



Longitudinal Diagnostic Classification
Models: Theory and Applications

Matthew Madison, University of Georgia





Access session materials at
www.matthewmadison.com/workshop





- Preliminaries: LCDM framework and estimation in R (1:15 – 2:00)
- Break 1 (2:00 – 2:05)
- Longitudinal DCMs I (single group) (2:05 – 3:00)
- Break 2 (3:05 – 3:15)
- Longitudinal DCMs II (multigroup) (3:15 – 4:15)
- Break 3 (4:15 – 4:20)
- Q&A, discussion, evaluations (4:20 – 5:00)



A Little About Me

- From Columbia, South Carolina
- Research interests: DCMs, measurement of growth
- Hobbies: sports (basketball, golf), true crime



A Little About You

- Hometown/country
- Graduate program or professional position
 - Background and research interests
- Why you want to learn DCMs and longitudinal DCMs

NCME 2024: Longitudinal DCM Workshop

DCM Preliminaries

LCDM Framework and Estimation in R

NCME 2024: Longitudinal DCM Workshop

Conceptual Example

- Suppose a basketball coach wants to test basic basketball ability
- Using traditional psychometric approaches, the coach could obtain an ability estimate for each player
 - » Classical test theory: obtain a test score (# items correct)
 - » Item response theory: obtain a latent ability estimate
- Both these approaches are normative in nature, only provide information on students' location on a continuum

NCME 2024: Longitudinal DCM Workshop

Conceptual Example

- Traditional testing procedures measure an **overall** ability in an area with a **continuous** latent variable:



Basketball Ability

NCME 2024: Longitudinal DCM Workshop

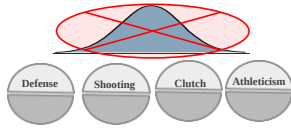
Conceptual Example

- Using traditional, norm referenced approaches, the result is a weak ordering
 - » Ordering is weak because of error in estimates
 - » Matt < LeBron < Jordan
- Questions that traditional psychometrics cannot answer:
 - » Why is Matt so low?
 - How do we help? What specific skills does he need to work on?
 - » How much ability is "enough" to pass?
 - What score is proficiency, or mastery?
 - » What skills have the players mastered, not mastered?

NCME 2024: Longitudinal DCM Workshop

Basketball Ability Example

- Instead of measuring an overall basketball ability, let's break basketball down into a set of skills or *attributes*:
 - Defense
 - Shooting
 - Clutch
 - Athleticism
- These attributes are categorical (not continuous).
- Often, they are dichotomous:
 - Model places examinees into two categories (e.g., Proficiency vs Non-proficiency)



NCME 2024: Longitudinal DCM Workshop

Diagnostic Classification Models

- DCMs use item responses to place students into groups according to proficiency or non- proficiency of multiple attributes

Player	Defense	Shooting	Clutch	Athleticism
	✓	✗	✗	✓
	✓	✗	✗	✓
	✓	✓	✓	✓

NCME 2024: Longitudinal DCM Workshop

Diagnostic Classification Models

- Questions from before:
 - » Why is Matt so low?
 - Because he can't shoot and he's not clutch. He needs to work on these skills.
 - » What skills are the players proficient, not proficient?
 - DCM provides these classifications directly.

NCME 2024: Longitudinal DCM Workshop

When are DCMs appropriate?

- Short answer: not always
- Successful implementation depends on:
 - 1) domain-specific theories/content experts.
 - What are the set of attributes we are trying to measure?
 - Can items be written to measure these attributes?
 - 2) alignment of purpose between assessment and model
 - Is classification the purpose?

NCME 2023: Advanced DCM Workshop

When are DCMs Not Appropriate?

- When you want to scale and desire a finer level of granularity
- DCMs do not distinguish within classes (limitation)
 - LeBron is more athletic than Matt, but they both possess the skill

Player	Defense	Shooting	Clutch	Attitude
	✓	✗	✗	✓
	✓	✗	✗	✓



NCME 2023: Advanced DCM Workshop

Conceptual Foundation Summary

- DCMs are psychometric models designed to classify
 - Diagnostic and criterion-referenced interpretations
- DCMs provide valuable information with more feasible data demands than other psychometric models
 - Higher reliability than IRT/MIRT models
 - Naturally accommodate multidimensionality
 - Complex items structures possible
 - Alignment of assessment goals and psychometric model

NCME 2023: Advanced DCM Workshop

Questions? Comments?



NCME 2023: Advanced DCM Workshop

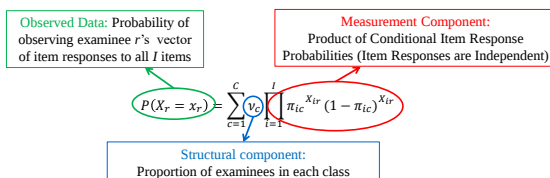
Statistical Foundation

- Latent class models use responses to probabilistically put respondents into latent classes
- DCMs are confirmatory latent class models
 - » Latent classes specified a priori as attribute profiles
 - » Q-matrix specifies item-attribute structure (confirmatory)
 - » Person parameters are attribute proficiency probabilities

NCME 2024: Longitudinal DCM Workshop

DCMs as Latent Class Models

- DCMs are constrained latent class models



NCME 2024: Longitudinal DCM Workshop

Statistical Foundation

- Latent class models (LCMs) use responses to probabilistically put respondents into latent classes
- DCMs are confirmatory latent class models
 - » Latent classes specified a priori as attribute profiles
 - » Q-matrix specifies item-attribute structure (confirmatory)
 - » Person parameters are attribute proficiency probabilities
- DCMs are the categorical latent trait analogue of IRT models
 - » Item parameters model relationships between attribute proficiency and responses

NCME 2024: Longitudinal DCM Workshop

Item Response Function Basics

- 2-PL item response function
 - » Discrimination (a)
 - » Difficulty (b)
 - » Ability (θ)
 - DCM
 - » Intercept (λ_0)
 - » Main effects (λ_1)
 - » Attribute proficiency status (α)
- $$\text{logit}(X_i = 1) = a_i(\theta - b_i)$$

$$= \frac{a_i\theta - a_ib_i}{\psi_i + \delta_i\theta}$$
- $$\text{logit}(X_i = 1) = \lambda_{i,0} + \lambda_{i,1}\alpha$$
- α is either 0 or 1
 - λ_0 is log-odds of a correct response for non-proficient
 - λ_1 is the increase in log-odds of a correct response for proficient

NCME 2024: Longitudinal DCM Workshop

Development of DCMs

- Over the past several years, numerous DCMs have been developed
- Each DCM makes assumptions about how mastered attributes combine/interact to produce an item response
- We're going to focus on one, the Log-linear Cognitive Diagnosis Model (LCDM; Henson, Templin, & Willse, 2009)

NCME 2024: Longitudinal DCM Workshop

LCDM Multi-attribute Items

- Simple structure item: $\text{logit}(X_i = 1) = \lambda_{i,0} + \lambda_{i,1}\alpha$
- Complex item: multiple attributes being measured
 - » Let's say an item measures Attributes 1 and 2
 - » Parameters are like a reference coded two-way ANOVA model

$$\text{logit}(X_i = 1) = \lambda_{i,0} + \lambda_{i,1(1)}\alpha_1 + \lambda_{i,1(2)}\alpha_2 + \lambda_{i,2(12)}\alpha_1\alpha_2$$

Intercept
Main Effect (Attribute 1)
Main Effect (Attribute 2)
Interaction (Between Attributes 1 & 2)

NCME 2024: Longitudinal DCM Workshop

LCDM Interaction Plots

- Play with excel file: [“Example LCDM Items.xlsx”](#)
- Manipulate item parameters and see how logits and item response probabilities are affected

NCME 2024: Longitudinal DCM Workshop

Statistical Foundations Summary

- DCMs are confirmatory latent class models
- DCMs are categorical latent trait IRT models
- Person parameters (α): attribute proficiency probabilities
- Item parameters (λ): proficiency ~ responses
- LCDM item response function
 - Logistic regression with binary predictors
 - Reference-coded ANOVA model
 - Parameters indicate increases in probability correct for different proficiency statuses
 - Many DCMs that place constraints on item parameters

NCME 2024: Longitudinal DCM Workshop

DCM Inputs and Outputs

- Inputs:
 - Item response data ($N \times I$ matrix)
 - Q-matrix ($I \times A$ matrix)
 - Type of DCM you want to estimate, plus options
- Outputs
 - Item parameter estimates
 - Person parameter estimates (proficiency probabilities)
 - Attribute parameter estimates (proficiency proportions, correlations)
 - Model fit statistics

NCME 2024: Longitudinal DCM Workshop

Estimation in R

- Several packages estimate DCMs
 - **CDM**
 - GDINA
 - mirt
 - blatant
 - hmcdrm
 - npcd
 - PACER
- Other software programs as well
 - Mplus, flexMIRT, Latent Gold, mdlm

NCME 2024: Longitudinal DCM Workshop

Estimation in R

- In the Examples folder, open “example0.R”
- Two attributes, 10 items, $N = 1000$

NCME 2024: Longitudinal DCM Workshop

Break

- Take a quick 5-minute break
- Next up: Longitudinal DCMs I



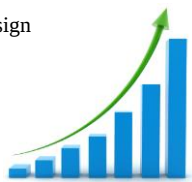
NCME 2024: Longitudinal DCM Workshop

Longitudinal DCMs I (single group)

NCME 2024: Longitudinal DCM Workshop

Assessing Growth in Education

- Recent shift in educational policy
 - » Meeting state standards → measuring student progress
- **Need to quantify learning**
- Single-group pre-test/post-test design
 - » Examinees take pre-test
 - » Intervention (e.g., instruction)
 - » Examinees take post-test
- Goal: assess growth in students



NCME 2024: Longitudinal DCM Workshop

Growth and DCMs

- Diagnostic classification models (DCMs)
 - » Classify on categorical latent traits: *attributes*
 - » Attribute mastery or non-mastery
- Growth in a DCM framework
 - » Individual: non-mastery $\leftrightarrow$ mastery
 - » Group: 30% at pre-test $\rightarrow$ 80% at post-test
 - » Previous work modeled testing occasions separately

NCME 2024: Longitudinal DCM Workshop

Latent Class Models

- DCMs are constrained latent class models

Observed Data: Probability of the observed item responses

Measurement Component:
Models how latent classes respond to items

$$\mathbb{P}(X_r = \mathbf{x}_r) = \sum_{c=1}^C v_c \prod_{i=1}^I \pi_{ic}^{x_{ir}} (1 - \pi_{ic})^{1-x_{ir}}$$

Structural component:
Models how classes are related, class proportions

NCME 2024: Longitudinal DCM Workshop

Latent Class Models

- There are longitudinal latent class models
 - » Latent transition analysis (LTA)

$$\mathbb{P}(X_r = \mathbf{x}_r) = \sum_{c_1=1}^C \sum_{c_2=1}^C \dots \sum_{c_T=1}^C v_{c_1} \tau_{c_2|c_1} \tau_{c_3|c_2} \dots \tau_{c_T|c_{T-1}} \prod_{t=1}^T \prod_{i=1}^I \pi_{ic_t}^{x_{rit}} (1 - \pi_{ic_t})^{1-x_{rit}}$$

Structural component:
Proportion of examinees in each class at Time Point 1

Transition Probabilities:
Probabilities of transition between latent classes from time to time

Measurement Component:
Product of Conditional Item Response Probabilities (Across time points)

NCME 2024: Longitudinal DCM Workshop

Transition DCM

- Transition Diagnostic Classification Model (TDCM)
 - » Combines LTA with the LCDM as measurement model
 - » General form (T time points):

$$\mathbb{P}(X_r = x_r) = \sum_{c_1=1}^C \sum_{c_2=1}^C \cdots \sum_{c_T=1}^C v_{c_1} \tau_{c_2|c_1} \tau_{c_3|c_2} \cdots \tau_{c_T|c_{T-1}} \prod_{t=1}^T \prod_{i=1}^I \pi_{ic_t}^{x_{rit}} (1 - \pi_{ic_t})^{1-x_{rit}}$$

- » Pre-test/post-test form (2 times points w/ invariance):

$$\mathbb{P}(X_r = x_r) = \sum_{c_1=1}^C \sum_{c_2=1}^C v_{c_1} \tau_{c_2|c_1} \prod_{t=1}^2 \prod_{i=1}^I \pi_{ic_t}^{x_{rit}} (1 - \pi_{ic_t})^{1-x_{rit}}$$

NCME 2024: Longitudinal DCM Workshop

Measurement Invariance

- Reflects the notion that measurement is happening the same way across groups or over time
- In a longitudinal psychometric analysis, person estimates should be determined the exact same way
- For the TDCM, we want that proficiency is being determined the same way over time
 - » Otherwise, change in proficiency status is meaningless

$$\mathbb{P}(X_r = x_r) = \sum_{c_1=1}^C \sum_{c_2=1}^C v_{c_1} \tau_{c_2|c_1} \prod_{t=1}^2 \prod_{i=1}^I \pi_{ic_t}^{x_{rit}} (1 - \pi_{ic_t})^{1-x_{rit}}$$

NCME 2024: Longitudinal DCM Workshop

Empirical Data Analysis

- Mathematics assessment measuring four attributes
 - » Ratios and proportional relationships
 - » Measurement and data
 - » Number system (fractions)
 - » Geometry (graphing)
- Sample of 879 middle school students
- Same 21 items at pre- and post-test

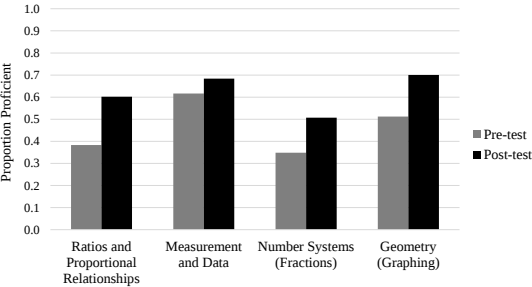
NCME 2024: Longitudinal DCM Workshop

Data Analysis Results

- Item parameters:
 - » Median intercept: -1.82 (14% chance of correct response)
 - » Median discrimination = .55 (masters – non-masters)
- Attribute Correlations:
 - » Pre-test: ranged from .58 to .84
 - » Post-test: ranged from .73 to .90
 - » Indicated multidimensionality

NCME 2024: Longitudinal DCM Workshop

Data Analysis Results



NCME 2024: Longitudinal DCM Workshop

Data Analysis Results

Attribute	Pre-test	Post-test	Growth	p-value	Odds Ratio
Ratios and proportional relationships	.384	.602	.219	<.001	2.43 (medium)
Measurement and data	.616	.684	.068	<.001	1.35 (small)
Numbers and operations (fractions)	.348	.507	.159	<.001	1.92 (medium)
Geometry (graphing)	.512	.701	.189	<.001	2.23 (medium)

NCME 2024: Longitudinal DCM Workshop

Data Analysis Results

- Individual Attribute Transition Matrices

Ratios and proportional relationships		
	Non-proficiency	Proficiency
Non-proficiency	.52	.48
Proficiency	.21	.79

Measurement and data		
	Non-proficiency	Proficiency
Non-proficiency	.59	.41
Proficiency	.31	.69

Numbers and operations (fractions)		
	Non-mastery	Mastery
Non-proficiency	.57	.43
Proficiency	.16	.84

Geometry (graphing)		
	Non-proficiency	Proficiency
Non-proficiency	.42	.58
Proficiency	.19	.82

NCME 2024: Longitudinal DCM Workshop

Estimation in R

- In the Examples folder, open “example1.R”
- Four attributes, 10 items, N = 1500

NCME 2024: Longitudinal DCM Workshop

Estimation in R

A group of biology educators were interested in the progression of students’ proficiency in foundational biology concepts (mitosis and meiosis) using an innovative blended learning instructional approach. They developed three forms of a test measuring both concepts and administered the test at three different points during the semester. Each form had ten items, with the forms sharing three common items (Items 3, 5, 7). They had a total of 875 Biology 101 students take all three tests.

- In the Examples folder, open “example2.R”
- Two attributes, 10 items, N = 875
- Different Q-matrices at each time point
- Items 3, 5, 7 are common items across the tests

NCME 2024: Longitudinal DCM Workshop

Longitudinal DCMs II (multigroup)

NCME 2024: Longitudinal DCM Workshop

TDCM Applied Study

- National Assessment of Educational Progress, 2011
 - » 23% of students w/o disabilities scored below *basic*
 - » 65% of students with disabilities scored below *basic*
- Basic: operations with rational numbers and real world problem solving
- Research suggests that students with disabilities struggle with problems embedded in text
- University of Kentucky (Bottge, et. al) focus on improving mathematics instruction and assessment for students with disabilities

NCME 2024: Longitudinal DCM Workshop

Enhanced Anchored Instruction

- Acronym (EAI)
- Instructional method developed to improve problem solving
- Realistic problems are embedded in interesting contexts
 - » 8-15 minute videos
 - » Hands-on applications
- Include modules and interactive media
 - » Help students visualize abstract concepts

NCME 2024: Longitudinal DCM Workshop

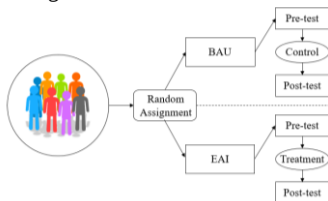
EAI Study Design

- Overall research questions:
 - » To what extent does EAI improve students' problem solving abilities?
 - » How well does it work compared to traditional instruction?
- Team developed an assessment targeting multiple areas
 - » 21-item problem-solving test
 - » Measured four Common Core *strands*
 - Ratios and proportional relationships
 - Measurement and data
 - Numbers systems (fractions)
 - Geometry (graphing)

NCME 2024: Longitudinal DCM Workshop

EAI Study Design

- Business as Usual (BAU, $N = 456$) vs. EAI ($N = 423$)
- Sample of 6-8 graders



- Is EAI better than traditional instruction?

NCME 2024: Longitudinal DCM Workshop

EAI Study Design

- Items not initially designed with a DCM in mind
 - » But since items were prospectively written for each strand, well-suited for a DCM analysis
- Methodological research questions:
 - » How do we analyze growth with a general DCM?
 - Prompted development of TDCM
 - » How do we analyze intervention effects in a DCM?
 - Multiple-group TDCM

NCME 2024: Longitudinal DCM Workshop

Analysis

- Evaluate intervention effect
 - » Use multiple-group TDCM
 - » Analyze growth for BAU and EAI
 - » Compare growth
- How does growth for EAI students compare to growth for BAU students?
 - » For which attributes does EAI appear to be better than BAU?
 - » Where can EAI be improved?

NCME 2024: Longitudinal DCM Workshop

Analysis

- To analyze the intervention effect, extend TDCM to multiple groups
 - » Attribute proficiency proportions and transitions depend on group membership (g)

$$\mathbb{P}(X_r = x_r | \mathcal{G} = g) = \sum_{c_1=1}^C \sum_{c_2=1}^C \underbrace{v_{c_1|g} \tau_{c_2|c_1,g}}_{\substack{\text{initial} \\ \text{transition}}} \prod_{t=1}^2 \prod_{l=1}^I \pi_{lc}^{x_{rit}} (1 - \pi_{lc})^{1-x_{rit}}$$

NCME 2024: Longitudinal DCM Workshop

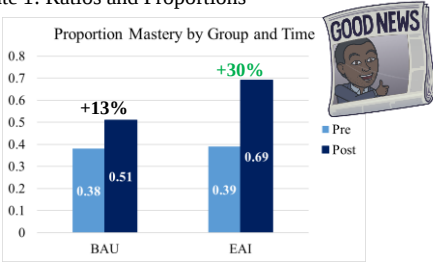
Analysis

- Let δ_{g_1} and δ_{g_2} be the attribute mastery growth for both groups, respective
- To test for differential growth, can use a Wald test:
$$z = \frac{|\delta_{g_1} - \delta_{g_2}|}{se(\delta_{g_1} - \delta_{g_2})},$$
 - z follows a standard normal distribution
 - Similar to a “difference-in-difference” test
 - Use Hochberg (1988) multiple tests correction to control type I error rates

NCME 2024: Longitudinal DCM Workshop

Results: Differential Growth

• Attribute 1: Ratios and Proportions



NCME 2024: Longitudinal DCM Workshop

Results: Differential Growth

Attribute	Growth (EAI)	Growth (BAU)	Difference	p-value
Ratios and proportional relationships	.31	.13	.18 (.05)	<.001
Measurement and data	.07	.07	.00 (.04)	.929
Number Systems (fractions)	.18	.13	.05 (.04)	.597
Geometry (graphing)	.26	.11	.15 (.05)	.008

NCME 2024: Longitudinal DCM Workshop

Results: Conditional Growth

- Proficiency transitional probabilities
 - » Were EAI non-proficients more likely to transition to proficiency than BAU non-proficients?
 - » Ratios and proportional relationships

BAU	
Post-test	
	0 1
Pre-test	0
	1
0	.63 .37
1	.27 .73

EAI	
Post-test	
	0 1
Pre-test	0
	1
0	.41 .59
1	.14 .86

NCME 2024: Longitudinal DCM Workshop

Alternative Analysis

- Can use gain scores and repeated measures MANOVA to analyze this data
 - » Interaction effects test differential growth
- Could also use a longitudinal IRT model
- Why use a DCM?
 - » Two main reasons

NCME 2024: Longitudinal DCM Workshop

Interpretive Value

- Effect size measures
 - » TDCM: difference in groups' growth in attribute mastery
 - » Gain scores: average difference in groups' gain scores
 - » Longitudinal IRT: difference in groups' growth in ability
- Effect sizes:

Attribute	TDCM	Gain Scores	Long IRT
Ratios and proportions	.18 (72)	.49	.51



NCME 2024: Longitudinal DCM Workshop

Reliability

- Refers to the consistency of the growth estimates
 - » Are they trustworthy?

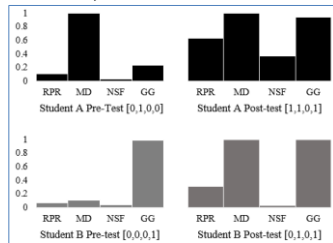
Attribute	TDCM	Gain Scores	Long IRT
Ratios and proportions	.915	.488	.550
Measurement and data	.968	.651	.566
Number Systems	.928	.603	.631
Geometry (graphing)	.911	.619	.638



NCME 2024: Longitudinal DCM Workshop

Interpretive Value

- Individual estimates
- Two examinees (Student A and B):
 - » Pre-test = 9/21
 - » Post-test = 12/21
 - » Gain Score = +3



NCME 2024: Longitudinal DCM Workshop

Conclusions

- Analyzed data from a randomized controlled trial
- Evaluated efficacy of EAI
 - » Used TDCM, which has benefits over traditional analysis
 - » Analysis suggests EAI better than traditional instruction
- Overall, EAI was successful
 - » But...could be improved for number systems and measurement and data
- The TDCM provides a meaningful measure of growth



NCME 2024: Longitudinal DCM Workshop

Estimation in R

The same group of biology education researchers wanted to gauge the degree to which the blended learning approach could be successful for other foundational biological concepts. To evaluate the efficacy of the blended learning approach, they conducted a quasi-experiment where instructors (and the students in their respective classes) were either assigned to the control group (traditional instruction; $N = 1057$) or the treatment group (blended learning; $N = 1104$). The pre- and post-test were identical and consisted of 20 items measuring four attributes (respiration, reproduction, DNA structure, protein synthesis). The test was administered at the beginning and end of the semester.

- In the Examples folder, open “example3.R”
- Four attributes, 20 items
- Two time points
- Two groups: control (1057) and treatment (1104)

NCME 2024: Longitudinal DCM Workshop

Break

- Take a quick 5-minute break
- Next up: Q&A and discussion



NCME 2024: Longitudinal DCM Workshop

Wrapping Up

- We discussed longitudinal DCM advancements
 - TDCM is an extension that combines LTA with the LCDM
 - Provides criterion-referenced estimates of growth
 - Can be more meaningful and interpretable
- All estimated in R
 - Much simpler than Mplus or other software

NCME 2024: Longitudinal DCM Workshop

Questions/Discussion

- Open forum for questions and discussion regarding:
 - » Experiences with DCMs items, tests, models, estimation, software, etc.
 - » Ideas within additional material
 - Further explanation/questions
 - » Ideas outside of this workshop material so far
 - Is there a topic you had hoped we would include?



NCME 2024: Longitudinal DCM Workshop

Questions/Discussion

- Other questions?
 - » Email me: mjmadison@uga.edu
 - » **Join DCMNET** (www.matthewmadison.com)
 - Network for DCM research and practice ($\approx$ SEMNET)
 - Access to DCM experts
 - » Join the diagnostic measurement SIGMIE
 - » Access the DCM [database](#)



NCME 2024: Longitudinal DCM Workshop

Last Favor

- TDCM package is under revision
 - » This is the first workshop where we have used it
- Would love your feedback
 - » Click [here](#) for a brief feedback form



NCME 2024: Longitudinal DCM Workshop
