second is that hepatocytes are injured and not biliary obstruction. Third, jaundice results because of the first two, indicating that enough hepatocytic injury has occurred that remaining cells cannot clear plasma bilirubin.
- How many times do we need to repeat this?
So, the protocol and the case report cannot or should not be used to create a narrative. You have to go back to the physician at the site, the investigator, who is a medical doctor. All of the investigators of clinical trials are MDs, for a good reason, because they have responsibility to protect the safety of the subjects participating in the study.

If you are experimenting with a new drug, there might be injury. It is important that the physician in charge of the study site be ready to interpret what is going on and, if necessary, stop the drug or interrupt treatment with the drug, or do something else, maybe take further action. But certainly, diagnostic consequences follow what the data show at the study site.

So, Hy's Law is not diagnosable by just chemistries. You need more information. You need the time course, not just one day, but the changes over time, and supplementary information. You need to know what other information the physician in charge at the study site used to determine whether or not the drug should be stopped or not. It is not in the chemistries; you can't do it automatically by chemistries alone.

So, preparing the narratives should be done thoughtfully by the person who is responsible for the welfare and safety of the study subjects, not by a data analyst, not by a clerk, not by a project manager, but by a physician who has the legal responsibility for the safety of the subject.
Clinical trials are more than just gathering papers and numbers to get an approval. The safety aspect of clinical trials is very important, particularly in the field of liver injury, because any drug is potentially capable of injuring the liver, sometimes fatally. In 1997, the FDA approved eight drugs that had to be removed from the market because they were killing people. Four of those were for liver problems, three for cardiac problems, one for muscle problems. But four out of the eight were liver toxicity. That was the year that triggered off the first of these conferences. I will say, as a result of that, the FDA has raised the consciousness of its reviewers, and the reviewers have raised the consciousness of the pharmaceutical companies. As a result of that interaction between the reviewers at the FDA and the sponsors making the new drugs, no drug has been approved by the FDA since 1997 that has had to be removed from the market for fatal liver toxicity.

It is not that minor toxicity doesn’t occur. Sure, we get transaminase elevations. So what? Mild injury is not killing the patient until the function of the liver is so badly disturbed that it can no longer do its job. We can cut out two-thirds of the liver and throw it in the bucket, or we can injure chemically two-thirds of it so that it is not working. And the liver is still able to regenerate, remarkably, more than any other organ that we know of, and the person lives.

So, all of this has to be factored into what Leonard was just talking about, how you make the right diagnosis of what caused the problem. How do you know that what is

---

Get a clinical narrative - Step 3.4

- Clinical trials are not just data gathering exercises to get information to support approvals of new drugs, but are real world tests of drug effects, both good and bad in real people who may vary in how they respond;
- It is difficult to specify all the details needed for writing a good narrative. Go to medical school first, then practice a while, and you may begin to know what’s needed;
- The purpose of eDISH is to assist reviewers to scan over all the subjects, learn which ones may need special attention and further investigation to understand the cause of their test abnormalities and clinical findings.
happening was caused by the drug and not by some disease, not by something else? That is not so easy. It must be done by somebody who has spent a whole lifetime making diagnoses, and that is a physician.