Multilevel models for longitudinal continuous outcomes, part 1—2010 MCH EPI pre-conference training workshop

DON HEDEKER: that you're listening to today and yesterday, will actually be on the website, www.mchcom.com in a matter of weeks or so, so you can hear it again, and again, and again, and again. Also, the material that I'm going through this morning will be the material from the second day. So yesterday when we left off, there's still one lecture remaining on cluster data and that's pertaining to ordinal data. I thought that's sort of optional, if we get to that later on today we will, if not, it's better to move to the Day two material on Longitudinal. And if you're interested in the – some of my written articles and things at my root website, which is just http://tigger.uic.edu/~hedeker/ -- my last name – there is a number of links to several of my websites, one of which is the ML HTML where these PDF files and the datasets reside, but some of my articles, like the one down here that I'm describing, you can get the PDF via the root directory there. So if you want some additional reading material on this you can get that there.

So the material that we're gonna talk about this morning is about multi-level or mixed models of longitudinal continuous data. I've worked on a book with my colleague, Robert Gibbons called *Longitudinal Data Analysis*, and much of the material comes from there. And then before that I had worked on a chapter, *An Introduction to Growth Modeling*, that – and this PDF you can get for free at my website that describes much of the same material. So if you want to read a little bit more, there are some places to go to.

So, once you start getting into longitudinal data, it's a long, long process. I felt I entered this building in – when I started graduate school in the '80's or so and I still have only made my way partially through it because there's always more things to learn, subtleties in the data to guide us. But it could be worse. You could be studying one of these two subjects. So, longitudinal data analysis is not quite as, plain geometry, that would be a little bit worrisome if you're on top floor, I think.

So what are some advantages of longitudinal studies? Why do you see so many studies these days where you are getting repeated observations across subjects? Number one, it economizes on subjects. In as sense, subjects serve as their own control because you can see how a subject performs relative to say, some baselines situation. How they change across time and therefore you can see change relative to some starting point. Okay, so that's a very useful thing is to examine not just how a subject is say at one time point, but how they, at a given time point relative to where they were to assess that change.

As a result of that, statistically what we can do is to extract from the error the between subject variation. So in our models we will have between subject variation, how our subjects are different from each other and within subject variation. Within a subject, how much variation is there. If you only have one time point, you can't partition those two. So the fact that we have repeated measurements from subjects allows us to say how each individual is different from each other.

As a result of this, we can get more efficient estimators than cross-sectional designs with the same number and pattern of the observations. Again, because we can subtract the between subject variation from the within, we can get more statistical power essentially with essentially less data. And importantly, from an interpretational point of view, we can separate aging effects, changes over time within subjects from cohort effects; differences between subjects at baseline. So if you have a cross-sectional study and you notice that there's an age effect, you know, people of different ages respond differently. You don't know that as a person ages that the change happens or just if people of difference ages are different from each other due to generational and age cohort effects.

In longitudinal study, where you measure them across time, you can separate and find out, how are people different as they change, as they age. So that's a nice advantage that the longitudinal studies allow. And throughout all of this we can get information about individual change. Not just about how populations might change, but actually how each and every individual in this sample can change across time.

So we're going to be using mixed or multi-level models to analyze such longitudinal data. Here's some advantages of mixed models on multi-level models for longitudinal. Again, we're gonna explicitly model how individuals change across time. Okay?

If you do analysis via a nova kinds of models. A nova models are all about means. Basically comparison of group means, group means across times. So you get information about how groups change across time, but you get really, no information about how individuals change across time. In our mixed or multi-level models we're going to be finding out about both. How do groups change across time, how do individual change across time. So that additional information can be quite useful.

Now in terms of the repeated measurements, here's really where the multi-level models are very useful. They're much more flexible. Subjects can have different numbers of observations, number one. So there's going to be no assumption that every person is measured, let's say, at all five or six time points in a study. It's a big advantage because in collection of data from humans, you would never get the situation where you measure everybody at every single time point.

What did people do in the old days? They threw away the subjects that were incompletely measured across time. What's the problem with that? You have a biased sub-sample then. People who are measured incompletely across time could be systematically different then those that are completely measured across time.

So in our analysis, we're gonna allow people that have measurements at any time points essentially. That's a huge advantage.

Also, we can treat time in a continuous matter. Now in some studies, you have fixed time point measurements. You got them at baseline, six month, a year, two year. Something like that. So time – the values of time are the same for every subject. But there's other studies in which you might get subjects that's sort of the baseline and then you follow them across the two-year period and some subjects might be measured at three months, some at four months, some at five months. The measurement of the variables across time might be at different – for different subjects.

Well, in our models, we're gonna allow time to be a continuous variable, therefore what that allows us is that the values of time can be different for each and every subject. Okay? So there's no requirement that we have this fixed time point measurement scheme. If we have it, fine, that's great. But we can also be more flexible. Subjects don't have to be measured at the same time point.

Now for the various co-variant structure; what's that about? Well, you've got this dependent variable and you're measuring it across time. So you have the variance of what's the variable at time point one, two, three, four. What is the co-variance? What's the association? And really, most of the models for longitudinal data, but an important aspect of them is how they treat that variance/co-variance structure. Do they assume that the variance is the same across time? Do they assume that all the correlations are the same across time? What are they saying about that?

Well with the mixed models, we can specify many different structures and test them to get a more parsimonious statistical model. Basically what the means is, using up as few parameters as we need. Okay? So that's going to be a nice advantage as well.

Really, in longitudinal, there's essentially you can think of it – there's sort of like two models. There's a model for the mean structure and then there's a model for the variance/co-variance structure. We want to pay attention really to both in order to get two models and also to get good statistical tests.

Now, I'll be talking really, about two level situations in which we got repeated observations within subjects. But as I mentioned yesterday, we could really have this combination of longitudinal and clustered in which we have repeated observations within subjects, subjects within providers, or with states or within something like – some sort of clustering. And we can – as we ran three level models yesterday, we can do the same thing with longitudinal and clustered. And so we can have higher order models.

A little bit later, we're gonna start again with normal outcomes because in a way, that's a bit easier in terms of the computations and the assumptions and blah, blah, blah, but we're gonna also get to dichotomous outcomes as well and of course there is methodology out there for ordinal outcomes, for count outcomes for all the different kinds of outcomes we encounter.

All right, so here's the equation. Here's what's called a two-level model for longitudinal data. So let's go over this. On the left-hand side, what I have is the vector of responses from a given subject, "YI." Okay, so what's a vector? It's just the collection of the repeated observations across time. So if there's say, five time points and they've got five. Well, notice that "YI" has rose "NI," there is "NI" repeat observations. That "I" is very important. What that says is that, each subject can have different numbers of rows, different numbers of time points. Okay, so the size of that repeated measures vector can be five, could be four, could be two, could be one. We're not assuming that every person has the same number of repeat observations by having "NI" by one. So it's the dependent variable vector from individual "I" because here I'm using "I" for individuals or subjects, "J" for repeated observations. And "J" goes from one to "N" sub "I". The number of repeated observations is different from each subject.

So just like yesterday, we were assuming that the number of individuals in a cluster could be different. The analogist thing in longitudinals, we're allowing that the number of repeat observations from a subject are different.

On the right-hand side, the first thing we have is XBeta. So our usual multiple regression. Again, I want to emphasize this point. This model is just an oddmented multiple regression model. If I don't have this part right here, we have ordinary multiple regression. So again, everything that you've learned about regression applies here. In "X" we could have **** variables, we can have continuous variables, we can have interactions. We can have all of those kinds of things that you can do in ordinary multiple regression modeling. That's in "X".

What's new about this model other than random subject effects, now yesterday we only really had one random effect for the cluster because again, the amount of correlation of the data within a cluster is not like it is in longitudinal. Longitudinal, the data within the subjects is more highly correlated. We're gonna do more work as a result of that. And so here what I'm essentially allowing for by having Z Epsilon and allowing Epsilon to be R by 1 is the possibility that there's more than one random effect per subject. We'll see examples of this. To give an idea of this is, you have a longitudinal study, you can allow a subject to have their own starting point. Okay, so what's our measure of the dependent variable at baseline, let's say. That might be one random effect.

And the second random effect is, what's their slope across time? So we were going to allow the possibility, it's that people start off at different levels of severity of illness, let's say. And they change differentially across time by having more than one random effect.

Okay, so that's the big new thing for the model today relative to yesterday. Yesterday, what we had was a reduced version of this where we only had essentially one random effect. Random intercept kind of model. Here we're gonna build upon that because, again, longitudinal, the data are more correlated, there's more work to be done and more interesting things to look at. Yes.

Question: [Inaudible]

Answer: It's sort of, yeah, the fact that individuals are different from each other. You're right. Yesterday we're sort of controlling for the context of the cluster. Today we're going to be kind of controlling for the heterogeneity in subjects. Yes. The subject today is the cluster. Yesterday level two was the cluster, level one was the subjects, today level two is the subjects, level one are the repeated observations. Yes.

Question: Would that have also worked yesterday since we have two different actuations that we expressed ****?

Answer: We could have, yeah. In a way, what I did yesterday was to gloss over to longitudinal aspect by modeling the post-test and using pre as a covariant. When you only have two time points, you can do that. You don't necessarily need to – sorry, the question was: yesterday we had pre and post tests and we kind of – we didn't really treat it longitudinally per se, because what I did yesterday was model post tests and treat as a co-variant pre test. Today instead, what we're going to be thinking about is this really applies when you have more than two time points. If you have just two time points, then doing that sort of pre test/post test analysis like we did yesterday, that's generally reasonable. But when you have more than two time points, you can't really do that because what you are you gonna, you know, you're gonna have post tests, pre tests, you'll have more than two. So that's the idea is that we're gonna expand that notion.

So here we've got situations where we have potentially more than two time points. Any more questions?

Okay, we we've got an augmented multiple regression model, again, subject is the clustered, repeated observations are the clustered observations. So here's some pictures and some equations of what the picture represents. Yesterday, we had a random intercepts model with a cluster as a random effect. Here is the analogous model with the subject as a random effect. And what I've depicted here in this graph, let's say we've got two groups. Those are the central lines here. And here I've labeled them as males and females. Okay, we've got some dependent variable, and here the waves, it's looking as if females start here and go up like that. Males start up here and go like that.

And around each of the group, I've put one subject above and one subject below to represent the different individuals. Now I just put two for each group for the purpose of illustration, but there would be as many as individual lines as there are individuals. So that's the heterogeneity across subjects is the relative to the group line, which is in the middle. The dispersion of the lines above and below the group. So on average, females look like this line, but there's a female here who's above it. There's a female here below it.

On average males look like this line, but here's a male above and here's a male below. So across the males, they average to this, but there's variation in that. Now in this model here, what I've got in terms of regression, I've got a time effect, linear time effect. I've

got a group effect, that is a groups differ when time equals zero and I've got a group by time interaction.

Okay, so this part of the model, the time, the group, the group by time, tells us about these two lines. There's a time effect, there's a group effect, there's a group by time interaction. Additionally to that, I've got the subject effect. Subjects are different than the group trend lines. But they're different in a constant manner. There's only one subject effect. Okay, so the subject difference at a given time point is the same at every time point. Okay.

That's the parallel lines kind of model. It's like the blondie model, parallel lines. They're parallel to the group across time in this random intercept model.

So this is sort of, the picture, this is the schematic of the model and this is the actual model with the Greek parameters and everything. And so we have, time, group, group by time. Now, notice the subgroups for the variables. Here for time, I'm saying that the value of time could be different for each subject and at different time points. Okay, so "I" represent subjects "J" represents time points. To the values of time could vary from subjects and obviously values of time change across time.

Here in this picture, I've got group varied only with subjects. So this would be a situation where subjects are say, randomized, well males and females are going to be the same across time, that's typically not a variable that changes across time, though I mention that it could. And then group by time interaction.

And then the random subject effect. And this is exactly like the statistical model we were treating yesterday except that our regressors here are time, group, group by time. But exactly the same thing. Exactly like what you did yesterday. We've got one random effect and we assume that it's normally distributed with zero mean and variant sigma epsilon. This is telling us how disbursed are the subjects. If every person is the same as their group trend line, this gets small. The between subject variation gets small. If subjects different from their group trend line. This gets large. Subjects are different from each other. So the between subject variance is telling us how much heterogeneity is there across subjects, between subjects, how much variance is there? The **** variance is the within subject variance. Within a subject… within a subject's data, how much variation is there? So between subject variance, within subject variance.

Now this model says that each person has exactly the same slope as their group. How realistic is that? Not too realistic. Okay, so this is a starting point, but this is probably not the ending point in a longitudinal data analysis because it's unlikely that every person's going to have exactly the same slope as their group. All right, it just doesn't work that way, some people go up fast some people go up slower. So typically, a longitudinal -- the random intercept model is not, it's not sort of again where you would end up. You might start of analyzing looking at that, but again it makes this assumption that every person has exactly the same slope as their group, it's not so plausible with real

data. So here is one step up for man, one giant step up for statistics, or something like this.

This useful model, we're now allowing one more random effect. So, a random subject intercepts and a random subject time effect. So now, again, we have the same two average group trends for females and males respectively. And then we've got variation in terms of their intercept. How are subjects different at times zero at the baseline point? And then, how are they different in terms of the slope? Again, for illustrative purposes I've put for each group one subject above and one below, but of course you'd have as many lines as there are subjects.

Now we've got these two sources of variation from subjects their intercept, their slope. And so now, in terms of the random effects we've got two, we've got a bi-variant normal distribution for that distribution. Okay, what's that? Bi-variant normal, we've got two variables here is variance of their intercepts. How different are subjects when time equals zero? When subjects come into the study, how different are they? How much subject heterogeneity is there? That's basically about the dispersion right here. Now, how different are the subject values at time equals zero?

The variance of the slopes tells us how different are there changes across time. Are those slopes very different or not? Realize, the last slide, the random intercept model would say that the slope variation is zero because it said everybody has the same slope. The last slide said there was no variation in slopes at the individual level. This model says, wait a minute, there is. Here's how we're going to estimate it.

So you can see what we can do that is to statistically compare this model to the latter one to test that hypothesis. Do people have different rates of change or not? If they have the same rate of change across time then this parameter would equal zero. And by comparing this model to the previous one, we can do a test of that idea.

Okay, so we've got this, and usually the way it works in longitudinal is that this model is favored relative to the previous one because the last one says that everybody has to same rate of change, and again, often that's not the case. Now, there is one other parameter in this variance/covariance matrix for the random effect, that's the Association, the covariance between the intercept in the slope. Okay so that's this one down here which is sigma epsilon zero epsilon one. Epsilon zero is the intercept, epsilon one is the slope.

Okay. And what that one is about, and what most people kind of ignore that one, but it's kind of an interesting parameter. It's saying, is the person's slope associated with where they start out. In a way it's like can you predict the rate of change based on where they are at times zero because that's what it's about, the association between their slope in their time zero value.

So the next one we want a little more attention to and maybe because it's you know again, it has some interesting qualities to it. So this is a very useful model in longitudinal and

we'll be applying this sort of model to data sets again, it's really not a huge step over what we've dealt with yesterday, the random intercept model but it is a very useful step. Yes.

Question: I have a question. And here you assume the group is not changed, it's fixed. Is that right? What I am studying now is that even the group is changing across the time. I'm looking at the impact of feeding patterns on infant growth, so the feeding pattern changes from time point one to time point two. So how can this model and –

So the question is, here in this model, I've got that group, the subject is in the same group across all time points, but in other studies you might have were subjects change group membership across time. That would be a situation where the group is what's called a time varying covariance. And yes, we're going to get to an examples of where you have time varying covariance in your model and yes, this model can definitely accommodate that. The picture wouldn't look the way I've drawn it here, but yeah you could have we're group is different across time. They also have that in what's called crossover studies were sometimes people are observed under one condition for a certain number of time points and then observed in a different condition for different number of time points. That's another situation where group membership changes across time. And yes that can definitely be accommodated.

Okay, any other questions? Yes.

Question: **** examples of groupings. You mentioned gender, right?

Answer: Yeah, sure, sure. In clinical trials, for example, where I've done some of my work is people are randomized to either get a control situation were some intervention situation across time. Right, so that would be a natural thing where you've got group A, group B, subjects that are in this. So it might be you know like yesterday in the educational example, the subjects were random or the schools really were randomized to either get control situation or this enhanced knowledge saying and then they were measured across time. And then he who your interest there was to find out is the rate of change across time different? Is it the case that the people who get the enhanced knowledge, do they actually do better across time than the people who don't get that?

So in clinical trials it's very common that they have two or more groups so that it stays in that group and their measured more longitudinally across time to see the differential change in the groups across time to assess whether or not the group is a beneficial thing or not.

Okay any other question?

Okay, well, here is, you know, that random intercept and trend model you see in a lot of places. I even saw it as an example in the Onion. Do you guys get the Onion down here? I saw this, I couldn't believe it. It's like wow! That's a random intercept and trend model. Okay, it said that the rise in teen sexual activity comes as a surprise to area teens. And it shows in the graph, you know, normal teens going up like this and then there is this ****

he’s at zero. And he stays at zero, right? And he's different in terms of his intercept, he's differ then in terms of his slope. It's terrific. It's random intercept and trend model. I'm sure the editors at the Onion didn't think of it in those terms but you know it's like wow that's what it is. And so there are a lot of examples of it.

Now I'm going to be showing you as well, the way to write this out as a multilevel and hierarchal linear model. And the basic idea is that in the level I core within subjects models you put in any level I variable. So for example, time is a level I variable, it changes values across time. Now in the example I just had, group was a level II variable, it only changed values from subject to subject, it didn't change values across time. Now in the example you asked about, group might be a level one variable, a time varying variable. So this is a general schematic, the idea is that in level one you put in all the time varying covariates, level II you consider the subject varying covariates, things that don't change across time. And so ultimately what you getting your model and sometimes this was, in the early days, this used to be called the slopes is outcome model, you get level I coefficients for the time varying covariance, level II for the time invariant and then cross level interactions.

So for example, in the simple slide we've been considering, time was a level I variable, group was a level II variable, and group by time was a cross level interaction, it involves group, which is subject level, level II, time level I, that interaction is a cross level interaction. And again, anything that's at level one could potentially be varying at the subject level. And thus far in the slide what we've considered is random intercept and random time effect, but really the effect of any time varying covariant could be random at the subject level. Okay, so that's the general schematic, we'll see specific examples of the way we were writing it that way.

Now here's the way I like to write it. I always think in matrix terms. I know most people don't like matrices, but I’m always… I’m kind of struck by that because people don't like matrices, but a lot of people watch Hollywood Squares or at least they used to. And what's Hollywood Squares but a 3 x 3 matrix of washed up celebrities, right. So I don't understand; I get a disconnect there.

Okay here, the way I've written it, is I've got on the top essentially the Y=XBeta part of things. Really, the multiple regression part of things I have under the top part. So "Y" is the dependent variable vector. Again, the number of rows here is going to be the number of time points for a given subject. And then I've got these regressors that I'm considering, time, group, group by time. Now if you write it this way you can easily see that the values of time can be very different for every subject because time is just regressor. Think about this, but if you have blood pressure as an independent variable those values are different from subject to subject. Time is just like that, it's just a regressor here. So the values of time can be different for each and every subject. The number of time points can change, that's the number of rows here, and the actual values of time could be different too. Yes question?

Question: Does each subject, start at their own zero or do you have one zero point and everybody's kind of in the same time?

Answer: Right. The question is about the zero point. Does everybody start at baseline or do different -- it depends on the study. In some studies you know people come in baseline is baseline, that's times zero. In other studies it's more of an observational studies and then let you consider times zero to be could be different from subject to subject. It kind of depends on the given subject and the aims of the studies. But you know, in clinical trials sorts of things, which I guess maybe not -- you know times zero is sort of like it's fixed in time. But again observationally, you might think of time zero as a given age for a given you know something like that and fix that as times zero. So it kind of varies.

Okay, so this is the regression part of thing. And down, on the bottom I've got the random part of things; the random effects, the intercept and time effect, and then the random errors. And so the random effects are all about the between subject variation, how are subjects different in terms of their starting point, in terms of their track. The within subject variance, the error variance, within a subject how much variation is there. So between subject variance within subject variance. And again as you see time is just here a predictor so it can be measured in years, months and it could differ from subject to subject. So there is great flexibility here with how we can treat time especially relative to if you've done any work with Anover models, **** variance models, time variance is treated like a factor. And everybody's got to have you know four time points. If you don't have four time points and you're not in the analysis.

Now this slides a little bit much but let me give you an idea of this. This says in this model with two random effects, random intercept and time, what are we saying about the variance covariance matrix? That is, of the dependent variable, whether we saying about the variance across time and the covariance of the repeated time points? So here in this slide I'm saying, let's say we get a random intercept and trend model and here I'm just quoting time sequentially in zero, one, two, four say three time points, here's the four of variance-covariance matrix for the three repeated measurements. Now the first thing you notice is this ugly right. I want to convince you that ugly as beautiful. Why? It's complicated. Right? But notice, on the diagonal what we have is the variance of the three repeating measures, it changes. In most studies, the variance of the dependent variable is not constant across time, it often grows across time. People get more disparate as you go across time points. So this model is allowing that. A simpler random intercept model alternatively, would say the variance of the dependent variable is the same every single time point. So again, that's not a very plausible structure usually.

So again, this one giant -- this one little step of having a random trend model is allowing us to say something a bit more realistic are the variance and also for the covariances is in the repeated measures. Covariances are like this. If you have two time points that are close the covariation is usually much stronger than if two time points were far apart.

It's not rocket science, if you measure people at one for zero and one to one, the Association is going to be stronger than one to zero and one year later. So again you want a model that says that the covariances is far stronger if they're closer in time and weaker if they're further away in time. And this model accommodates that this simple random intercept model does not. It says that all of those covariances are the same. So the random intercept model that we used yesterday can be faulted on several counts in longitudinal. Number one, it says every person has the same slope across time relative to their group. Okay, so that's **** realistic. As a result of that, it's also saying that the variance of the dependent variable is exactly the same at every single time point and all of the covariances are exactly the same. Now none of those things are generally seen as real data. And so moving up to a random intercept and random trend model of is an advantageous step for longitudinal.

Now, more general models than what I'll talk about today, also allow things like auto correlated errors, so there's additional work that can be done in this realm. We'll be dealing mostly with this sort of model today.

All right, so you're probably say like this, "I've followed you up to a point where you had a picture of oregano." But I would argue this, you see on the blackboard there this person's got integrals, it's got square roots it's got all kinds of complicated math, you haven't seen yet. Just kidding.

Okay, so let's get to an example. Again I apologize if this example is not necessarily in your substitutive research area, but it's a good teaching example again. And one I've used for a while and it has some nice aspects to it that highlights the differences in the modeling techniques. It's an old study in psychiatry from **** and Associates, and they've looked at the relationship between Imipramine and Desipramine plasma levels and response from depression in 66 depressed in-patients. So, what's Imipramine and what's Desipramine? Those are classic tricyclic antidepressants. And actually, the way it works, is that a person was given Imipramine, that's the drug, and then in the body it metabolizes to Desipramine, both of those chemicals are known to have antidepressant properties, so they measured both in the subjects.

Now this study was carried out in 1977 and Imipramine was kind of like the classic antidepressant at that time. If this study was carried out today what would they be given instead? Prozac, right. Prozac. So the Imipramine has sort of side effects like dryness of mouth and other things that Prozac seems to handle better. So these days Imipramine is not used as commonly as it was back then.

Okay, now there are 66 subjects, not a terrible -- for an epi study that's nothing, right. But in psychiatry that's actually a pretty reasonable sample. They were classified in terms of endogenous, non-endogenous standards. Okay, does everybody know what that's all about?

Question: Self-selecting or self-selected versus non?

Answer: No. Not exactly. It is --

Question: Is it if there's an external basis for the depression, like a loss in the family or something?

Answer: Yes. Exactly that's exactly right. The way it works is that if the depression was precipitated by some tragic life event, right, like a spouse dies, like White Sox lose the World Series, something of that nature then that would be called a non-endogenous depression, or reactive depression. Conversely if there is no precipitating life event that is classified as an endogenous depression. And the thought was that the drugs might work better for endogenous depression because that's kind of more of a chemically, you know, organic sort of development. It has nothing to do with tragic life events, let's say. So they separated the subjects in terms of those two diagnostic groups and are interested in knowing if the response across time are different for those two groups.

So here's the study design, there was a one-week drug washout period in which the Hamilton Depression Scale, and this is kind of a common relating scale for measuring depression was obtained. And the baseline at week zero people were diagnosed in terms of being either endogenous or non-endogenous. Not after the first week, people were given the drug Imipramine and they were given it at four time points, at each time point at the end of the week they were measured at how much Imipramine was floating around in the body as well as how much Desipramine. Okay, so what happened is that people would take pills, but that gets metabolized to Desipramine. Soon your body you would have both of these things floating around and they both have antidepressant properties so they would measure them both.

So what's interesting about this study design is that first of all, we've got a longitudinal -- this is what our dependent variables are going to be, the depression rating scale. We've got covariants, like diagnosis, diagnosis doesn't change across time. People in the same within a depressive episode they're either endogenous or non-endogenous. And then we've got the time varying covariance of Imipramine and Desipramine. And notice the sample size at the bottom time point, by the way, that should really be large and interpretive of my notation that it refers to the number of subjects. It's a pretty well-designed study, 66 depressed inpatients, but not at any of time point do you have all 66. So, incomplete data across time.

Okay, so here's the outcome variable to the Hamilton Depression Scores. The independent variables will be diagnosis, Imipramine and Desipramine for diagnosis we've just have the two groups so we can use a simple dummy code to represent the two groups. For Imipramine, that's the drug that was given, 225 mg per day in those last four weeks. Desipramine was a metabolite of Imipramine that was also measured.

Okay, so here are some observed descriptive statistics for this data set. Here's two groups, endogenous/non-endogenous, again, it's small and it really should be large at how many subjects there are at different time points, here is the means. Incidentally, higher values indicate more depression, lower values indicate less depression, so what they were

hoping to see is that the means go down across time and if you look at these means across time they do. And on the bottom we have this pool standard deviation. Okay, so this is the standard deviation for the dependent variable across the six time points. Notice the values of it. They grow. Why? In the study, people come in and there are differing levels of severity of illness, across time some improve in some don't improve. So it's common sense that you're going to see increased variance across time. So you want a statistical model that can accommodate that. Again, a random intercept model would not accommodate that. So if it's a random intercept and time model it can't.

Now in terms of endogenous, non-endogenous, do we see a differential pattern across time? Does it look like the drug works better for endogenous or non-endogenous? Well if you look at the means you can see that at every single time point the endogenous group there's a little bit higher than the non-endogenous group. But the mean difference doesn't seem to change much really across the different time points, it's about two or three points at every single time point. And so it looks like both groups are of going down across time and it doesn't necessarily look like the rates of change across time for the two groups are any different. So were gonna test that but just the eyeball statistical test that kind of suggests that both groups are going down, there doesn't seem to be a great deal of change in their slopes across time.

Now here is a measure of association for the Hamilton across the six time points. And what I put in bold are the **** **** correlations and above the diagonal are the paralyzed deleted correlations. Okay, what's the difference? When you have six variables like this and you go to a say, Proc core or whatever, what it will do by default is, list wise delete it. It will throw out any person that doesn’t have one of the six. And then, you this study you'd be left with 46 subjects. So think about this, any statistical analysis that would throw out subjects with incomplete data, and you could be throwing out a third of your data set here. You know if you only have 66, you don't want to lose 20, that's rather a lot. So analyzing only the completers has two major problems with it. Number one, you're going to have less statistical power, number two, the question of bias; the people who are completely measured across time can be systematically different than those who are a incompletely measured across time. Okay, so we don't want to do that. And we’re not to do that day but this will give you an idea of what would happen should we go with an analysis that throws out incomplete subjects. We would only be left with the 46 out of the 66.

Now in terms of the correlational structure, what you see here is what you often see in longitudinal. First of all, as you go away from the diagonal, you are going to correlations that are further away in time. And so let's say, week zero and week two, associations is .22. Week zero and Week one is .5, about. It’s stronger. If the measurements are closer in time, they’re more highly correlated than if they are further way in time. Again, it’s not like rocket science. It makes a lot of sense that measurements are going to be closer in time or they’re be more correlated than those that are further apart. That’s exactly what you see here. In other words that as the lag increases, you know, if the lag is five weeks apart, four weeks apart, the association is less than if it’s one week apart.

Additionally, within a lag, that is, within a sub diagonal, it seems like the association might be increasing. Okay? That's maybe not so dominant, but there's a suggestion of that.

Now if you look at the pair-wise, which – I'm sorry, the peer-wise – and the way the pair-wise are computed is for any given correlation, the people who are – this is based on, are people who have week two and week five. And so that's why the number at any one of these could be between 46 and 66. So, the number of subjects used for a given correlation depends only on how many subjects have those two time points. In bold, it depends on how many subjects have those six time points. Okay, so the data requirements for the pair wise, you know they're kind of less.

It doesn't look like there's anything systematically different between that. What that suggests is that maybe subjects are incompletely measured are not systematically different than those who are completely – again, this afternoon, we're actually going to get into a little bit of testing of that idea, but for the moment it does look too bad with relation to that.

Okay, but so far, these two slides I think have been pretty daunting evidence that a random intercept model is not a good model for these data. Again, the random intercept model would say that the variance across time is all the same. The last line show that the variance increases across time. The random intercept model would say all these correlations are the same. They don't look to be the same, they differ in a very systematic way and so random intercept in time model makes much more sense here.

Okay. Now here's what's called Spaghetti Plots. Some of you might have had spaghetti last night here's the data analysis version of that. What's a Spaghetti Plot? It's a plot of all the data, all the dependent variable values for a given subject. So for example, this guy in blue, or this person in blue, they started off here at 30 and went up to like 40 and then went down. What this gives you a flavor of is the gross pattern of the data across time and the variation across time. So by and large what you see is that well in general the pattern of the data is that it is going down across time. So on average people are improving. Not for everybody, notice these couple of subjects out here. They're actually getting worse.

Okay, when you do an analysis just that means, you don't observe these things. When you focus on the individual as well you get information, not just about the means, but individual. There are some subjects out here who do worse there are some subjects here who are doing rather remarkably well, getting down to zero. You know that's like Bobby Mc Ferret, "Don't Worry Be Happy," ****. Like who's like that?

Notice also the increasing variance across time; just look at the range. Twenty-five to 35 at baseline, at the end of the study it's like zero to 35. We had that standard deviation at every time point it went up, okay fine. Here's the visual representation of that. The variance at every time point is growing. Why? Again, some people do great; some people don't do as well.

So Spaghetti Plots are kind of useful for examining getting a feel for the data and it also can be useful for finding any outliers, some really weird values because again if you do the analysis you won't necessarily know that some subjects at this time point has an impossible value. You do this and you can say, oh wow, there is somebody like really weird, what's that data point all about.

Okay let's run some analyses. We're gonna start with this random intercept in trend model and this first model is akin to the mixed model version of the simple linear regression model. Now in the linear regression presumably the first model you learn with simple linear regression. What simple linear regression? 1Y, 1X right. You've got one dependent variable, you've got one X variable, you can draw in that little line, Beta Knot, thats the intercept, Beta 1 is the slope. Right? We've got the analogous version of that here in longitudinal. On the top part, I've got the intercepted slope for the effect of time. So Beta Knot here is what's the mean of the Hamilton when time equals zero, Beta one is how does that mean of the Hamilton change with every increment of time? Here I've quoted time simply as zero, one, two, three, four, five for the six time points. So what we're gonna get with Beta one is what's the per week change in the mean of the Hamilton? How much does the Hamilton mean change with every week? That's what Beta one will address. Zero will say, what's the mean of the Hamilton when time equals zero, the intercept?

So Beta Knot and Beta one will be about the population. Where does the population start out? How do they change across time? Down below here I have at the individual level the intercept and time effect. How does a given subject start out? How does a given subject vary across time? So the subject intercept and the subject slope. Now the epsilon are subject parameters are deviations relative to the over all population. So for example, if the subject has exactly the same starting point as the population epsilon zero would-be zero. These are deviations. Epsilon zero deviations is relative to Beta Knot. There is subject deviations relative to the population trend line.

Okay, so this is simple linear mixed regression you might call it because we have information about how the population is changing, how the individuals are changing. But we've only got one covariant here and that's time. But we have information about the slope at the population and at the subject level. Now if we write this is a multilevel model, here's how it looks. We've only got one regressor, one independent variable, one explanatory variable whatever you want to call it, one covariant, one X, okay it's time. Time changes values across time, it's a level I or a within subjects variable. It changes values within subjects. So I've got that at level I, so I'm modeling the Hamilton from subject "I" at time point "J." in terms of an intercept and a slope. The intercept and slope come down at level II between subjects to be modeled in terms of any subject level variables. Now in his first example I have none. So all I'm saying with this model is that a person's intercept is made up of the population intercept, Beta Knot, and their own deviation to that population intercept.

A person's weekly change or slope is made up of the population's slope in their own deviation relative to that. Okay, so this is a simple random intercept and trend model with just time as a predictor. The random intercept version of this would be like that. And what that would say is every person has the same slope. Again, that doesn't -- that doesn't make sense, so a lot of that second random effect here is rather useful saying every person has a different slope. So this is the way to write it as a multilevel model, okay, in running yet we don't necessarily have to do this, but it is helpful for us to understand what we're doing to write it this way.

Here are the results, and in a little bit will show how to do this. We can do this with Proc Mix because all we have really is just one regressor and we've got two random effects so here we've got the estimate of the intercept and the slope, the maximum "Y" period estimate, the standard error, the ratio of the estimate to the slope that's the "Z" or "T" value that's assessed that indicates -- and then the "P" values.

So first of all let's look at it row by row. The estimate is 23 1/2. So the population is starting off with depression levels at about 23 1/2. That's pretty depressed. That's highly significant. Is that test, that first test, is that of any interest? This is a depression study, what is it testing? You're testing that Beta Knot is equal to zero. Could that ever be? No. It's impossible. Why would -- I mean again, you're talking about a Bobby **** territory if you have a Hamilton zero. And if you have a depression study, there is no way that the average at the start is going to be zero. It makes -- so it's significant, but who cares. So oftentimes in regression modeling testing the intercept is of no consequence really in this case.

But the next one, the slope, that's basically saying, does the mean of the Hamilton change across time? And literally here, I'm considering linear change across time. We'll get to nonlinear change but for the moment I'm saying does the mean that the Hamilton change linearly across week. We get the estimate of -2 1/3 about what each week people are improving by about 2 1/3 points. Okay the mean is changing by -2.38. That's highly significant. What that means is that we can reject the horizontal line in favor of a line that goes down by over two points per week. So on average people are improving. I should point out, we haven't yet considered a diagnostic group, we haven't yet considered the drug levels. Okay. So we're starting off just addressing, is there change across time? That's all. So there is statistically significant change across time.

Now, here we have our random effect variance for the intercept, the slope, and the covariance. Now first let's look at the intercept variance, it's 12.63. What is that telling us? What's the square root of 12.63? It's somewhere between three and four, right? Okay, let's call it 3 1/2 let's make it simple. Well that's approximate to what we wanted, 3 1/2. What I'm doing is this. The intercept is estimated to be 23.58, this is about how much variation there is in the intercept. So let's think about getting a 95% confidence limit for where the intercept resides, it would be 23.58 plus or minus approximately 2 or 1.96 times the standard deviation. So the standard deviation is about 3 1/2, if we multiply that by two we get seven. So on average people are 27.5, but the 95% upper level would be about 30, lower would be about 16. Okay so, on average 23, but there's variation

between 16 and 30 thereabouts. That's what this is helping us to identify. Is there a significant variation at times zero people are not all starting at 23 1/2. You know when you see that intercept estimate of 23 1/2 you're like, oh everybody's at 23 1/2. No. Not everybody is at 23 1/2. On average everybody's at 23 1/2. But there's a lot of variation and this intercept variance is telling us about that.

Likewise, the covariance. On average, you know you see this, like whoa. 2 1/3, everybody's improving, no, everyone's not improved. Or you don't know that, let's put it that way. All you know is that on average people are improving, you don't know that everybody's improving. I mean you're kind of drawn to that. You think about oh the drug works that it's great. On average it works. Here's the variance of the slope. Two. What's the square root of two? To put it on a standard deviation metrics it's like 1.4, or something like that. That's a number I should know, the square root of two, that comes up a lot in statistics. Let's say it's 1.4, I think it's something like that. So again, I'll multiply it by 2.8. Okay so some people, -2 plus 2.8, they're rather down in the positive region. In other words, they're increasing. -2.8 minus 2.8 minus 5, some people are having tremendous improvement across time. Okay so again, this is -- you know you're seeing ****, people are improving, this says well wait a minute, there is range of net improvement, there's variation. Some people are improving a lot, some people are improving not so much. That's what the variance reflects.

Now, we'll get estimates of those individual random effects that will also bear this out but when you just see this you can also get a feeling for that.

Now here's the covariance, it's negative, and expressed as a correlation it works out to be -.3. Okay and I say most people kind of ignore that, but what that's suggesting is what? It's saying a negative relationship between an intercept and the slope, in other words, if the intercept is higher, the slope is more negative. Now what could be producing that, higher intercept, more negative slope?

Right. So they're higher. And that's associated with having more of a negative slope. More depressed, more improvement. Right. That's what that's suggesting. What could be causing that though? What? The drug? Maybe. Ever heard of anything called regression toward the mean? Right. If you have a sort of abnormally high value at one time point, you're next time point could be lower just for regression toward the mean. So that could be causing this negative association.

Another thing, as you mentioned is the drug. In psychiatry, there is this notion that the drugs don't work as well for people who are mildly depressed relative to people who are really depressed. This would bear that out as well. My here's another thing to consider. If you start off at 30 you can go all the way down to zero. If you start out at 15 you can go all the way down to zero. Who's gonna have more of a negative flow? So the rating scale, the floor effect could also be having something to do with this negative association. With the data we cannot necessarily separate those three ideas, regression toward the mean, better you know work for the drug or floor effect. Really, all of them could be

yielding this sort of negative association. So we don't necessarily want to make too much of that.

Now as mentioned yesterday, for the variance/covariance parameters the wall test, taking the estimate dividing by a standard error, that's kind of a dubious test so instead of that we prefer likelihood to ratio test for the variance/covariance parameters. Now here I've done this right here. With this statistic, what I've done is compare this model to a simpler model that didn't have the variance in the slope and therefore didn't have the covariance between the intercept and the slope. It didn't have these two parameters; the simple random intercept model. If I fit that model and compare it to this model, I get a chi square of 66 and two degrees of freedom. That's highly significant. And that's rejecting this null hypothesis. The null hypothesis is saying that there is no variation in the slopes. Everybody has the same slope. So we can easily reject that statistically. I mean, common sense we reject it because it's saying everybody has the same rate of change across time. That doesn't make much sense.

Here's the statistical test of that. By comparison of this model to a simpler model that doesn't have the random slope, we get a highly significant improvement in model fit for this model. Yes?

Question: Can you review how you came to that very last **** again?

Answer: Right. What I did was I compared this model and this log likelihood value to a simpler model at which I did not allow slopes to vary from subject to subject. In other words, I had a model without this. And so will see -- will have SAS syntax to show exactly how one runs those two models. Here I didn't run that first model and show you the results because again, it's logically implausible, generally. And so what I'm showing is a result of that test. But it's just comparing the model that I have here to a model that says, there is only a random intercept and that's it. And the random intercept model is resoundingly rejected in favor of this model.

Okay, so again I said it was kind of like one small step, but you can see the improvement by that small step by saying the subjects vary in terms of their slopes, that fits a lot better than saying the subjects have the same slope.

Another point is that in some of the software you sometimes see these walls statistics squared and then expressed as a chi-square statistic. Actually where you see that mostly is with logistic regression when they take that ratio and they squared and they represent the test as a chi-square. The two, it's the same thing; it's just a re-expression of the same test statistic.

Now let's take a look at the estimates we got in now while they fit the observed quantities. Here in red, what I have are the observed means across the six time points. And the green line is what was estimated for the means across the six time points. How did I get those? Let's go back to the estimates. Based on this, what is the estimate of the Hamilton mean when time equals zero? 23.58, great. Perfect. What's the estimate of the

Hamilton mean when time equals one? It's the sum of those two, right? And then at week two is the intercept +2 times the slope, etc. etc. right. Those first estimates tell us basically what's the Hamilton across time, the mean of the Hamilton. So that's exactly what I calculated here in this row here, what's the Hamilton mean and it's only based on the intercept and the slope estimate.

And here's the observed quantities. And when I saw this I felt like Gregor Amendol. Does anybody remember Amendol? He postulated those laws and then it was found out that he actually, I think, had thrown some of the data away because it looked like it was -- I couldn't believe the fit. It's like perfect linear fit of the means. Right is quite amazing that the model, the Hamilton mean is really changing in a linear manner across time. So quite remarkable how well the model picks that up okay. But really I didn't throw out any of the data. Maybe **** did I don't know. But I didn't.

Now here's a variance/covariance matrix. That's the observed variance covariance matrix. Now on the diagonal, what you have are the variance across the six time points and on the off angle you have the covariance. So for example, the 602 is a covariation between week zero and week five. And now what I've done is also to calculate what's the estimated variance/covariance matrix based on my model. Well, that's this form here, this kind of ugly form, but it involves only, let's see, four parameters. Error variance is one and then in the variance/covariance random effects I have intercept variants, slope variance, intercept sloped covariance. So four parameters. Now how many elements are in this variance/covariance matrix? It's a 6 x 6 symmetric matrix. Anybody know the rule? It's 6x7 divided by two, so that's 21. There’s 21 elements here in this variance covariance matrix, I'm fitting it with for parameters. How well am I doing? Not bad, not bad. You can see that the model is picking up that increased variance across time the diminished variants as we go away from the diagonal, so the covariance, the increase covariation with in a lag. With only four, it's doing a pretty reasonable job.

Now on the bottom I have with the simpler random intercept model estimated for its two variances, 16 and 19. Okay we saw that chi-squared statistic of 66 that rejected this model in favor of this model. Well here's why. What the random intercept model would be saying about the variance, the diagonal, is that it's the sum of these two numbers. It's like 35.2. Now look at those variances, are they equal to 35.2? On average they are, if you average those variances up and get an average it's like 35.2. But they differ systematically from that. At earlier time points less variance, increasing variance across time. So again, this simpler random intercept model which says that the variances are the same across time, you know it doesn't fit very well. Likewise, the co-variances. The co-variances in the random intercept model well it would say all of those co-variances are equal to 16. Again, they're not they've differ systematically from that in that in the way that they are smaller as you go away from the diagonal and within a lag they tend to increase.

So we got that test statistic of 66 with two degrees of freedom and we resoundly rejected the random intercept model. When you look at the estimates and what it would imply about the variance/covariance structure, you know you can really see why it was rejected.

It doesn't fit the data. And going up to this model with these two more parameters because really that's what we're talking about here. This model had two, our model over here now has four; a lot, lot better fit by adding that random trend in the model.

Now here are the random effect estimates. I'll show you in a bit how to get those out. Yesterday there was some question about how you get the random effect cluster estimates out. We didn't show you then the syntax, today we will because longitudinal often, it's a little bit more normal to look at those things. Here shows each person, their intercept and their estimated slope. Now what I've done here is to put in dash lines the average of the intercept and the average of the slope. And each dot represents one of the 66 subjects, so for example, this subject right here, let's say their intercept is 25, their intercept deviation relative to the average intercept is like about 1 1/2, their slope is -5 or so, so this is a subject who has the greatest change relative to the average change. Here's the average change.

Now if you look at this, you'll find some interesting subjects. These subjects out here first of all. These are subjects that are starting off pretty close to the average in terms of their depression level. But look at their slope values. Theirs are actually positive. These are people that are not doing well you know they're actually getting higher values across time. So what this can be useful for is, matter of fact, who are these subjects? Why are they doing so poorly? Also, these subjects appear, these are kind of interesting individuals. They're starting off at the highest level of depression, they have very negative slopes. These are people that are really improving a lot across time. They're getting kind of interesting subjects; you might want to take a look at.

So looking at these estimated random effects can be useful in exploratory kinds of ways of try to help you understand the data a little bit more and suggest some other analyses. Now here's another thing, remember when we saw that negative covariance between the intercept and the slope? And I said it could be due to regression toward the mean, could be because the drugs worked better for more depressed subjects, it could be due to a floor effect of measurement. Now if I take these two subjects out, do you think there is an association between the intercept and slope? Does it look like there's any association between this and this if I take these two points out? Not really. Not really.

So in this sense that negative covariance was in some part at least to those two subjects who they started out really depressed, because they were quite a bit above their peers, and they have really good responses across time. So again when you do this it kind of informs that covariance as well. So you shouldn't make much of it because it's really -- it's kind of being driven by a few subjects.

Now let's look at how you run these models actually. And we're gonna get into doing that. Here is some SAS code, Proc mixed that we will use just a bit. Let me explain the code that we have here. The data set that we're reading in that you have available is called resbe.dat. Now it also says the resbe data Hamilton Depression Rating Scores Across Time. Again, we're going to have to modify the in file statement because we have it in a different location than what I currently have there. The variables I'm reading in ID,

is the subject ID, HamD is the depression level. So one thing to realize is in this way it's in that long format. There's as many rows as the subject has data. So repeated measures don't go across as they do if you're using something like M plus. Other software where the repeated measures are considered different variables. Here the repeated measures are different lines.

ID the Hamilton Depression, again, I've got an intercept in there for historical reasons, but for some other programs I needed that for SAS you don't need that so just ignore that. Then the value of week. Week goes from zero to 5, so zero, one, two, three, four, five. Endogenous. We're going to look at that in just a moment or if no endogenous, endogenous, status, endogenous by week, the product of endogenous by week. It's the interaction term.

Now the random intercept model that I didn't show the results of, but we used in the tests can be estimated like this. In Proc Mixed a specified maximum **** estimation, I'm here asking for tests of the various covariate parameters when one gets the output I'll say a little bit more about that.

Class, I mentioned ID, model statement, Hamilton Depression is equal to week because I only have week as a regressor at this point. And again, the solution. Again in Proc Mixed by default you will not get estimates of the datas, right, the regression coefficiency; you have to ask for it. So that's what solution is all about. Give me the solution, why would I be writing this if I didn't want the solution. I don't know.

Then the random statement. Here I'm asking for a random intercept model first. Subject equals ID, type equals unstructured and "G". Okay, let me say little bit about that. What SAS calls the variance/covariance matrix of the random effects, it calls it the "G" matrix. I'm not sure why they call it "G", but that's what they call it. So here I'm asking, printout that "G" matrix. Now by default when you run Proc Mixed if you have more than one random effect it assumes that they are uncorrelated. It assumes what would be called type equals diagonal. There are no -- there is no association between your multiple random effects. Now strictly speaking, in this first model, I only have one random effect. So there is no multiple random effects, there's no covariance so I don't need this type equals unstructured. Type equals unstructured basically tells it, estimate those covariances between the multiple random effects. At this point I don't need it, here I will but I think it's useful practice to always put on the random statement type equals unstructured. Why? If you don't put this and you have more than one random effect, you're going to be estimating a model in which you assume, for example, that the person's intercept and their slope are uncorrelated. It won't estimate the covariances by default. Now for other kinds of mixed models, that might be appropriate. For longitudinal, I can't see why a priori you would assume that.

So SAS has this default that I think is a rather bad choice for longitudinal because it by default, it is going to assume that you have more than one random effect that they are uncorrelated. I guess it does it for computational reasons, I don't know for sure. But I would always, if you're running Proc Mixed and you've got more than one random effect,

always have type equals unstructured. That will allow estimation of all variances of the random effects as well as the covariances.

So for example, we have that negative covariance between the intercept and slope. If you don't put this it's not been estimate that and you're gonna be saying that a person starting point and their slope are totally independent. Why would you want to make that assumption? There is no good reason.

Okay, so that's the syntax for the random intercept model and then for the random intercept and time model, I simply add week on to random statement and that's all you got to do. So again, I want to emphasize that type equal unstructured really is not essential here because you only have one random effect, but here it is. Now the other thing that I've added here is "G" core. What that does is the random effect variants/covariance matrix it expresses it as a correlation matrix as well. So you'll get it in covariance form as well as correlation form, that's all that that's about.

Any questions about the syntax? Yes.

Question: I didn't see the variable for the endogenous.

Answer: Right. The question was, I didn't see the variable for the endogenous. I haven't yet considered that yet in the problem. So what I'm showing here is just the models I fit thus far in terms of the PDFs. Okay, so thus far the results I've shown you were for this analysis. And I then did a likelihood to ratio comparison test comparing this to this. Yes.

Question: I'm a little confused. Time refers to the structure?

Answer: Yes.

Question: so is that truly a compound symmetry structure?

Answer: No, no. The question was, about this type structure. Type equals unstructured. All this is saying is that, again, it really pertains if you have more than one random effect. And then it says like down here, it's saying, let the covariance be estimated. In other words it says that this two by two variance covariance matrix is unstructured.

Question: So if you were estimating, say a compound symmetry structure, you wouldn't have the CS there, how to address it, you would have AR?

Answer: Let me say with regards to those kinds of structures the question is about these auto regressive or compounds symmetry structure. You would generally not have that for the random effects, you might have that for the error variance/covariance because for the random effects you know there is no reason to think that the intercept and time would follow, let's say an error one structure. You're right. Or you're talking about more about the error variance/covariance matrix where you might have an AR1 or a compound

symmetry or something of that nature or for the random part of things she would never really want to have anything other than this.

Okay, well this then shows the analysis that which also produces the random effect estimates. Okay so the random effect estimates I used specifically in creating this plot, how did I get these dots? Because we didn't see them for example, yesterday we ran these mixed models; we didn't see the individual random effect estimates. But here I've got this plot using those random effect estimates. For each person their intercept, their slope. How did I get those? By default you're not going to get those so if I run this, I won't get an estimate each person's intercept and slope what it is a variance of the intercept. What's the variance of the slopes. So let's say I want that but I also want every person's intercept estimate, every person sloped estimate.

On this slide here I have how you can do that. What it involves is the SAS statement ODS. Does anybody use ODS? As its name implies. It stands for Output Delivery System. And basically it's a way of getting all kinds of things from a given SAS procedure to be output to a new data set. In this case I'm going to be using it to give a solution other random effects, put that out to a file, and then I can do some work on that.

Okay, so everything here is the same as before in terms of the random intercept and time model, except I should say in the random statement I also have to add here, ask for solution. Okay, so I need that there if I want to get the solution for the random effects. In other words, to get each person's intercept, each person's time, effect.

Okay on the next line, what I'm directing it to do is this. First, I'm saying, well don't list those estimates out in the output, rather put them out to an output dataset that I name Randest. So what I'm doing with this line here is really two statements, I'm saying, in terms of the physical output, don't list those random effect estimates, instead put them out to this data file. Okay? Then what I do is I print them out. Okay, so you might say, why did you print them out of you asked not to list them here? It's true, it's like, I'm sort of doing what this would have done.

Okay, but this is how then to get these random effect estimates of each person's intercept slope out to a new data file. That's why I work on to produce that plot that showed each person's intercept slope in that bi-variant plot.

Now here is some additional work. Maybe I can skip over this. What I'm doing is to then reorganizing the data in multi-variant form. Okay, so here's a way. You know, sometimes you want to get the data in multi-variant form where each variable given time points is a different variable. Here is what I'm, that's what I'm doing here is I'm creating then a dataset that has a different Hamilton values at different time points with different variable names. Okay? So this slide accomplishes that.

And then this slide puts together the person's estimates of the random effect with the estimated interceptance slope and creates estimates of the person's Hamilton across the six time points. Okay? So with this slide what I'm doing is I'm taking each person's

intercept deviation, slop deviation, which I get from this file or this dataset randest, and using the estimated intercept and slope, I'm creating, what's the estimated Hamilton at every time point for every subject.

Okay, so I don't know that you necessarily want to do this in all cases, but this shows how you can then create for each person, what is their estimated dependent variable value at every time point? And then I plot those out and here's how I got that bi-variant plot of the intercept and the slope.

Okay, so that's how one can get these plots out. It's unfortunately, it's not a simple thing of just you know, plot this versus that, you have to do a little bit of work with the result of the analysis.

Okay, let's – before we try some of this, let me go a little bit further – yes, a question.

Question: So is that last part, is it **** you've got the other graphic on slide 14?

Answer: Yeah,

Question: The one you're time **** you're still creating the other one with the dots.

Answer: The one I'm creating here is the one with the dots.

Question: Oh, that's still the dots.

Answer: Yeah, that's the one with the dots. That's right. To get that Spaghetti Plot, I actually did that in SPSS and years ago, I went into it more recently and I can't seem to recreate it. Then change – you know, they change things and I don't know how. If you want to do that, let me get on the selling platform for a moment.

I've worked out a program with Scientific Software called Super Mix. Super Mix is available for, I think, it's $425. It does a lot of these plots that I'm illustrating today pretty automatically. Now, why didn't I talk about Super Mix for yesterday and today? More people have access – that know about SAS. So I'm describing everything in terms of SAS, but actually, the plotting aspects, if you like that, you might want to investigate this program, Super Mix, it has better plotting facilities, easy plotting facilities. The other advantage it has is for categorical outcomes like dichotomous outcomes, ordinal outcomes, counts. It's very fast relative to SAS. It's much faster than that. So, some of these plots are pretty easy to get with Super Mix or with other software than with SAS, per se.

Okay, let's now answer the question of diagnosis. Does it matter if you're endogenous or not? Yes, a question.

Question: I think it's big time break –

Answer: Oh, it's break. I'm sorry; I thought we go to 10:00. No?

Question: You said you didn't have to One hour 15 minutes.

Answer: Oh really. I'm just looking at the listing, it says from 8:00 to 10:00 and then 10:00 to 10:15 there's a break. Yeah, I don't know. Yeah, maybe can you check? They do have… so it's time for a break? Okay, so I guess we'll break then. Sorry, take a break.