# Multilevel models to study intersectionality

## Transcript MAIHDA - Examples and visualisations (video 4)

Part of the resource: <https://www.ncrm.ac.uk/resources/online/all/?id=20849>

Hello and welcome to Part 4, Examples and Visualisations of MAIHDA from the literature. And welcome back. I'm George Leckie, University of Bristol. So, let's get cracking with this.

 So, the very first published application of MAIHDA was by Clare Evans in 2018. She had an application on MAIHDA to individuals’ body mass index in America with health survey data. And so, what were the social identities of interest in her study? So, yeah, as usual, I get my pens out and we'll go for a lime tree green seems nice. She had five variables of interest. They were gender, ethnicity, income, education and age. So, five social identities. Gender had two categories, male and female. Ethnicity had three categories and you can see what they are up here, white, black, Hispanic. Income at four categories, low income, low/mid income, high/mid income and high income, and so that may well have been a continuous measure which has been categorised, that’s something you have to with MAIHDA analysis, we're working categorical variables. Education has four categories, less than high school, completed high school, some college but no degree, and I guess this is maybe, yeah, okay, and then college degree or more. Okay, fine. And then age has four categories as well, 18 to 29, 30 to 44, 45 to 59, and 60 plus. So, with two by three by four by four by four categories you multiply it up. What do you get? No, not 385, you get 384. So, that's a typo. That should say 384. That's the theoretical maximum number of intersections.

 Now it could happen that in the data you might observe, you might not observe individuals in all of them because some of these are going to be pretty rare. So, if you choose, for example, the rarest ethnic group, which perhaps is the Hispanic Latino one, and then you might perhaps 60 plus is, well, no, let's take the 18 to 29 year old Latino. And then let's say the high income. Okay, well, they're only 18 to 29. Are they really, you know, not going to have many high income. And then let’s combined high income, 18 to 29 with less than high school. Okay, so they're a high school dropout, yet they've got really high income and they're very young. You know, so you're going to get very few individuals in particular combinations. And more generally in those kind of multiply marginalised, marginalised are often rarer groups, and so if you're rare on this category, rare on that, rare on that, rare on that, there might be a particular interest because of that, but you're not going to observe so many individuals and it might even be empty in which case you might want to oversample and so on if you're collecting your own data.

 Okay, so, that's what's going on there. Okay, so let's shift on to the next slide.

 So, we've got 384 intersections, we're interested in how mean BMI varies across them in all its nuance. So, the first thing you might do is, well, the first thing that you will do is to fit Model 1. So, let’s get a new pen colour and we're going to go for, let's go for red. Okay, that sounds not particularly exciting in the colour palette, but there we go. So, another model simply estimates the intercept, which is the overall average BMI across all individuals. I think there's around 5,000 individuals nested within the 384 intersections and there's a two level data sets. So, the mean BMI is 28, but the mean BMI can vary across those 380-odd intersections and here is the variance, 1.8. But crucially, within any given intersection, people are heterogenous, and there's a lot of individual heterogeneity. Okay? So, the variance of individual BMI around BMI averages is 34.5. .And so, we're really interested in the relative magnitude of those two numbers and if you do the maths and say, “Well, what proportion of the total is due to the strata effects?” you'll see that it’s 5%.

 So, what we can say is that 5% of the variation in individual BMI is attributable to mean differences across these 384 intersections. And that's our effects size, our BPC effect size is saying what is the magnitude of intersection in inequalities? How important are they? Well, 5%. So, maybe we do the same study in another country, you get 8% and another one 2%. So, that's kind of interesting to compare.

 Okay, so that's Model 1. So, you do that first of all. And then what you do is you jump on to Model 2 and there that's the model where you put in the additive main effects and you find kind of probably what you expect in terms of the main effects, you know, the point of MAIHDA to go beyond that. So, we're not necessarily going to learn much new from the beaters(?). We see that he does have higher BMI than men, all that’s equal. So, do blacks versus whites, so do Latinos versus whites. Then we have education, completed high school, this is not significantly different from less than high school. Ditto some college, no degree. Their BMI is not significantly different from the reference category. But if you have a college degree then you are significantly lower BMI on average in the reference which was less than high school.

 Then there’s income, is there an income gradient in BMI? The answer to that is you're partially, so all else equal, relative to the low income, the low to mid income, the lower BMI, but not significantly so, the high to middle income are also lower and by a greater extent, but still not significant. It's only when you get to the high income versus low income contrast that the high income are half a BMI point less, and that is statistically significant.

 Age gradient versus the young. Yeah, here, we've got some significant effects. All age categories are heavier, although that comes down again when we shift from middle age to the 60 plus.

 Okay, so the standard kind of patterning that you might see in terms of that of the main effects, but the real thing of interest here is to what extent have we explained away the initial inequalities between the 384 intersectional means which was a variant to 1.8, by the inclusion of added main effects. Well, it dropped from 1.823 to 0.643. Okay, that's a reduction of about two-thirds. Okay? So, two-thirds of the variation in intersection means is captured by the additive main effects that will be introduced into the model, one-third is not. And indeed, you can see the statistic, a third approximately is given here. The percent of the variation unexplained by main effects, which means it's explained instead by the interactions conceptually, the two-way, three-way, four-way and five-way interactions between these five social identities in the model.

 So, it's saying you do a main effects model, yeah, you're only kind of capturing two-thirds of what's going on. You completely know nothing about the remaining one-third is what that's saying. But because we've done this inclusion of main effects within this multi-level MAIHDA model, we are going to learn about the interactions and we already have started to do so by quantifying their importance with this PCV statistic.

 And then we control down, okay. So, then we can look at our 384 intersections and predict their values and here the interaction effects from Model 2 are capturing how does the mean BMI at a given inception, okay, for example, low income white males who are 60 plus who have got high school education, what's their mean outcome? And how does this deviate from what the model expects for that combination of characteristics based on the general patterns that we see across the nature of the whole. And so, we're capturing outliers, unusual mean outcomes for intersections versus expectations based on a kind of additive thinking.

 And so, in her paper, Clare's done this. Now she's got five dimensions, which is quite hard to plot, okay? So, let’s go for a blue. Okay. I think Zoom should improve the colour palette and have some more exciting colours there. So, with five dimensions, so she's explicitly looking at the interaction effects from Model 2, how intersections deviate from model expectations based on the main effects, and she's plotting this, she's really highlighting low income, so that's the income dimension versus high. She's not considering the two income bands in the middle, just to simplify what's going on, so looking at the extremes of income. And then she's also looking at ethnicity, white, black, Hispanic, so that's two dimensions. And then gender, third dimension, male or female, male or female? Then we've got age and education as the two remaining variables. For age there’s four categories, education there’s four, four by four is 16, so there’s 16 dots on each of these lines. So, this is showing how the age by education combinations, young and highly educated, young and less educated, middle age and you know all of them, how their mean outcomes differ from model expectations. Okay? And so, across the data as a whole, these are centred on zero because they're deviations, right?

[0:10:34]

 But you see particular combinations strike you. So, low income, black women, nearly all of these age by education combos are showing higher average BMI than what we'd expect given the overall additive main effects for ethnicity and gender. Remember back to that Kimberly Grenshaw plot, I think it was, with the bar chart at the bottom which showed mean effects on the left and then a bar chart of interactions on the right and how have you adjusted the main effect to get the wrong story? Well, this is what this is speaking to. There's something about black, low income females and BMI which leads them to show different BMI than what we'd expect in general by ethnicity and gender and income effects estimated across all the data where we've pulled all the individuals together to get those after main effects. There's something unique going on here. And so, that's what MAIHDA is flagging.

 Ditto kind of high income, white females, okay? They are all, all 16 of them, age by education combos are showing lower BMI than what we would expect. Now, crucially, these predictions are shrinkage predictions as well, so we're not just kind of getting persuaded by something variability showing some chance results. This is really a kind of conservative approach. And so, the fact that we're seeing action here means there really is action, if you like, because all these dots are being implicitly pulled towards the zero line in what we're seeing here. They’d be much more wild if we didn't do that. So, we're protecting ourselves. And so, we're quite confident that there's something going on there.

 Also, what else? White females, high education. So, on the bottom plot, so Clare's flipped from contrasting income to contrasting low education, higher education, ethnicity and gender. And so, we've got age by income combos being shown here.

 So, one of the difficulties here is that we are kind of trying to plot 384 means and we're trying to kind of discern patterns visually from it and we are kind of eyeballing at this stage. And so, you know, you've got to try different plots and see where the action is. So, there's a little bit going on here and I guess there’s a bit of subjectivity going on here as well. Okay, that’s the criticism, direct criticism of this approach. But we are trying to kind see where the action is in terms of interactions in a way which allows that more than really conventional regression which is very hard to work with if you've got a five-way interaction.

 Okay, let's clear those drawings out. Let's go on to another example. Here’s another paper, again by Clare Evans. This is a 2019 paper. She's obviously done a logistic regression MAIHDA analysis because the outcome is the predicted probability of cigarette use in that graph. So, let's go with, I’ve used all the colours, haven’t I? My favourite was lime tree green. I thought that was very nice. Okay, so, predictive probability of cigarette use here on the percentage scale. And so here she obviously had gender as one of the social identities, ethnicity, education, okay. But then in addition to that, quite interestingly in this study, they've combined those kind of classic social identities with something more traditionally contextual about the neighbourhood that you grow up in and the school you grew up in. And so, it's whether neighbourhood’s poor, whether the school’s poor, whether you're both in a poor school, poor neighbourhood. Sorry, non-poor or both poor or I guess no context is we don't know. And so, if we cut the data that way into those five groupings and that's what the five error bars are in each little panel, Okay? And you can see at any given pattern, sometimes you have kind of different shapes going on, sometimes a gentle increase, sometimes up and down and so on.

 So, again, you've got to be a little bit careful because we're eyeballing the data to talk about the patterns here to try and see where the action is. We're helped in that this is a conservative approach with the shrinkage which dampens down the noise in the data so that if we're still seeing action then its more believable there's something going on. But I guess an important message here is you need to do quite a lot of plotting and trying different things and being careful and considered with your interpretations.

 So, some important points to consider from these predictions from the model. So, in Model 1 you can you can use that model with the predicted random effects included to make the predicted mean outcome at each and every intersection. You can do that, but it turns out that doing the same thing for Model 2 where the predicted mini outcomes will also incorporate the active manufacture better because of shrinkage. Shrinkage in Model 1 is just towards the global average BMI, whereas in Model 2 shrinkage for each intersection is towards the expected outcome for that intersection based on the combination of characteristics and the added main effects and that in general will be a better prediction.

 So, if you wanted to predict mean outcomes and compare them, do them from Model 2 is what we’re saying. Okay? If you want to kind of study, well at what intersections is something going on which is unexpected, then you need to look at the Model 2 predicted random effects. And you want to say is it positive, is it negative, is it large, is it small, is it significant for any given intersection? But when you do that, don’t look at them just in isolation. You need to compare it towards what the predicted mean based on the additive main effects is. So, an example of this is if I discover that a particular intersection, the outcomes, “Hey, that they're kind of two BMI points higher than what we'd otherwise expect,” that means something very different if that is a group where what we're expecting is a higher BMI and it's even higher versus an intersection where actually we're expecting very low BMI and it's higher. Okay, so, just being two points higher, the interpretation of that depends on two points higher than what, depends on what. So, you need to look at that as well.

 Okay, great. Okay, so let’s bin those drawings and jump to this one. So, it’s Dan Holman at Sheffield, a collaborator of ours like Clare as well. So, this is a 2019 paper. I wonder what colour he would like? So, we'll give him a kind of an orange. Why not an orange? I haven’t had that for a while. He studied obviously one, two, three, four, five, six outcomes. I see there’s six for BMI, but we've got other ones in there as well. Big old sample of individuals. And he was studying intersectionality. I see we've got ethnicity, gender, education and income. Okay, so four social identities there. They're clearly kind of quite coarse here because it looks like ethnicity is just white and BME, which is black and other minority ethnic groups. Gender is just male and female. Education, we've got high education, low education being shown on this plot at least. Income we’ve got low, medium and high. Okay, so it's two by two by two by three, three twos are six times two is 12 times two is 24. Okay, so there are 24 intersections in this study, and that's what the graphical(?) league table is. Yes, should be 24 of those. I hope that's right.

 And what he found was when he did the kind of the null MAIHDA model and then the Model 2, the main effects model, he found that in the null model the importance of the inequalities for BMI is at just 1.6% compared to Clare's 5% for HbA1c, think that’s a marker of diabetes and it's more like 5% cholesterol, more like 9%. So, that's a key statistic to compare across outcome variables, yeah. And it's like an effect size. So, that's really cool. But when you put in the additive main effects for these four social identities, ethnicity, gender, education and income, what he found was that the importance, really, the kind of the remaining importance of the interactions versus individual heterogeneity, which is substantial, was pretty trivial. So, there might be interactions going on, but actually, you know, perhaps we shouldn't really be focusing on them because they're accounting for almost none of the variability in individual health outcomes having kind of controlled away the output of main effects, those four social identities.

 And so, what he did at that point was he kind of reverted to a linear regression model. Okay, so this isn't actually a MAIHDA analysis in the end, so he considered the MAIHDA analysis and actually wasn't supported. So, this is a nice application to say that just because you're learning a new fancy method doesn't mean it makes sense for your particular data, and if it doesn't, you should say so and pull back because it's not going to give anything more. And so, in this league table of the HbA1c, which is where the action, where there was quite a bit of variability across the intersections, you can see there's a graphical league table and it's lowest amongst white women, high education, high income, multiply advantaged, with respect to this, whereas BME men, low education, low income and multiply disadvantaged with respect to this health outcome with much wider confidence intervals here, because these are BME who are going to have a minority ethnic group versus the white majority. And then obviously get tighter precision with bigger samples. So, that’s the thing.

 I guess a nice genuine application where MAIHDA wasn't required, okay. And then a form of statistical tests you can do for seeing whether it's needed as well which I guess we might do in the statement of our videos.

 And here's the last slide from me on this little video. This is actually one of my papers. This is a paper I did with Lucy Pryor. I work in education, so it's an application to education. I'm going to choose a colour for me. I think I'll go for ice blue. The outcome was in England, age 11, achievement or test scores are standardised. That means zero standard deviation one. We defined, were interested in the following social identities. Age, which is term of birth, either you're August born, spring or summer. We’ve got autumn born, older in the year, so do better. Gender, male or female. Free school meal marker of low income. Either you are poor or not poor, effectively. Special educational needs, either you have them or you don't. And then ethnicity, we had a white, black, Asian students, mixed ethnicity students, other and unclassified, which is affecting missing data, unknown.

 And so, that's quite a lot. One, two, three, four, five social identities. So, conceptually we'll want to be allowing for up to five-way interactions and we've got three, oh God, help me come and do the maths on here. Five by two is 10 times two is 20 time two is 40 times three is 120. I hope I've got that right. Embarrassing otherwise. So, let's say we've got 120 intersections. And in each intersection we've got a number of students and some intersections are much bigger than others. So, and I'm showing you here the top ten in terms of their mean achievement on the standardised outcome and the bottom ten. Okay, so the very top intersection is scoring 0.57 of a standard deviation higher than the overall average. Who are they? They are autumn born, so they're old in year. That's an advantage. They're female. They're not poor. Okay, and you can imagine the rich do better. Okay, they don't have special educational needs, so not surprisingly that they score higher. But in terms of ethnicity, they are mixed ethnicity, which means that they might be white and black, white and Asian, some kind of two parents of different ethnicities. And there's 790 of those students and that's the top.

 Now in terms of the top ten, there's some commonality and this is kind of where things go a bit interesting. We can see seven out of ten are autumn born, and if you're not autumn born, you're spring born. We have no summer born kids in the top ten of the league table because they're at a disadvantage that you know, the half a year, three-quarters of the year younger than they're older peers. They don't score so well at age 11. Makes sense, yeah? Gender, girls are six, they win six to four in the top ten, so it means that gender isn't really kind of crucial for determining your ranking, if you like. And now in contrast, the really important factor is clearly free school meals because the top ten, none of them are on free school meals, none of them are poor, whilst the bottom ten, all but one are poor. Likewise special educational needs, none of the top ten have special educational needs, all of the bottom ten have it. So, these facts are really important. Okay, so, they've got big effect sizes in terms of the additive main effects, if you like. But ethnicity, the patterning is not clearcut. It's more like gender and we see actually in the top ten, all ethnic groups are represented apart from black students and in the bottom then the lowest performing, all ethnic groups are represented somewhere. Okay, so ethnicity is not so kind of determining this league table.

 There's lots of kind of interesting interpretations you can make. I was making a mention that some of these intersections are larger than others. So, the largest one is this one here, 4,022 kids. So, these are autumn born, well, a third, a third, a third, a third are autumn, spring, summer. So, that's not very important. Gender, there's a 50/50 splits, so that's important. Now they are not on free school meals. Now maybe three-quarters of kids are not free school meals. So, the majority group, and not SEN, maybe that's about 90%. So, again, majority white, yeah, and white’s the biggest ethnic group. So, that makes sense because it’s the biggest. Whereas the rarest with only 15 kids, okay, now again, age and gender don't matter because the data's balanced with respect to that. Now they are poor, again, only a quarter of kids are poor overall. They are also SEN, and only 10% are SEN overall, so rare group combined with rare group combined with the rarest ethnic group, missing ethnicity, and there’s only 15 of those when it comes to the males and autumn born before that.

 So, I'll predict a mean outcome -1.12, which is a shrinkage one, so it’s got much better statistical properties than simple mean and it's still going to have a wide confluence on it as there’s only 15 data points. Okay, but often we're interested in those multiple rare categories and multiple marginalised.

 So, hopefully this was a useful plot to show you, kind of you know different ways to interpret the data, different things you can do to try and communicate your results. Hope that was useful. And that's me done on that video. It’s going to be back to Andy for the fifth video, and I believe the fifth are going to be, we're going to have the R and Stata videos. Okay, so, you'll see one of us, depending on which software package you pick, which data. Okay. Bye bye, thank you.

National Centre for Research Methods (NCRM)
Social Sciences
Murray Building (Bldg 58)
University of Southampton
Southampton SO17 1BJ
United Kingdom

**Web** www.ncrm.ac.uk
**Email** info@ncrm.ac.uk
**Tel** +44 23 8059 4539
**Twitter** @NCRMUK