Context and Change in Health Research: A Multilevel Model Approach

Julia McQuillan

University of Nebraska,

Goals of this Lecture

- **Provide a conceptual overview of the *kinds of questions* that Multilevel Models (MLM) can answer and the *kinds of data* for which MLM is useful**
- **Encourage research that can take advantage of MLM applications for health research**
- **(Part II) Practice connecting research questions with MLM output and introduce advanced issues.**

Overview Of the Lecture

- Introduction
- Kinds of DATA
- Kinds of QUESTIONS
- The Basic Logic of MLMs
- Substantive Examples
- Questions/Explorations
- Useful Web pages, books, articles

Introductory Information

My Entry into MLM: Trying to model Change

- Dr. Susan Reisine and Dr. Judith Fifield sent me to a training by Stephen Raudenbush and Anthony Bryk (authors of the Hierarchical Linear Modeling (HLM) book and software)
- Their goal was to use an HLM approach to their 10 year panel study of individuals with Rheumatoid Arthritis (RA).

Examples of “change” questions we explored:

- Does the general level and rate of change of fatigue over time (the fatigue trajectory) differ for those with and those without a history of affective disorder?
  - Fifield, Judith, Julia McQuillan, Howard Tennen, T. Joseph Sheehan, Susan Reisine, Victor Hesselbrock and Naomi Rothfield. 2001. “History of Affective Disorder and the Temporal Trajectory of Fatigue in Rheumatoid Arthritis.” *Annals of Behavioral Medicine.* 23(1):34-41.

More Change Questions

- Are changes in daily negative and positive events associated with changes in daily stress? Are these associations modified by work context (i.e. demands and control)?
  - Fifield, Judith, Julia McQuillan, Steven Armelli, Howard Tennan, Susan Reisine, and Glenn Affleck. 2005. “Strain, Daily Work Stress and Pain in Workers with Rheumatoid Arthritis: Do Job Characteristics Make a Bad Day Worse?” *Work & Stress* 18(4): 275-291.

Did the intervention work?

- Do increases in guideline adherent medication (GAM) improve asthma control? Does electronic feedback to physicians regarding guideline adherent medication increase the likelihood of prescribing GAM?
  - Fifield et al, manuscript in progress

Applying MLMs to questions of context

- Dr. Berdahl and I are examining occupational context (occupational racial composition) on the odds of a workplace injury
- Dr. Torres Stone and I studied labor market effects on the association between Hispanic subgroup membership (Puerto Rican, Mexican and Cuban) on wages, and labor market effects on the motherhood wage penalty
  - Stone, Rosalie Torres and Julia McQuillan. 2007. "Beyond Hispanic/Latino: The importance of ethnicity-specific earnings analyses." *Social Science Research.36:1:175-200.*

Many Possibilities for MLMs in Health Research

- Many major health improvements have come from efforts at higher levels (e.g. clean water)
- MLMs are becoming more common in health research, for example:
  - Kim and Kawarchi (2007) show that state level **context** (social capital) is associated with **individual** health related quality of life using a MLM approach

Why do people use MLMs?

- They have the *type of data* that is appropriate for multilevel models (multiple observations within people (longitudinal/panel data), multiple family members, multiple patients in therapy practices, or physicians, or clinics, or hospitals,…)

AND/OR

- They have the *types of questions* that are appropriate for multilevel models (more on this later)

In the second part of this lecture (advanced/applied)

- We will learn how to answer research questions with MLM output, and examine the equations underlying MLMs in more detail

Other names for MLMs

- **In education they are usually called Hierarchical Linear Models (HLMs), or multilevel linear models (MLM) and nonlinear models (HGLM) (sociology),**
- **Mixed-effects models and random-effects models (biometry),**
- **Random-coefficient regression models (econometrics)**
- **Covariance components models (statistics)**

Caveats

- Interrupt me if you have a question
- This is a conceptual introduction, I'm not a statistician
- I use the terms HLM (software brand name) and MLM (generic "multilevel model") interchangeably, you may have heard other names

If you can learn OLS regression

- Than you can learn MLMs
- They use the same logic
- BUT they do require a considerable investment to understand well enough to do them well.
- This talk should help you decide if it is worth your effort to learn MLMs

Software

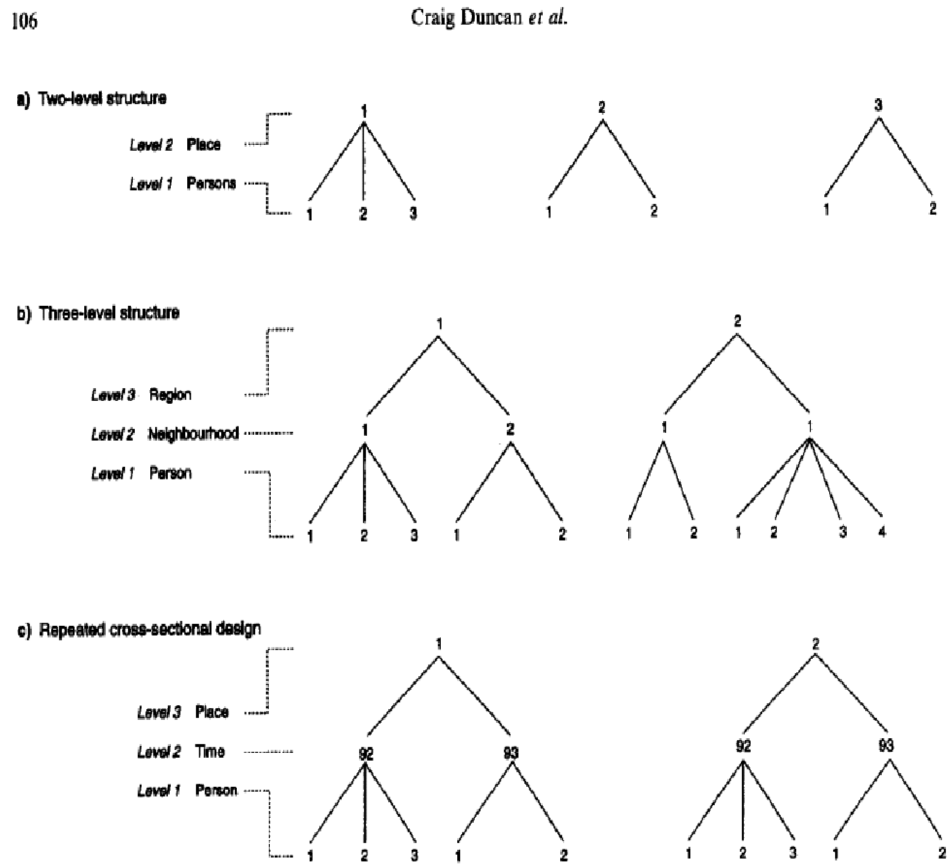
- I learned MLMs on HLM, it is most comfortable for me
  - It has limitations
- Many software packages do MLMS (SPSS Mixed, SAS, STATA, MLwin, MPLUS, etc.)
- Some are more flexible, some are a little more difficult to correctly set up
- It is always a good idea to run models in at least two programs

Kinds of Data

Examples of Types of Data

- multiple observations within individuals **(repeated measures)**
- multiple observations within couples
- multiple family members within families
- multiple clients within therapists
- multiple families within neighborhoods
- multiple patients within hospitals
- multiple diease outbreaks within states

Multilevel Data Structures Duncan, Jones, Moon (1998)



More Multilevel Data Structures Duncan, Jones, Moon (1998)

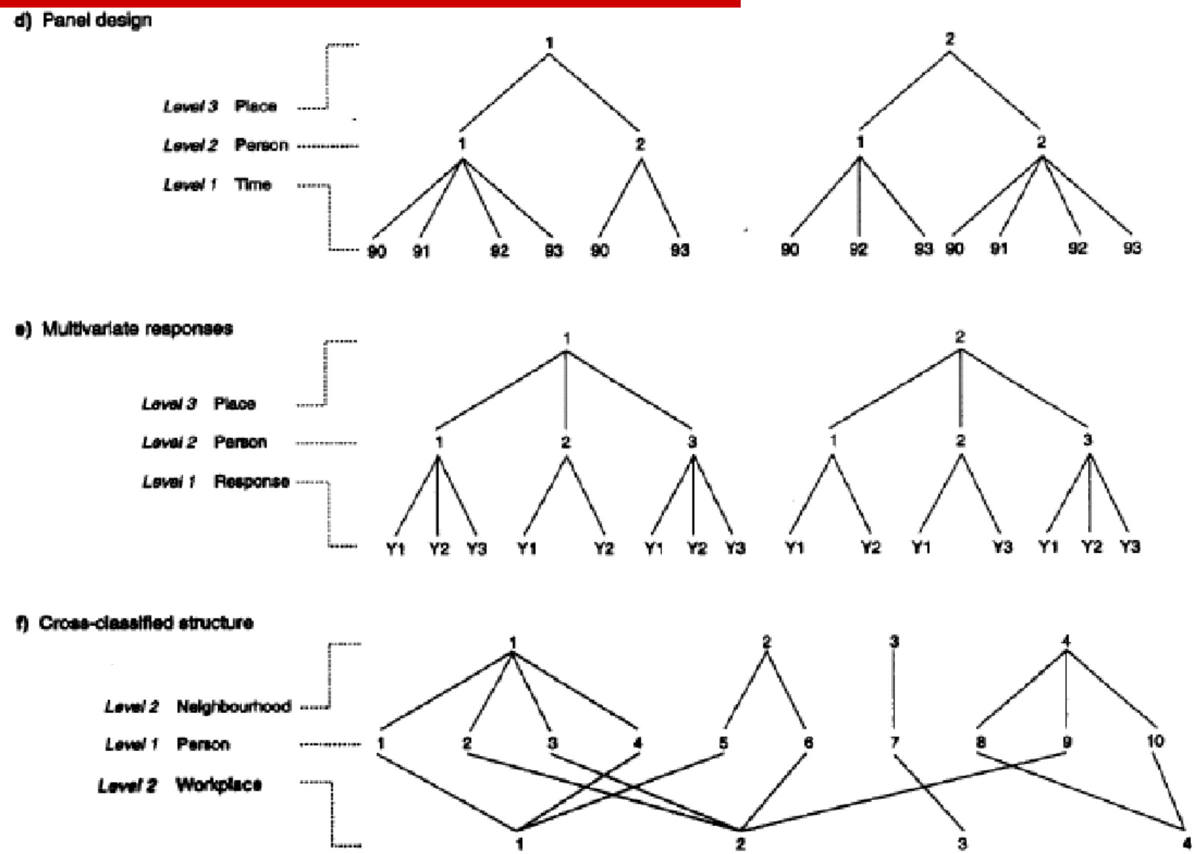


Fig. 8. A range of multilevel data structures.

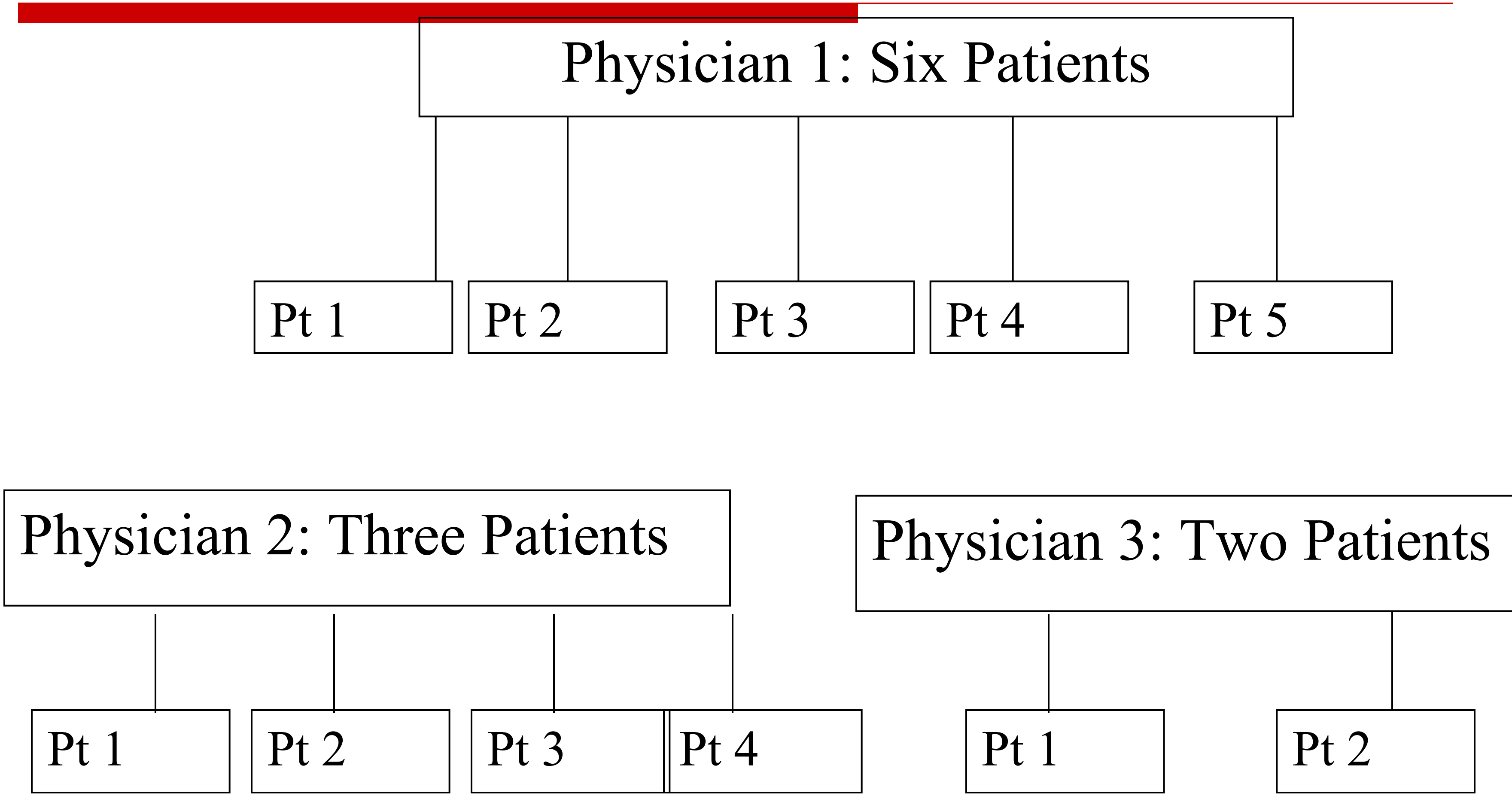
Made up example to show data issues because I have no published context and health papers

- Over simplified
- Hopefully this makes sense
- Pretending that physicians were randomly sampled and patients were randomly sampled within physicians
- Have data on multiple patients within doctors, and multiple doctors

This data can answer questions such as:

- Do some physicians spend more time with their patients than other physicians?
- If yes, why do some physicians spend more time than others?
  - Is it because they’ve had the patients for a longer or shorter time?
  - Is it because they were trained as family physicians or not?

Nested Data: Patients in Practices



Example of a Physician Flat file: Each Patient has a column of data for each response

Physician ID	Family Phys.	Pt Volume	N Years 1	N years 2	N years 3	N years 4	N years 5	Min1	Min2	Min3	Min4	Min5
01	0	200	12	11	10	.	.	20	15	25	.	.
02	1	100	6	5	8	7	9	30	60	55	70	45
03	0	50	1	3	2	5	.	15	20	30	16	.
04	1	150	8	9	.	.	.	35	40	.	.	.
05	0	60	5	10	.	.	.	15	30	.	.	.
06	0	90	6	4	5	2	.	12	16	5	22	11.

This is the standard data structure for OLS data.

This would also be common if data was collected annually, and new variables were added, e.g. Pain1, Pain2, Pain3….

Example of Patient level **Stacked** Data (Level 1) File (Each patient has a row of data)

Phys. ID	PT ID	N Years	Minutes	Fam.Doc.
1	001	12.00	20.00	0
1	002	11.00	15.00	0
1	003	10.00	25.00	0
2	004	6.00	30.00	1
2	005	5.00	60.00	1
2	006	8.00	55.00	1
2	007	7.00	70.00	1
2	008	9.00	45.00	1
3	009	1.00	15.00	0
3	010	3.00	20.00	0
3	011	2.00	30.00	0
3	012	5.00	16.00	0
4	013	8.00	35.00	1
4	014	9.00	40.00	1
5	015	5.00	15.00	0
5	016	10.00	30.00	0

The stacked file can be aggregated up to the physician level to get average physician data

Phy ID	PtID mean	Yrs Mean	Min Mean	Fam Doc	N Pts
1	2	11	20	0	3
2	6	7	52	1	5
3	10.5	2.75	20.25	0	4
4	13.5	8.5	37.5	1	2
5	15.5	7.5	22.5	0	2
6	18.5	4.25	10.5	1	4

Which Data set should we analyze?

- If we want to run a regression equation with:
  - minutes as the outcome
  - years as a patient and family doctor status as predictors
  - which is better:
    - The flat physician file
    - The stacked patient file
    - The aggregate physician file?

What if we analyze the stacked file?

- We could analyze the stacked file with a row for each patient and include physician information.
- Largest sample size this way.
- But we’d likely violate the assumption of independence, because patients are clustered within physicians

What if we analyzed the aggregate physician data?

- We'd have more reliable measures (based on multiple patients)
- We would not violate the assumption of independence
- But, how could we know if there are differences within physicians between patients? For example, how could we know if patients who have been with a doctor longer spend more or less time with the physician?
- No way to learn "within physician" information

What’s an analyst to do?

- Fortunately, you don’t have to chose
  - MLMs provide a way to do both, simultaneously, and to take care of the lack of independence
  - Conceptually, you can:
    - run the regression on the data from each physician separately,
    - and then use the estimated intercepts and slopes as variables (latent) to answer between physician questions (e.g. Family docs spend more time?)

Recall the basic regression equation

- $Y = a + BX + r$
- Y = the outcome (minutes per visit)
- a = the intercept (mean of Y when X = 0)
- B = the association between X and Y (presuming X = N yrs in the practice, it is the change in Y when X increases one unit)
- X = independent variable (e.g. N yrs in the practice)
- r = error = Y^-Y (Predicted – Observed)

Imagine running a regression for the data for each physician (N=6)

- You’d estimate 6 intercepts
- You’d estimate 6 slopes
- You could treat those intercepts and slopes as a new Physician level variables in the aggregate data set (a = the physician average length of time with a patient when X, years in the practice, = 0, and B = slope)

Examples of possible findings: 1.Constant Intercepts and Slopes

- We could find that all doctors spend the same average amount of time with patients, and that the association between length of time in the practice and time per visit is similar across Patients

$a1=a2=a3=a4=a5=a6$ intercepts same

and

$B1=B2=B3=B4=B5=B6$ slopes same

2. Varying Intercepts, Constant slopes

- We could find that some physicians spend more time with patients on average and some less time, but that the association between years in the practice and minutes per visit is the same for all physicians

  a1>a2>a3>a4>a5>a6 intercepts vary

  and

  B1=B2=B3=B4=B5=B6 slopes same

3. Constant intercepts, varying slopes

- We could find that physicians tend to spend the same amount of time with patients on average, but that the association between time in practice and minutes per visit differs between physicians

  a1=a2=a3=a4=a5=a6 intercepts same

  and

  B1>B2>B3>B4>B5>B6 slopes vary

4. Varying intercepts, Varying Slopes

- It is most likely that physicians differ both in average time with patients and with the association between time in the practice and minutes per visit

a1>a2>a3>a4>a5>a6 intercepts vary

and

B1>B2>B3>B4>B5>B6 slopes vary

Hoffman (1997) provides an illustration of these four possibilities

Examples of Intercept and Slope Patterns from Multiple Clusters

(Hofmann, David A. 1997. An Overview of the logic and rationale of hierarchical linear models. Journal of Management. 23:6:723-744.)

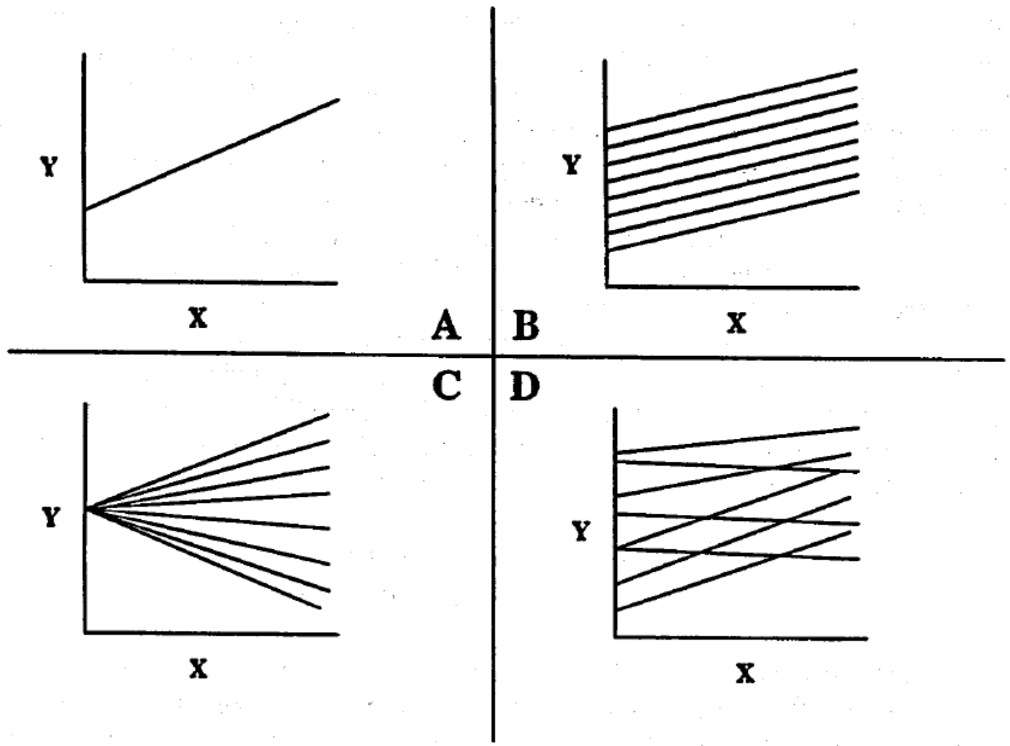


Figure 1. Four possible patterns for intercepts and slopes when level-1 models are estimated separately for each group.

Take home message regarding data

- If you have data that requires you to chose between analyzing stacked data (and violating independence) or analyzing aggregate data (and forfeiting knowledge of within unit variance) then MLMs are probably for you!

Kinds of Questions

Kinds of Questions I

- Is neighborhood disadvantage (context) associated with adult depressive symptoms, controlling for individual characteristics?
  - (Ross, Catherine E. 2000. Neighborhood Disadvantage and Adult Depression. *Journal of Health and Social Behavior*. 41: 177-187).
- Does individual education level protect against daily stressors, physical symptoms and psychological distress?
  - Is the stressor to health association weaker for those with more resources?
  - Grzywacz, Joseph G., David M. Almeida, Shevaun D. Neupert, Susan L. Ettner. 2004. Socioeconomic Status and Health: A Micro-level Analysis of Exposure and Vulnerability to Daily Stressors. *Journal of Health and Social Behavior.* 45:1-16.

Kinds of Questions II

- Are mental health symptoms higher over the life course among those whose parents divorced compared to those with continuously married parents?
  - Cherlin, Andrew J., P. Lindsay Chase-Landsdale, Christine McRae. 1998. "Effects of Parental Divorce on Mental Health Throughout the Life Course. " *American Sociological Review.* 63:239-249.
- Does stigmatization from receiving mental health services depend upon provider characteristics (context) or individual characteristics or both?
  - Verhaeghe, Mieke, and Piet Bracke. 2007. Organizational and Individual level determinants of stigmatization in mental health services.*Community Mental Health Journal*. 43(4):375-400.

Kinds of Questions III

- Is health related quality of life higher in states with more social capital?
  - Does individual socioeconomic status mediate the association?
    - Kim, Daniel and Ichiro Kawachi. 2007. U.S. State Level Social Capital and Health Related Quality of Life: Multilevel Evidence of Main, Mediating and Modifying Effects. *Annals of Epidemiology*. 17:4: 258-269.
- Does HIV/AIDs education increase condom use confidence?
  - Does the association depend upon weekly meetings or an intense weekend **intervention**?
    - **O'Connell, Ann A and D. Betsy McCoach. 2004. Applications of Hierarchical Linear Models for Evaluations of Health Interventions: Demystifying the Methods and Interpretation of Multilevel Models. *Evaluation & the Health Professions*. 27:119-151.**

Kinds of Questions, IV

- Are residents of poorer regions at greater cardiovascular risk than residents of wealthier regions?
  - What explains more variance in cardiovascular disease, regional or individual characteristics?
  - Breckenkamp, Jeurgen, Andreas Mielck, Oliver Razum. 2007. Health inequalities in Germany: do regional-level variables explain differentials in cardiovascular risk? *BMC Public Health* 7:132

Kinds of Questions V

- Are adolescents in poorer communities (low SES context) at greater risk of depressive symptoms and engaging in delinquent acts?
  - Does perception of social support from peers and adults weaken any association between community context and adolescent outcomes?
  - Wright, Richard G., Amanda L. Botticello, Carol S. Aneshensel. 2006. Socioeconomic Context, Social Support, and Adolescent Mental Health: a Multilevel Investigation. *Journal of Youth and Adolescence*. 35:1:115-126.

How did they answer these questions?

- They had at least 2 levels of data (e. g. people in regions, people in communities, observations within people)
- They used MLM analyses
- If you have similar questions or data, keep listening

Basic Logic of Multilevel Models

(Building on ideas introduced in the "kinds of data" section, now using a "**change**" or longitudinal example)

Fatigue Trajectory Example

- 415 Survey Participants
- Up to 7 years of data on each (“year0”
  - Skipped one year, not everyone has 7
- Diagnostic interview schedule measures of meeting criteria for lifetime Affective Disorder (AD), specifically Major Depression (MD), Generalized Anxiety (GA), or both (Comorbid)

Basic logic of MLMs 1

- Using the Fatigue Trajectory example:
  - Should we analyze all of the time points and ignore that they are nested in individuals?
  - Should we analyze all of the individuals, and ignore that observations vary within individuals in levels of fatigue?
  - How can we capture the trajectory over time?
  - Similar to Physician context data, we face the question - which data should we analyze?

Of course, MLMs can do both!

- Called “latent growth curves” when analyzing repeated observations within units (in this case people)
- We can estimate overall average fatigue and how much variance there is around the average (do individuals differ?)
- We can estimate average change in fatigue (the trajectory), and how much variance there is around the average (do individuals differ?)
- And if levels and slopes are associated

Basic Logic of MLMs 2

- If we analyze the stacked data (each observation/year has a separate row of data),
  - we ignore nesting in individuals, and
  - we get the OLS regression results that I’ve estimated, summarized on the next slide

Regression results using Stacked data, *ignoring dependence within individuals (N=2,905)*

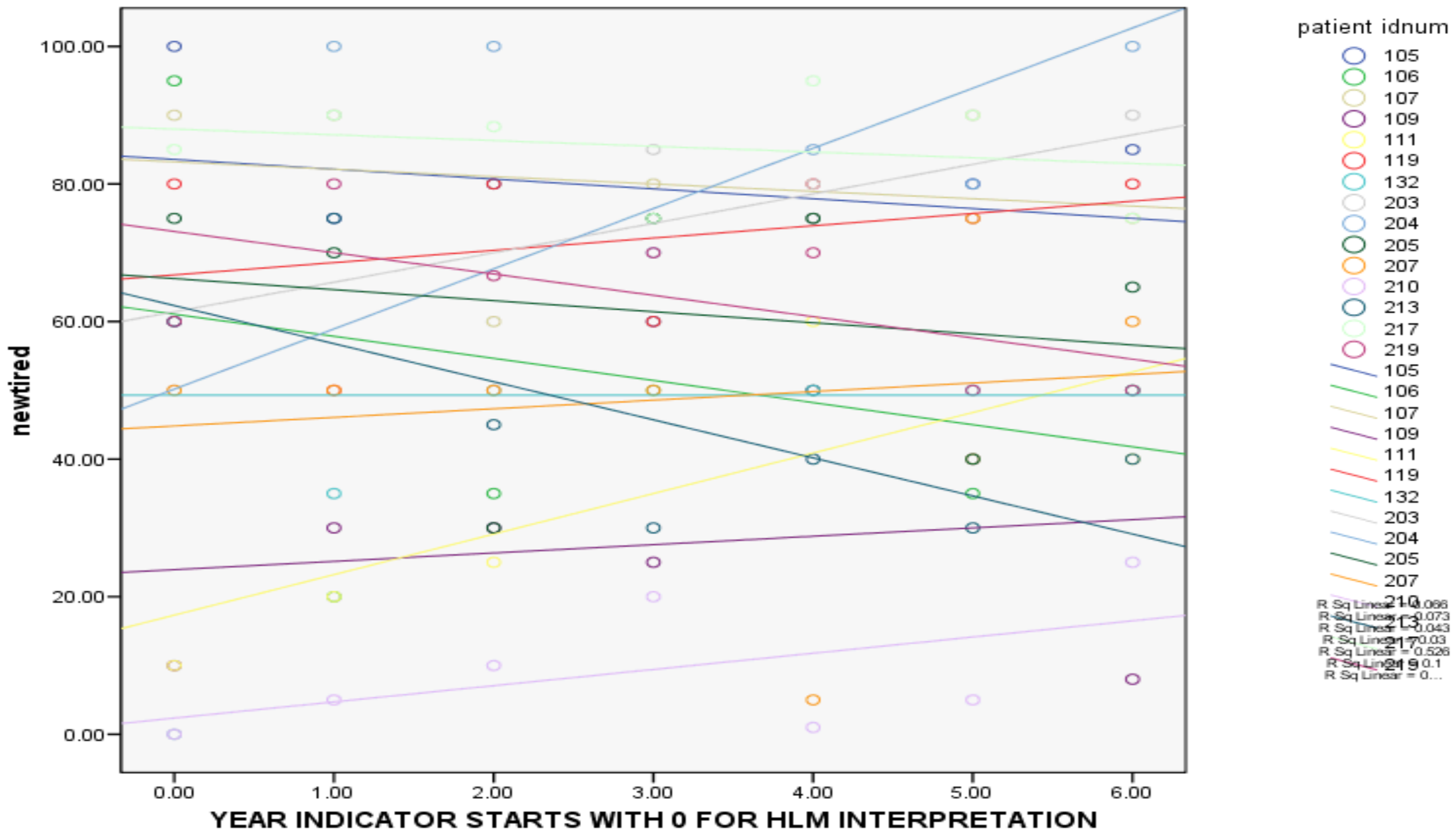
Stacked	B	SE	
(Constant)	57.11	1.37	***
Year0	.51	.26	
ad	-9.61	1.27	***

Dependent Variable: newtired

Basic Logic of MLMs 3

- If we analyze each person separately, we find that fatigue increases for some, decreases for others, and doesn't change for the rest
- See the amount of variation in the regression lines on the next slide (selected the first 15 participants)

Regression Lines by Person



Regression equations summarized, *Separate for each person*

Person		B	SE
105	(Constant)	83.57	8.70***
	Year0	-1.43	2.41
106	(Constant)	61.07	18.53*
	Year0	-3.21	5.14
107	(Constant)	83.21	8.16***
	Year0	-1.07	2.26
109	(Constant)	23.93	11.10
	Year0	1.21	3.08
132	(Constant)	49.29	5.46***
	Year0	.00	1.52

Bickel (It’s just regression)

- He shows how to combine the separate regressions for each person into an overall analysis that, using weights, approximates MLM analyses.
- Recall that the regression actually created 415 constants (intercepts) and slopes. We could then save these and analyze them as outcomes

It is important to find out if nesting matters

- When relationships differ between level 2 units, we lose considerable amounts of information when we ignore nesting
- As the previous slide shows, relationships can even change direction between individuals

What if we just look at individuals?

- We can ignore the variation within individuals, and avoid dependence between observations by **aggregating** separate year information into overall person information

- This is similar to analyzing schools, ignoring within school variation between students

Example of person level analysis

“Year” can’t be estimated because it is a constant at this level (mean years).

	B	SE	
(Constant)	58.65	2.16	***
ad_mean	-9.61	2.46	***

a Dependent Variable: newtired_mean

Comparing within and between

- The separate regressions for each individual focus on within person change over time
- The aggregate regression across individuals focuses on between individual differences in fatigue

Which should we use?

Basic Logic of MLMs 4

- MLMs can do both at once
  - They estimate an individual line for each person, and an overall average across all individuals (weighted by how many observations they have), creating less biased and more informative estimates

  - They also provide a way to see if individual level characteristics (e.g. a History of AD_ moderate individual associations (cross level interactions)

The Basic Logic of MLMs 5

- Conceptually, the program estimates a separate regression for each person (level 2 unit)
- Level 1:

  For example, for each person:

  Fatigue = B0 + B1Time+….r
- THEN comes the TRICKY PART…

The Basic Logic of MLMs 6

- The coefficients from all of the level 1 models (B_o, B_1, $B_{2,}$ etc) become the dependent variables at level 2.

  $Fatigue_{ij}=B_o+B_1Time+r_i$

For example:

$$B_o = G_{oo} + G_{o1}\,(AD)+U_0$$

$$B_1 = G_{1o} + G_{11}\,(AD)+U_1$$

(AD = Affective Disorder)

The U0 and U1 are very important – these are “random effects”

Basic Logic of MLMs 7

- How does it do this?
- Something very special, called the EM algorithm and Bayesian estimation
- It uses iterative maximum likelihood procedure to fit the model to the data
- Fortunately, most of the time, you don't really need to understand how the under lying math works, *just* how to interpret the results

Basic Logic of MLMs 8

- Understanding what MLMs tell us can be challenging
  - It's very important to carefully specify the level 1 model
  - It's also very important to know what a 0 value on each independent variable means
- This requires careful thinking about each coefficient and what information it has.
- This requires thinking about centering

Why Center?

- It provides more meaningful coefficients and helps with the multicollinearity produced by interactions
- It is related to the constants (intercepts) becoming outcomes at level 2 – it helps if they are meaningful.
- We work the predictors so that the answer to the question: What is the value of the constant? Is meaningful.

How Center?

- Usually, subtract the mean from each value (Sometimes grand mean (over all cases), sometimes group mean (mean within the group, for example a family). I'll save which to choose for questions….
- Sometimes you might choose another strategic value (First year, middle year, last year of a panel study) – For the fatigue study, year 1 = 0, making the constant equal to "initial" fatigue

Error terms (Variances)

- Indicate variance with in individuals and between individuals (what percent at each level?) In the fatigue example, it was about 50% within and between.
  - Recall what “variance” is – average squared deviation
- You can also relate the variances of the errors to each other (Are slopes and intercepts correlated? Do individuals with higher averages fatigue tend it increase or decrease in fatigue over time?)
- 1.96 +/- the SD of the Variance Component = 95%ci – the confidence interval – most values in this range

There’s so much more!

- But now you have the basic idea
  - Nested data
  - Questions about context
  - Levels of analysis, with coefficients as outcomes
  - Variances at each level
- The formal equations for the fatigue model are on the next slide
- Just to confuse you, when time is the focal predictor (growth curves) the symbols for the coefficients are no longer B (OLS regression) nor Gamma (MLM) but Pi

From the article, the growth curve:

36 **Fifield et al.**

$$Y_{ti} = \pi_{0i} + \pi_{1i}\, a_{ti} + e_{ti}$$

for i = 1 … *N* individuals and Y_{ti} = fatigue score for person i at time t

π_{0i} = the intercept parameter indicates the estimated fatigue level for each person in the 1st year of this study, and π_{1i} = the linear slope parameter indicates the growth rate for each individual for each year

a_{ti} = the measure of time (in this study, the year of the interview for individual i)

e_{ti} = the error, usually assumed to be independently and normally distributed with a mean of 0 and a constant variance σ^2

From the article, main effects (pi0) and cross level interactions (pi1):

The following equations describe the Level 2 relations (32)

$$\pi_{0i} = \beta_{00} + \sum_{q=1}^{Q_0} \beta_{0q} X_{0qi} + r_{oi}$$

$$\pi_{1i} = \beta_{10} + \sum_{Q=1}^{Q_1} \beta_{1q} X_{1qi} + r_{1i}$$

X_{qi} = measures of characteristics at the person level for each individual i

β_{0q} = the effects of the characteristics on the slope and intercept parameters

r_{oi} and r_{1i} = error terms that are assumed to be uncorrelated with the characteristics and are multivariate normally distributed with means of 0. The error terms represent unmeasured characteristics of individual i that do not change over time.

Complex Error Structures

- For over time (longitudinal/latent growth curves) you also need to check the error structure
  - E.g. Are observations closer together more highly correlated than those farther apart? (AR 1?)
  - Is some other complex error structure required?

HLM output: Levels 1 and 2

The outcome variable is NEWTIRED
Final estimation of fixed effects:

Fixed Effect	Coefficient	Standard Error	T-ratio	P-value
For INTRCPT1, B0				
INTRCPT2, G00	47.305954	1.470472	32.171	0.000
HASAFHX, G01	10.463765	3.089703	3.387	0.001
For YEAR_0 slope, B1				
INTRCPT2, G10	0.579059	0.262613	2.205	0.027
HASAFHX, G11	-0.285328	0.551792	-0.517	0.605

Now the intercept is smaller because it is the mean for those without an AD (Has an Affective History (HASAFHX) = 0), the slope for “year”

Variance Components from HLM

Final estimation of variance components:

Random Effect	Standard Deviation	Variance Component	df	Chi-square	P
INTRCPT1, U0	22.87484	523.258	414	1563.907	0.000
YEAR_0 slope, U1	2.75814	7.607	414	631.330	0.000
level-1, R	20.14347	405.759			

Statistics for current covariance components model

Deviance = 26716.49204
Number of estimated parameters = 4

Deviance = -2LL, relative fit (Compare to baseline model, chi-sq)

Chi-square P shows that there is significant variance between individuals in average initial fatigue and in the rate of change in fatigue (measured by the coefficient for "year_0".

Luke (2004)Mixed Equation

Level 1: $$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j}$$

Level 2: $$\beta_{1j} = \gamma_{10} + \gamma_{11}W_j + u_{1j}$$

Equation 2.2 Mixed Model

$$Y_{ij} = [\gamma_{00} + \gamma_{10}X_{ij} + \gamma_{01}W_j + \gamma_{11}W_jX_{ij}] + [u_{0j} + u_{1j}X_{ij} + r_{ij}]$$

[fixed] *[random]*

Noticed what has happened – the betas (β) are gone – but what they equal (the right of the equations) is still there. It's also more explicit that W is actually a modifying effect for the association between X and Y.

Early books were formidable, newer books are more accessible…

- Bickel’s (2007) “Multilevel Analysis for Applied Research: It’s Just Regression!” builds from researcher’s knowledge of OLS – although MLMs are a little more complicated than the title implies
- Luke (2004) has a short and clear Sage book on MLMs

Many good introductory resources are available if this talk motivates you to learn more…

- O’Connell and McCoach provide an accessible introduction to MLM for evaluating interventions with longitudinal data
- Duncan, Jones and Moon (1998) also provide a basic, brief introduction
  - Details in the “resources” handout

We won’t have time for

- Exploring data before analyzing (but you should and several books will tell you how)
- Assumption checking (e.g. that the errors are normally distributed)
- Power Analyses and project design (how many level 1, how many level 2 will I need? – Short answer: consult an expert who can do simulations for you)
- Non-linear outcomes (binary, count, etc)

Nor do we have time for

- The complexities of missing data
- Setting up equations
- Meta Analyses
- Graphing interactions
- Plots for assumption checking
- BUT – many of these topics are covered in archived workshops at:

SSP.unl.edu

And other sources listed in "MLM Resources"

Alternate approaches to change

- *Repeated measures* (but would need to exclude cases involved less than 7 years, and could not answer all of our questions) **(See O'Connell and McCoach for the benefits of a MLM approach)**
- *Fixed effects pooled time series* (could not examine between people, only within person change – useful for 2 waves)(Johnson 2005)
- *Latent growth curves with SEM* (Mathematically equivalent, see Chuo et al 1998 for a comparison – sometimes a better choice)
- See http://www.longitudinal.stir.ac.uk/

Your turn

- Questions?
- Ideas?
- Explorations?
- Data?