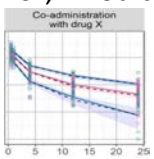
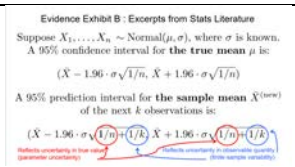
OPENING

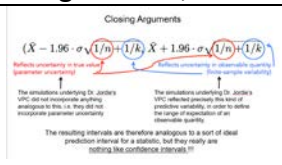
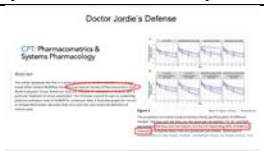
NAME	ACTION
S/J	• We will now call Courtroom 6C to order. • Opening credits with music
S/J	• "All rise for the Honorable Judge KK"
Ken	• *Walks in, sits down* • "You may be seated. • The plaintiffs and the defendants for these cases have all agreed to appear in my court and to abide by my rulings. To those of you in the gallery, I expect decorum and will not tolerate any disruptions. • I will not hesitate to call on Bailiff French or Bailiff Tannenbaum to escort you out of the courtroom! • Bailiffs, I know we have a busy docket today so let's get started with the first case"

CASE 1: The Case of the Crisis of Confidence- Jim and Eric

NAME	ACTION
S/J	• **The next case is "The Case of the Crisis of Confidence".** • This is Jim Rogers. He is suing his colleague and ex-friend for disrespecting statistical principles. Jim's colleague's description of the intervals in his visual predictive check caused such anxiety that Jim had to take a 2 week leave of absence from his job at Metrum Research Group.
S/J	This is Eric Jordie. He has several aliases including Bob Ross and Wicked Lobstamatrician. His Reckless Indifference to the Proper Use of the term "Confidence Interval" caused pain and suffering for his once friend and current colleague, Jim Rogers. Eric refuses to listen to reason, and for that, Jim is taking him to a court of law.
Ken	Bailiffs, please swear in both parties.
S/J	Statistician, please place your left hand on XXX and raise your right hand. Pharmacometrician, please place your left hand on Rowland and Tozer, and raise your right hand. Do you both swear to tell the truth, the whole truth, and nothing but the truth, so help you? [I do] Your honor, the parties have been sworn in.
Ken	Dr Rogers, would you please present your case
Jim	Your honor, this all began when my colleague asked me to review this material that he had written about one of his models …
Eric	(interrupting) OBJECTION!!! It isn't overconfidence when you are right! Puts sunglasses on!
Jim	ORDER!! IN THE COURT!!!

NAME	ACTION
Ken	*(BANG GAVEL)* Pipe down, both of you. First of all, Doctor Rogers, I am the one who calls order in this court. Secondly, Doctor Jordie, please remain silent while Doctor Rogers presents his case- and remove your sunglasses.
Eric	***Indignant and dramatic eye-rolling and gesticulation***
Jim	Your honor, I would like to point out a few essential features of this evidence ... [SLIDE]  Now your honor, there are two types of intervals represented in that figure. The interval formed by the lower dashed line at the 10th percentile and the upper dashed line at the 90th percentile in an interval that, in general terms, is meant to contain the observations themselves. My colleague has referred to that as a prediction interval. And frankly, even that is a debatable use of terminology, but Lord knows I've got to pick my battles when I'm working with this guy, your honor, so I'm just going to focus on the other type of intervals that you can see in the figure. My colleague has referred to these other intervals, the shaded areas around each percentile, as "confidence intervals". Now, your honor, when I hear about a confidence interval for a percentile, I assume that the confidence interval is meant to contain the true value for that percentile.
Jim	And I read through the whole report on the assumption that they were in fact confidence intervals and that they were in fact meant to contain the true values of the population percentiles. *<voice rising>* And then, after I struggled through the whole darn thing, he and I were talking, and he explained how those intervals were made, and it turns out they aren't confidence intervals at all! *<voice rising again>* Your honor I could have been saved so much time and so much consternation if he had just used words in the normal way so I could understand what the heck he was talking about !!! Your honor, please make him speak like a normal person!!!
Ken	Calm yourself, Doctor Rogers. Let's get to the bottom of this. Do you have evidence that these are not, in fact, confidence intervals?
Jim	Yes. Your honor, I would like to submit into evidence some examples to illustrate how these terms are used in the statistics literature

NAME	ACTION
Jim	[SLIDE]  *[Show intervals, highlight distinction between parameter uncertainty and finite-sample variability.]*
Jim	And now, your honor, I would like to call the defendant to the witness stand!
Ken	We actually don't have a witness stand, but feel free to question him where he is.
Jim	Very well. Let the record reflect that defendant is disrespectfully sitting at the witness stand.
Ken	I'll worry about the record, Doctor Rogers. You do you.
Eric	***Makes a childish self-satisfied gesture towards Jim.
Jim	Hey hey hey!!! Your honor, did you see that? Permission to treat the defendant as hostile!
Eric	If you think I am hostile now...
Ken	*Gavels* Both of you- PIPE DOWN! Permission granted to treat the defendant as hostile. Dr Jordie, if you show disrespect in my court, I will hold you in contempt!
Jim	Doctor Jordie, in order to make your so-called *(<eye roll to audience>)* "confidence interval", am I correct that you first generated many simulation replicates, each one representing a hypothetical replication of the data set that you used to fit your model?
Eric	Yes.
Jim	I see. And would I also be correct in assuming that, from one simulated replicate to the next, you used exactly the same parameter values every time, namely the point estimates of those parameters from the initial fit?
Eric	Yes.
Jim	I see. So, would it be fair to say that you simulated as if the parameter values were known?
Eric	Yes.
Jim	I see. In other words, you simulated **without parameter uncertainty?**
Eric	[to Ken] Objection! Your honor, asked and answered. This is badgering.
Ken	Please move on, Doctor Rogers.
Jim	Yes, your honor.
Jim	And so, Doctor Jordie, the fact that you obtained a different estimate of the quantile for every simulation replicate, does that simply reflect variability in the observable quantities?
Eric	Yes.
Jim	I see. So, it reflects **data variability but not parameter uncertainty**?
Eric	Objection! Badgering again, your honor! What's worse than badgering? Mongoosing?

NAME	ACTION
Ken	Please bring it home, Doctor Rogers.
Jim	[SLIDE]  [*present argument on slide]*
Jim	Your, honor, I hope you can see why I am confused!! My colleague indicated to me that the interval in question was a confidence interval, and yet the width of his interval has nothing to do with parameter uncertainty! How on earth am I supposed to read his reports, when he says one thing and means another?
Eric	You want the truth Jim Rogers, you want the truth? I don't think you can handle the truth Jim Rogers!!!!! *Points crazily at Jim* Don't say I didn't warn you!!!!
Eric	[SLIDE]  The truth is Jim Rogers that I was running an ISoP tutorial figure and language about VPCs past you because I like seeing you get all riled up! Had I only known that I would end up on stage at ACoP as a patsy… So, Jim Rogers, I flip the question back on you...how on earth am I supposed to read ANY reports when the ISoP guidance says one thing but means another? If I had a microphone, I would drop it...
S/J	• We would like to do a poll of the gallery. You may cast your vote by applause. Dr Jordie has been accused of Reckless Indifference to the Proper Use of the term "Confidence Interval" o Who thinks Dr Jordie should be found NOT GUILTY? o Who thinks that Dr Jordie should be found GUILTY?
S/J	• "Judge Kowalski, you've heard the popular opinion. what is YOUR ruling?"
Ken	• Before I make my ruling I will present my thoughts on this case including some basic definitions and distinctions between confidence intervals and prediction intervals. [Present slides here]

NAME	ACTION
Ken	• **Will the defendant please rise?** • I find the defendant **guilty** of misleading use of statistical interval terminology for both CIs and PIs in performing VPCs • However, there are mitigating circumstances that inform the sentencing of the defendant: 1. The defendant was merely following guidance endorsed by ISoP. 2. The stochastic simulation procedure used to perform an internal VPC are appropriate and the resulting intervals that are constructed have diagnostic value. 3. While the resulting intervals do have diagnostic value, they do not maintain nominal coverage probabilities for repeated experiments. However, there is no commonly accepted terminology for these VPC diagnostic intervals. ▪ I suggest that the SxP SIG survey its membership to publish new terminology for these VPC diagnostic intervals so as not to be misleading to a broader statistical audience who may not be familiar with pharmacometrics practice.
Ken	• I sentence the defendant to <u>revise</u> his report to provide more detail regarding the simulation procedure used to conduct the VPCs and acknowledge the limitations in the use of the statistical intervals constructed as having diagnostic value but are not intended for valid inference for repeated experiments. • Moreover, I encourage the defendant and other pharmacometricians to read the following articles by Gerald Hahn as well as attend my workshop tomorrow. [present slides here] • In addition, Dr Jordie, I am holding you in contempt of court. Bailiff French, please take Dr Jordie into custody! • **This case is adjourned! (gavel strike)** • *[guilty slide]*
S/J	• Thank the participants, call the next case!

CASE 2: The Case of the PHARMACOMETRICIAN's PITIFUL PLANNING

NAME	ACTION
S/J	**The next case is "The Case of the PHARMACOMETRICIAN's PITIFUL PLANNING".** This is Professor Mentré, from Paris Polytechnical. She is particularly proud of her protocol planning prowess. Period. She's properly pissed at her pretentious pharmacometrician for poorly preparing plus not practicing proper power procedures.
S/J	Defendant Dr Jin delights in describing difficult datasets. She thinks that Distinguished Dr Mentre is being dramatic in her dogged defiance of her design decisions driving drug development! She is accused of dastardly deeds… as she doesn't depend on D-optimality.
Ken	Bailiffs, please swear in both parties.
S/J	*[direct- left hand on book, right hand in the air]* Do you both swear to tell the truth, the whole truth, and nothing but the truth, so help you? [I do] Your honor, the parties have been sworn in.
Ken	Dr Mentré, you are accusing Jin Jin of "Poorly planning her pharmacometrics study". Please present the evidence for your case
France	I read your report on the population pharmacokinetics of drug FIM in patients with Seurat disease. I found the following table in your report summarizing your population PK modeling results: *[model parameter table]* Based on the study design and assumptions, *<change slide and review contents>* this is a poorly designed study. You should have better evaluated the number of subjects and number of samples per subject *[next slide] –PFIM table* Since you didn't do this evaluation, I believe that you are guilty of poor planning!
Ken	Dr Jin, how do you respond?
Jin	Your honor, this accusation is unfair! Thank you for the opportunity to defend my case. The primary objective of our Phase 3 trial is to evaluate efficacy and safety in the target patient population. PK is secondary objective and cannot be evaluated at the cost of extra patient burden and inconvenience unless absolutely needed. First, regarding collecting only 2 samples per subject: • Patients with Seurat disease suffer from pain and immobility. Their site visit needs to be accompanied by their family or caregiver. PK sampling is a painful and burdensome procedure for patients and their families. • We proactively conducted a survey with clinicians and patients in this disease area. Based on their feedback, the team considered 2 samples per visit as maximally feasible.

NAME	ACTION
	• Unnecessary patient burden and inconvenience may also lead to slow trial recruitment and higher risk of patient drop-out. This may result not only in an increase of trial duration/cost, but also a higher risk of a failed trial. It will delay getting therapy to this rare disease patient population with unmet medical need. Secondly, regarding only 50 pts in the popPK analysis: • We did plan to collect PK data in all patients as stated in the clinical protocol. Unfortunately 50 of the 100 patients did not have any measured drug concentration, therefore the popPK analysis only include data from 50 patients. • Of the 50 patients without PK: o 10 dropped-off before Day 10 due to various reasons o 20 did not have any PK data due to unsuccessful seampling (e.g. too painful to complete the procedure etc) o 20 of the patients' PK samples were damaged due to inappropriate handling, storage, shipment, or assay. Overall, the handling issues arose because the recruitment was performed in multiple centers from several countries for this rare disease, and some centers are inexperienced with PK sampling. Presumably, if France's recommended design was implemented she would have also encountered the same logistical/conduct issues that we faced. *[go to table slide]* Referring back to the Accuser's simulation exercise, N=100 design cannot be implemented therefore is not a fair comparison. Comparing the lowest 2 rows in the table, there's no meaningful additional benefit for collection of 3 samples instead of 2 plus there is an additional patient burden. Given these legitimate reasons as detailed above, I plead NOT GUILTY to the accusations of poor design and planning!!
France	All this operational stuff may be true, but pharmacometricians should better understand the challenges of such a study, and work with the study statistician to run the calculations ahead of time in preparation. Regardless of the design look at how much you are losing in power by allowing the trial conduct to lose half the subjects. If you would have performed the PFIM analysis you would have known the impact of the loss in information if you lost half the patients ahead of time and could have worked harder to communicate to the sites how important it was to try and make sure patients comply and that the PK samples were treated with the utmost care. Therefore I urge judge Kowalski to consider the impact of not doing such an analysis and to find the defendant GUILTY!

NAME	ACTION
S/J	Judge Kowalski, what is your ruling?
Ken	I think there are valid arguments on both sides. I have <u>some</u> experience in this area but I am sure that other industry scientists have experienced similar logistic and compliance challenges to the ones that Dr Jin has encountered. For these reasons I would like to bring in the collective experience of a larger pool of PMXs and statisticians, and send this case to the jury. *[Gesture towards jury]* Matt Zierhut, I understand that you have been selected as the jury foreman. Please distribute the ballots to your fellow jurors and collect their votes. When the vote is completed, please tally the votes and come to the podium to announce your verdict and sentencing. Wil the defendant please rise? *(as Matt comes to Podium)*
Matt Z	we the jury find the defendant …" *(click button for guilty/not guilty)* Suggest the sentence *The jury sentences the defendant to visit a clinical site that is participating in a Pop PK substudy, so that the defendant can appreciate the challenges they face as they are not as well equipped to obtain PK information as Phase 1 clinical sites.* *Consider sending her to an optimal design course and/or working more closely with her statistical colleagues.*
Ken	Thank you to the jury for your service. This case is adjourned (gavel bang)
S/J	Thank the participants, call the next case!

CASE 3: The Case of The Extrapolated Prediction

NAME	ACTION
S/J	**The next case is "The Case of The Extrapolated Prediction".** • This is Dr. Lei Nie. He recently moved from the company Phase-3-Trials-Inc. to work on early development trials as a statistician at Extrapolations-R-Us. Dr. Nie presented the statistical analysis plan to the CEO, but he was deeply humiliated when his PMX colleague suggested that Dr Nie's work was not only not necessary, but that there was no need to do the study at all! He is asking for $400K in damages
S/J	This is Dr. Chao Liu. He claims that the study wasn't necessary because the outcomes could be predicted based on data that is already available in house from previous trials. He thinks that Dr Nie is completely overreacting and should stop extrapolating his problems at work onto his colleagues! He is being accused of making drug development decisions based on incomplete data.
Ken	Bailiffs, please swear in both parties.
S/J	*[direct- left hand on book, right hand in the air]* Do you both swear to tell the truth, the whole truth, and nothing but the truth, so help you? [I do] Your honor, the parties have been sworn in.
Ken	Dr. Nie, you are suing Dr Liu for making drug development decisions based solely on extrapolated predictions. You are suing for quite a large amount of money. How did you come to that figure?
Lei	I took his current salary and extrapolated it 10 years into the future using a Bayesian Gaussian process.
Ken	Wow, you must be good friends that you know how much Dr Liu makes! However if you're going to sue him for $400K that might end up being the end of your relationship!
Lei	Well… it's too late now! We're here so let's do this!
Ken	OK. Dr Nie, please present your case.
Stat (with LS)	• Thank you, your honor. Dr. Liu and I work for the pharmaceutical company ***Extrapolations-R-Us.*** Dr. Liu has been advocating for making drug development decisions based solely on extrapolated predictions without any empirical data to support these predictions. • As an example, Dr. Liu o developed a PK model and an exposure-response (ER) model based on data from a dose-ranging study with Formulation X. *[Show Fig slide 1]* o subsequently developed a PK model for Formulation Y which had better absorption than Formulation X (which had poor exposures at early time points). *[show Fig slide 2]* • Dr. Liu then o proceeded to use the ER model for Formulation X together with the PK model for Formulation Y to make PD predictions for Formulation Y. o recommended a clinical dose for Formulation Y for subsequent clinical pharmacology studies. • I claim that Dr. Liu is over-selling these extrapolated predictions

NAME	ACTION
Ken	Dr. Liu, you are accused of using models to make decisions rather than actual data. What is your response to Dr. Nie's concerns?"
Chao	• Dr. Nie is correct that we developed an ER model for Formulation X and a PK model for Formulation Y to make PD predictions for Formulation Y to determine a clinical dose. • These predictions were made on the very reasonable assumption that the ER relationship should be the same for both formulations. • This clinical dose prediction is supported by *Extrapolations-R-Us'* understanding of the pharmacology of this compound. • In addition, VPCs were performed for the ER model for Formulation X as well as for the PK model for Formulation Y confirming the adequacy of the predictions. • We make these kinds of extrapolated predictions all the time – I'm just doing my job!
S/J	• *Audience participation portion [time dependent]* o "Man on the street" question(s) of the audience – "how do you think Judge Kowalski should rule?" • "Judge Kowalski, what is your ruling?"
Ken	• I will share my thoughts on the role of extrapolation based on statements published in the statistical and pharmacometrics literature that inform my decision. o [Give presentation] • Before I make my ruling I have some advice for both the plaintiff and the defendant. First for the plaintiff: o We must recognize that drug development often requires making decisions based on extrapolations throughout the drug development process ▪ In Vitro → Animals → Healthy Volunteers → Patients ▪ Single Dose → Multiple Dose ▪ Special Populations (e.g., Elderly, Pediatric, Renal & Hepatic Impaired) ▪ Formulations (e.g., Capsules → Tablets) o Pharmacometric models leverage our understanding of the pharmacology and disease we are trying to treat o These models can and should inform drug development decision-making even under extrapolation

NAME	ACTION
Ken	• Now for the defendant: o The plaintiff is concerned that you are "over-selling" your predictions. This concerns me that there may be a communication issue. o When making extrapolated predictions explicitly acknowledge that you don't have data to confirm such predictions at the time and that these are "*hypothesis generating*" results rather than confirmatory conclusions. o Be an <u>advocate</u> for study designs and collecting trial data that would allow you to confirm these extrapolated predictions. o Once data becomes available under the new experimental conditions, perform an external VPC based on the model that provided the extrapolated predictions before updating the model with the new data. o Confirming extrapolated predictions on independent data from the new experimental conditions not used in the development of the model will help build trust between pharmacometricians and statisticians (as well as other team members). • **Will the defendant please rise?** • I hereby **dismiss** this case on the grounds of improper foundation for the charges that the defendant promoted the use of extrapolated predictions without empirical evidence to confirm these predictions. o Yes, the pharmacometrician performed such predictions, but I am in agreement with the defendant that he was just doing his job. o By definition, such extrapolated predictions cannot be confirmed with empirical evidence at the time that they are made. However, the predicted clinical dose would eventually be evaluated in subsequent trials to hopefully confirm its effectiveness and safety regardless of whether the predictions are confirmed. o The concern of over-selling the predictions is a subjective assessment that is difficult for me to rule on. I encourage both sides to more effectively communicate with each other regarding the value and limitations of such extrapolations. "CASE DISMISSED! " *(Gavel strike)* *[CASE DISMISSED slide]*
S/J	• Post-case interview- 30 seconds to 1 min each o Dr. Nie, what do you think of Judge Kowalski's ruling? o Dr. Liu, are you satisfied that justice has been served?
S/J	• Thank the participants, call the next case!

CASE 4: The CASE OF THE NOT SO CONFIRMATORY CONCLUSION

NAME	ACTION
S/J	• **The next case is "CASE OF THE NOT SO CONFIRMATORY CONCLUSION".** • This is Brian Smith from the Seven Miners institute. His dopey PMX colleague is making confirmatory conclusions from exploratory studies. Doc Smith is never bashful about voicing his opinion, but this colleague doesn't listen, and he's losing sleep about it. This is making him grumpy!
S/J	• This is Nag Chemuturi. He would be happy if Doc Smith stops badgering him about his Pharmacometrics analyses. His list of covariates is nothing to sneeze at, and he's doing his best and digging deep to mine the data. • He is accused of inflating the significance of a covariate effect in a population pharmacokinetic analysis
Ken	Bailiffs, please swear in both parties.
S/J	*[direct- left hand on book, right hand in the air]* Do you both swear to tell the truth, the whole truth, and nothing but the truth, so help you? [I do] Your honor, the parties have been sworn in.
KEN	Dr. Chematuri, you are charged with inflating the significance of a covariate effect in a population pharmacokinetic analysis. How do you plead?
NAG	Not guilty, your honor.
KEN	Can you tell me a little bit about the analysis in question.
NAG	I work at the Seven Miners institute for sleep related disturbances. We are a leading organization in helping patients develop effective strategies in order to get high quality sleep. We know that poor sleep patterns not only affect the quality of life, but also have long term health consequences.
KEN	(yawns) I know this to well, but tell me about your research.
NAG	You may have seen advertised on television the new natural product sleep aid, Sleep Beautiful, by Queen and Hag Inc.
KEN	Sure sure, they have that catchy jingle "Go to sleep, go to sleep, go to sleep little princess"
NAG	That's it. Our clinicians were very excited about this product since it seemed so effective. Although no one is quite sure how the active ingredient works, we have had tremendous success with our patients using this product.
KEN	Ok, that sounds exciting, yet we are no closer to the research in question.
NAG	Oh, I am getting to that. I can be a little dopey sometimes. Anyways, there is an apparently idiosyncratic adverse reaction that occurs in about 1 tenth of 1 percent of those that take the product. It appears that they go into an almost zombie like state.
KEN	That is terrible
NAG	It is actually not that horrible since we found out, quite by accident, that if a loved one gives them a peck on the cheek that they wake up good as new. We even had one incident where a person's dog licked them in the face and they were awake. It is quite a strange reaction.
KEN	I'll say.

NAME	ACTION
NAG	The thing is once you have gone into this zombie state, you are not very anxious about taking this product any longer. We have found that those in the zombie like state had accelerated sleep onset. This seems driven in part by higher than typical concentrations of potion x; however, we have seen patients that did not have as accelerated a sleep onset who had just as high of concentrations as those that did have accelerated sleep onset. This made us suspect that there could be patient characteristics that impact the maximum effect that they could receive from potion x. Thus, our research goal is to find what factors cause higher concentrations and what additional factors effect sleep onset .
KEN	This all sounds reasonable. Tell me more
NAG	We recruited 24 subjects to participate in a sleep study and we also had 47 concentrations collected from 47 patients 10 minutes after they were dosed. We started doing this in order to know if we needed to call a loved one to give them a quick smooch on the cheek.
KEN	Can you describe the design and analysis in a little more detail?
NAG	Certainly your honor. The sleep study was a randomized parallel group placebo controlled study 3 dose study with a sample size of 6 subjects per treatment group. Ten pharmacokinetic samples were collected over time intravenously. Subject's sleep onset was collected by EEG. Once our physicians started expecting the relationship between the adverse event and pk and/or sleep onset, new patients were first asked to try the medication in our sleep lab. That is how the 48 patient data was obtained. The PK model that we used ended up being a one compartment model with first order absorption. For the sleep onset we related the predicted concentration at 10 minutes with the sleep onset and used an Emax model.
KEN	And what did you find
NAG;	(drily) We found statistically significant relationship between body weight and clearance, a statistically significance between creatinine clearance and clearance. We also found a statistically significant relationship between the Emax value and gender, indicating women had were more effected by higher concentrations than men. We also found that the relationship between age and Emax to be statistically significant and indicated young people had a higher maximum drug effect than those older. This is consistent with what we thought since Aurora Rose and Bianca Snow, both young women, were two of the first that had the zombie life reaction. (More excited) The most interesting finding, however, was the relationship between wearing silk pajamas and sleep onset.
KEN	Silk pajamas?
NAG	Yes, wearing silk pajamas caused a 37% reduction in sleep onset. This was a landmark discovery.
KEN	(concerned) But, but what in the world made you think that wearing sleep pajamas might have an effect. Why was it measured and why would you examine this?
NAG	We at the Seven Miners institute are interested in all components that could impact an individual's ability to sleep. This includes sleep attire, but also accessories and beds.

NAME	ACTION
	We did some of the early work on My Pillow and we also studied the Sleep Number Bed. Anyways, we routinely collected all components that impact sleep in our clinical work.
KEN	(Still concerned) Do you have anything further to add.
NAG	No your honor.
KEN	(Looking at statistician) I could probably make a ruling at this point, but I would like to hear your case.
BRIAN	If it pleases the court, I would like to ask the defendant a few questions.
KEN	Dr. Smith in this court room, I normally ask all the questions but I will allow it. Dr. Chematuri (looking at the pharmacometrician), remember you took an oath to tell the truth.
NAG	Yes, your honor.
BRIAN	The publication that you wrote only discusses the covariates that were statistically significant and does not discuss to any extent those that were non-significant, is that correct?
NAG	That is correct we wanted to focus on the most important findings. There is limited word count in the journal that we submitted so we had to use some judgement here.
BRIAN	I see, but could you tell us here what covariates you examined?
NAG	We were very thorough in our work and wanted to leave no diamond unmined, so to say.
BRIAN	Yes, so how many covariates were examined?
NAG	There were 34 total covariates
BRIAN	So did you examine each covariate on clearance, volume of distribution, and Emax
NAG	That is correct with the exception that the sleep accessory covariates were only investigated on Emax in the PK/PD model.
BRIAN	Lastly, may I ask what was the type of covariates examined?
NAG	We had demographics, baseline characteristics, baseline lab values, a few concomitant medications, and our sleep accessory data.
BRIAN	Sleep accessory data?
NAG	This includes type of bed, type of pillow, type of sheets, and sleepwear.
BRIAN	Your honor, I have no further questions.
KEN	I will give you each some time for a closing statement.
BRIAN	• First, remember that something is statistically significant if the p-value is less than some pre-chosen value. In this case, the authors chose 0.05. This means that the probability of getting the effect that we observed or something more extreme if there is no effect at all is less than 5%. • Thus, suppose that none of the 34 covariates had an impact on either clearance, volume, nor Emax. In addition, for simplicity of argument that all of the covariates are independent of each other, then the expected number of statistically significant findings would be 5.
BRIAN	• Another way to think of this issue is with Bayes theorem. (Show slide) • We know the chance of having a positive result when there is none, we do not, however, know the chance of a positive result when there really is a positive result. This is called the power. I am going to assume it is 80% for each case.

NAME	ACTION
	• As you can see statistical significance is not the same in all cases and I think this proves that the Seven Miners Sleep institute strong recommendation for flannel nightwear when taking Sleep Beautiful is an over interpretation of their results.
KEN	(to pharmacometrician) Dr. Chematuri, any last remarks?
NAG	We just did what was standard practice. You can pick up any journal published and you will see the same practice as we followed. If we are guilty of anything it is that of following conventional scientific practice.
S/J	• "Judge Kowalski, what is your ruling?"
	•
Ken	• I will now present my thoughts on this case and pertinent precedent as published in the statistical and pharmacometrics literature before making my ruling. [present slides] • **Will the defendant please rise:** • I find the defendant **guilty** of inappropriately drawing confirmatory conclusions using "statistically significant" terminology in reporting his exploratory pharmacometric analysis. • The defendant should be aware of the following limitations with standard pharmacometrics practice in attempting to draw confirmatory conclusions. 1. It is difficult to fully prespecify pharmacometric models. Flexibility in postulating models while fitting the data is necessary in order to ensure good fitting models. 2. It is difficult to quantify the strength of evidence from the results of an SCM approach because of the multiplicity of testing. 3. If the defendant desires to make claims of statistical significance a full covariate modeling approach may be better suited for this purpose. ▪ However, prespecification may still be an issue since there can be a fair amount of exploratory base model development before a full model is postulated. Moreover, in practice, the data often cannot support fitting a full model with 50+ covariate effects. • I hereby order the defendant to <u>cease</u> using "statistically significant" terminology when referring to the set of covariate predictors included in the final model for **ALL** future pharmacometrics reports and publications. • The defendant should know that in my own pharmacometrics practice I hold myself to this same standard. In the following publications I advocate avoiding the use of "statistically significant" terminology and reporting of p-values for exploratory pharmacometric analyses. [Present slide citing references] • This case is adjourned! (gavel strike)
NAG	I really do not believe that I am guilty. I appeal to the Supreme Court!! (dramatically)
BRIAN C	*Brian C stands up from his seat- close to the front, pulls out a wig and gavel, and comes to the podium*
KEN	Justice Corrigan I thought I saw you in my courtroom! What are you doing here? Are you spying on me?
BRIAN C	*Brian C can make some pithy response and then proceed with his ruling.*

NAME	ACTION
	I have heard all the evidence of this case and have a similar reaction as the judge. Let me add that in 2019, the American Statistical Association released the editorial "Moving to a World Beyond "p < 0.05."" This editorial makes strong statements pointing out some of the issues of statistical significance and p-vallues. Yet, if you read their conclusion there is not uniform consensus on how to deal with the problem. As Judge Kowalski points out, for this case and situations like this case, avoidance of the term statistical significance and p-values is reasonable. But, in other cases you may also consider the following strategy 1) Be transparent in all of the steps taken in your research and analysis. 2) Avoid using the term statistically significant. If you must use the term, be judicial of your interpretation as to what it signifies. 3) If using p-values, give the exact p-value. It is more useful in judging the strength of evidence. 4) When considering what an effect signifies, consider all the non-study related factors that support or do not support the finding. Be fair minded. *[slide] The key is that all analyses deserve in addition sound scientific judgement. The results of the analysis have to be put into context. It is the results of the analysis along with scientific judgement that leads to conclusions that are meaningful.*
BRIAN C	OVERRULED! And I find the defendant GUILTY!
KEN	Wait a minute! *I* found the defendant guilty already. You just wanted to overrule me, My Lord*! [said with an ATTITUDE]*
BRIAN C	That's right! I've been wanting to do this for 20 years and this is my first chance! But actually… begrudgingly, I guess we agree. SIGH. Then, the ruling stands. The defendant is indeed GUILTY. This court is adjourned *[bang gavel]*
S/J	• Post interview- 30 seconds to 1 min each o Dr. Chemuturi, what do you think of Judge Kowalski's ruling? o Dr. Smith, are you satisfied that justice has been served?
S/J	• Thank the participants.

Ken	• Well, bailiffs, this appears to be the last of our cases for today, but I would like to make some closing remarks to those in the courtroom as well as the gallery. • Stay out of trouble, stay in school,… don't make the same statistical mistakes that your colleagues have! I don't' want to see any of you in my courtroom in the future!
S/J	• The issues and rulings are endorsed by the SxP SIG and are NOT the sole views of Judge Ken!
S/J	• Call up all participants to the stage for a group bow! • Send people off with a mission to spread the word, teach your colleagues, join the SxP • Closing credits