Good afternoon, everyone. First, thanks for inviting me here to introduce our work. I will talk today a little about the LDKB work. I am a toxicologist or I can say bioinformaticist, not a clinician. So, I will give you from my perspective whether drug properties can predict drug-induced liver injury.
We have talked many times in the DILI field. One big challenge, I think is the lack of a reliable predictive model. Especially today, we don't have a good animal model predict to predict human effects.
As we know, FDA still relies on the high-dosing healthy animal study. This study still can only identify 50 percent of DILI problems. This technology was developed more than 50 years ago, so we need some new technology to improve predictions today.

Preclinical toxicology often failed to predict DILI risk in humans
- “High-dosing of healthy animals” test identified < 50% DILI liability in humans
- Conventional in vitro assays have very limited prediction for DILI risk assessment
- No biomarker is available to identify the susceptible patients prior to drug treatment

So, we developed a project called the liver toxicity knowledge base. And this database provides a better predictive model. We have put some of the collected data in the particular model to a public domain. We can either use a LTKB such as Google to find that.
This slide gives you some more idea what data we have in our database. And basically, we have collected about 3,000 drugs. And basically these drugs, including almost all the academia drug, drugs that were pulled by the other agencies. Basically this we started to collect the human data and the non-human data or we collect part of the data. For the human data, we tried to collect all kinds of the DILI-related information, especially we have it noted as a DILI risk associated with the drug. For the drug property data, we also collected each drug from the chemistry property. DILI markedly related individual assay or some whole special biology risk poles using microRNA data or this other data.
At the end of the day, we tried to correlate these drug properties with human data, build a particular model. This is our goal to do the project. To develop a particular model, we need to list the drug have known DILI positive and DILI negative. The amount of the DILI drug in this model is to develop all kinds of translational biomarkers.

We tried all kinds of approaches. Finally we found that drug labels are good enough to serve our purpose. The drug label, basically, is an information tool. It provides certain data to the doctor and the patient. By the way, the FDA should inform the patients about the drug label.

**How to Assess DILI Risk in Humans for a Drug?**

- Three attributes of a drug are important for its DILI assessment:
  - Causality: was liver injury caused by drug or other cause
  - Incidence: how many case reports are considered significant
  - Severity: elevated ALT; Hy’s law; disability and hospitalization, liver failure; liver transplantation or death

- Risk = (How likely) x (How many) x (How severe)

*This is opinion based !!!*
We agree that the drug label is not perfect but it might be the most consistent, best information we can have to help us codify the drug.
We published a paper several years ago, describing our approach using drug labels to identify DILI drugs. The drug label has three sections to disclose a DILI risk: Box Warning, Warnings & Precautions, and Adverse Reactions. Dr. Temple discussed drug label a bit yesterday, so I don’t want to repeat today. If you are interested, go to our 2011 paper (Drug Discov Today 16:697-703, and get more details. This approach, classified each drug into most concern, less concern, and a non-DILI concern.

After we had risk classification by labeling and we know the drug is a DILI drug or a non-DILI drug, we then go to our LTKB data.
We tried to develop some predictive model based on our drug property data. The data we thought about was the daily dose, because most of the DILI drug we know was given -- but the daily dose alone basically is not predicting now because we know many signature, also given the 100 milligram.

We thought about whether we could find some other way to help. The LTKB database finally found that lipophilicity can also help for this purpose. If you could use, the DILI we are marking here, we found if the drug dose was more than 10 mg, then there was toxicity. Most non-DILI drugs got kicked out. Because of the rule of 2 there is a significant association with DILI risk.

The “RO2” models is a rule we found in the LTKB research. Through the survey of 164 FDA-approved drugs, we found that an oral medication with high lipophilicity (logP > 3) and high daily dose (DD > 100mg) was significantly associated with severe DILI risk. Basically, in the area that logP > 3 and DD > 100, the red one is DILI drug and green one is non-DILI drugs, we found most of drugs located in these area were DILI drugs with very few false positives. This so called “RO2” rule were validated by an independent dataset with 179 drugs,
and 5 drug pairs and some co-medications. The results has published in the hepatology, and got some positive comments from DILI experts. For example, Neil has comments that the “RO2” offered some added values to identify the potential of idiosyncratic DILI.
I show you some more examples to demonstrate the Ro2, using drug pairs. Drug pairs are basically two drugs capable of causing the same or similar effect and have similar structures, but show toxicity differences. For example, alpidem and zolpidem, two drugs with high logP, greater than 3, but alpidem had a much higher dose. Now look at troglitazone and two other glitazones: troglitazone, has a larger logP greater than 3 but only troglitazone had a much higher dose than the pioglitazone or rosiglitazone. Another example is bosentan. Dr. Temple mentioned yesterday this drug was also a RO2-positive drug. Its daily dose is 400 milligram, and AlogP also greater than 3.
We also show that logP helps in other cases, for example, tolcapone and entacapone. Those are drugs that have high doses but only tolcapone has the much higher logP. The same applies to nefazondone and trazondone.

But we don't say that RO2 always works. The RO2 only has limited sensitivity, about 30 to 35 percent. We have some false negatives, and false positives, for example, trovafloxacin, a drug we know was withdrawn. The daily dose is about 200 mg but logP is very low.
We wanted to know how to work on all FDA-approved oral drugs. So we collected all drugs approved by FDA before 2010, 748 oral drugs. And of these we had 168 drugs with most DILI-concern in labeling, but Ro2 identified only 72, about 43% sensitivity. Next, 193 drugs with no DILI-concern, of which only 11 drugs were ALT positive. That means that specificity was about 95%. There were 387 drugs of less DILI concern, but we only identified 13% as ALT positive.
We also wanted to know whether the Ro2 could help us identify drug failures in clinical trials or in drug development. Interestingly, in this model, Dr. Regev presented a drug with daily dose of 225 mg and AlogP of 3 to 4, a RO2-positive drug, a drug we discussed this afternoon. In this other drug, they had a daily dose of 120 mg and logP is 4.1, another RO2-positive drug. So, both drugs discussed today were RO2-positive.

You can see some more examples here, collected from the literature. but some are RO2 positive, some negative. But anyway, it shows that RO2 can identify some of the hepatotoxic drugs during drug development. We also want to call industry to study the failing drugs more, to learn if they can help give us a better predictive model. We know RO2 has limited sensitivity and we are trying to incorporate some more related data.
And finally, in this paper we use a high-content screen assay to improve sensitivity from 30 percent to 50 percent.
Going to the question John asked me: Are drug properties or host factors predictive? I think this cartoon is a very good answer to the question. In this cartoon, there are blind people who want to know what an elephant looks like. The first time, they don't agree because they are concentrating on a different part of the elephant. But very interesting, at the end of the story, original story, these blind men stopped talking and they started listening and collaborating. And then they envisioned the whole elephant.

So, we have some blind people discussing our chemistry. If we were to figure out what the data looked like, at least addressed, we proposed DILI basically an interaction between the drug property and the host factor. Drug properties and host factors work together to initiate cellular injury. In the individual patient, the host factors will contribute to the individual response and then finally determine the final outcome. So, I suggest considering in a DILI case not only the host factors but maybe also the drug properties, to help you understand what DILI is.
Overall, we believe that drug properties and host factors together contribute to DILI prediction, DILI development.

Although LTKB has collected diverse DILI-related drug property data, it can be helpful for understanding. We have developed a predictive model. A comment from Dr. Kaplowitz was that RO2 has added value to predict idiosyncratic DILI. We also believe if we incorporate more data, it can be improved. We still have a long way to go to make a better predictive model.

Finally, I give a summary for my talk today. Firstly, to answer the question we proposed, yes we believed the drug properties did significantly contribute to the prediction of DILI. Based on this belief, we developed a large drug database systematically collect the diverse DILI-related drug properties data and annotated with DILI potential in human based on FDA-approved drug labeling.
Finally, I want to thank the many people who helped me on the LTKB project, and especially the LTKB interest group. And also we thanke many people in this room. Especially I want to thank our collaborator Dr. Jurgen Borlak from Germany and my colleagues at NCTR. Thank you so much.