Innovative approaches to delirium measurement

Richard N. Jones, Sc.D.

Delirium in Older Adults: Finding Order in the Disorder
NIA/AGS Bench to Bedside Conference
February 9-11, 2014, Bethesda, MD
February 11, 2013

rich_jones@brown.edu

Disclosures

- **Current funding**:

  R01AG030618; P01AG031720;
  R03AG045494; U01AR057954

- **Other financial relationships**:

  None to report.

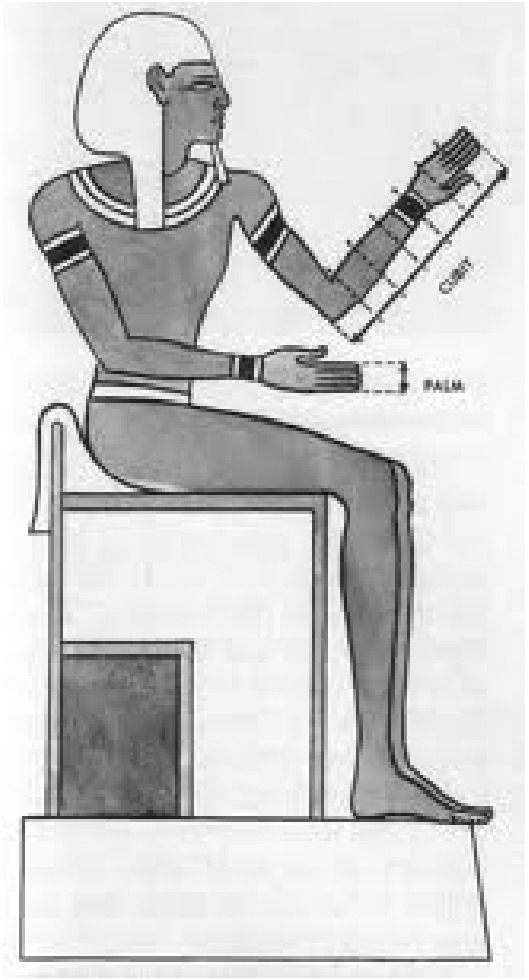
- **Conflicts of interest**:

  None to report.

Measurement

- The process of assigning numbers to observations
- Allows use of tools of math on subjective observations
- Fundamental for
  - Objectivity
  - Reproducibility
  - Validity
  - Progress of science

Measurements are constructed



http://abemkemet.blogspot.com/2012/06/ruler-for-ruler.html

- People construct measures
- Measurement, *like all of science*, is a <u>social phenomena</u>
  - What measures are used
  - How measures are interpreted


Icon Key
CAM
DRS
North Dakota
Minnesota
South Dakota
Wyoming
Nebraska
Iowa
Wisconsin
Michigan
Illinois
Indiana
Ohio
Kansas
Missouri
Kentucky
West Virginia
Virginia
Oklahoma
Arkansas
Mississippi
Alabama
Georgia
South Carolina
North Carolina
Tennessee
Texas
Louisiana
Colorado
New Mexico
United States
Kansas City
St. Louis
Columbus
Indianapolis
Denver
Albuquerque
Ciudad Juarez
Fort Worth
Dallas
Austin
Houston
San Antonio
Chihuahua
Monterrey
Gulf of
Jacksonville
Orlando
Tampa
Florida
Miami
Charlotte
Atlanta
Toronto
Ottawa
Montreal
Quebec City
New Brunswick
Maine
Vermont
New Hampshire
Massachusetts
Rhode Island
Connecticut
New York
Pennsylvania
Philadelphia
New Jersey
Maryland
Delaware
District of Columbia
Bermuda


Icon Key
CAM
DRS
Bergen
Oslo
Stockholm
Turku
Helsinki
St Petersburg
Tallinn
Estonia
Baltic Sea
Gothenburg
Riga
Latvia
North Sea
Denmark
Copenhagen
Lithuania
Kaunas
Moscow
Gdańsk
Vilnius
Minsk
United Kingdom
Isle of Man
Leeds
Dublin
Ireland
Liverpool
Sheffield
Hamburg
Bremen
Berlin
Bialystok
Belarus
Poland
Amsterdam
Netherlands
Poznan
Warsaw
Wroclaw
Lublin
Voronezh
Bristol
Cologne
Germany
Prague
Krakow
Kyiv
Kharkiv
Luxembourg
Czech Republic
Paris
Munich
Slovakia
Bratislava
Ukraine
Dnipropetrovs'k
Luhansk
Budapest
Zaporizhzhya
Donetsk
Zurich
Austria
Moldova
Nantes
Graz
Hungary
Cluj-Napoca
Chisinau
Odesa
Mykolaiv
Rostov-on-Do
France
Switzerland
Geneva
Lyon
Slovenia
Romania
Bay of Biscay
Zagreb
Turin
Croatia
Bucharest
Sevastopol'
Bosnia and Herzegovina
Serbia
Toulouse
Monaco
Italy
Marseille
Black Sea
Bilbao
Sofia
Bulgaria
Kosovo
Rome
Montenegro
Zaragoza
Macedonia (FYROM)
Plovdiv
Porto
Bari
Tirana
Istanbul
Madrid
Naples
Albania
Thessaloniki
Palma de Mallorca
Bursa
Portugal
Spain
Lisbon
Greece
Izmir
Turkey
Palermo
Athens
Konya
Catania
Seville
Antalya
Algiers
Tunis

Measurement in medicine

- Source of information
  - People
    - Patients, informants, clinicians
    - Verbal reports, performance tests, clinical exam
  - Tools, devices
    - Driven by advances technology and
    - Biological knowledge

Measurement in psychiatry

- Main goals
  - Case identification
  - Identification of risk factors
  - Severity measurement
  - Outcome assessment
- Main approaches
  - Self-report
  - Expert rating
  - Clinical observation
  - Biomarkers (imaging, function, biochemistry)

State-of-the-art in delirium measurement

- Case identification
  - Diagnostic criteria (e.g., CAM, DSM, ICD)
  - Count and cut (e.g., DRS, MDAS)
- Severity measurement
  - Count symptoms (e.g., DRS, MDAS)
- Outcome assessment
  - OPM (other people's measures; e.g., MMSE)

Please see this excellent systematic review:

Hjermstad, M. J., Loge, J. H., & Kaasa, S. (2004). Methods for assessment of cognitive failure and delirium in palliative care patients: implications for practice and research. *Palliative Medicine, 18(6), 494-506.*

An innovation in measurement

- Item response theory (IRT)
- Undergirds major NIH-funded measurement initiatives
  - PROMIS
  - NIH Toolbox
- Relatively new approach (1952)
- Computationally intensive (slows uptake)
- But, really, just a generalized linear mixed effect model like most everything else

Item response theory (IRT)

- Statistical model that relates
  - Responses or observations of patients, to
  - Theoretical underlying quantities
- A family of statistical models
  - Not just one model
- Invented in fields of educational psychology and psychological measurement
- AKA *latent trait theory*
- Widely applied in health and psychiatry

Original Article

Which Mini-Mental State Exam Items Can Be Used to Screen for Delirium and Cognitive Impairment?

Peter M. Fayers, PhD, Stein Kaasa, MD, and Jon H. Loge, MD, PhD
Department of Public Health and Clinical Research (P.M.F.), University of Aberdeen, Aberdeen, Scotland; Norwegian University of Science and Technology and Science, Trondheim, Norway; The Norwegian Cancer Society (M.J.H), Oslo, Norway; Department of Behavioral Sciences in Medicine (M.J.H., J.H.L.), University of Oslo, Oslo, Norway; Research Group of Geriatric Medicine (A.H.R., L.S.), and Oncological Department (J.H.L.), Ullevål University Hospital HF, Oslo, Norway; and Palliative Medicine Unit (S.K.), Department of Oncology, and Department of Anesthesiology (P.K.), St. Olav's Hospital HF, Trondheim, Norway

- **Used IRT to identify the best MMSE items to screen for delirium**
- **3 item test was best (year, date, backward spelling)**
- **Limitation: two data sets, in 1 delirium was defined as MMSE < 24**

Original Research Article

- Used IRT to evaluate measurement properties of MMSE
- Used logistic regression to devise short form
- Short form performed poorly relative to full MMSE

Yang et al. BMC Medical Research Methodology 2013, 13:8
http://www.biomedcentral.com/1471-2288/13/8

BMC
Medical Research Methodology

RESEARCH ARTICLE Open Access

Selecting optimal screening items for delirium: an application of item response theory

Frances M Yang[1*], Richard N Jones[2], Sharon K Inouye[3], Douglas Tommet[1], Paul K Crane[4], James A Rudolph[5], Long H Ngo[4] and Edward R Marcantonio[1]

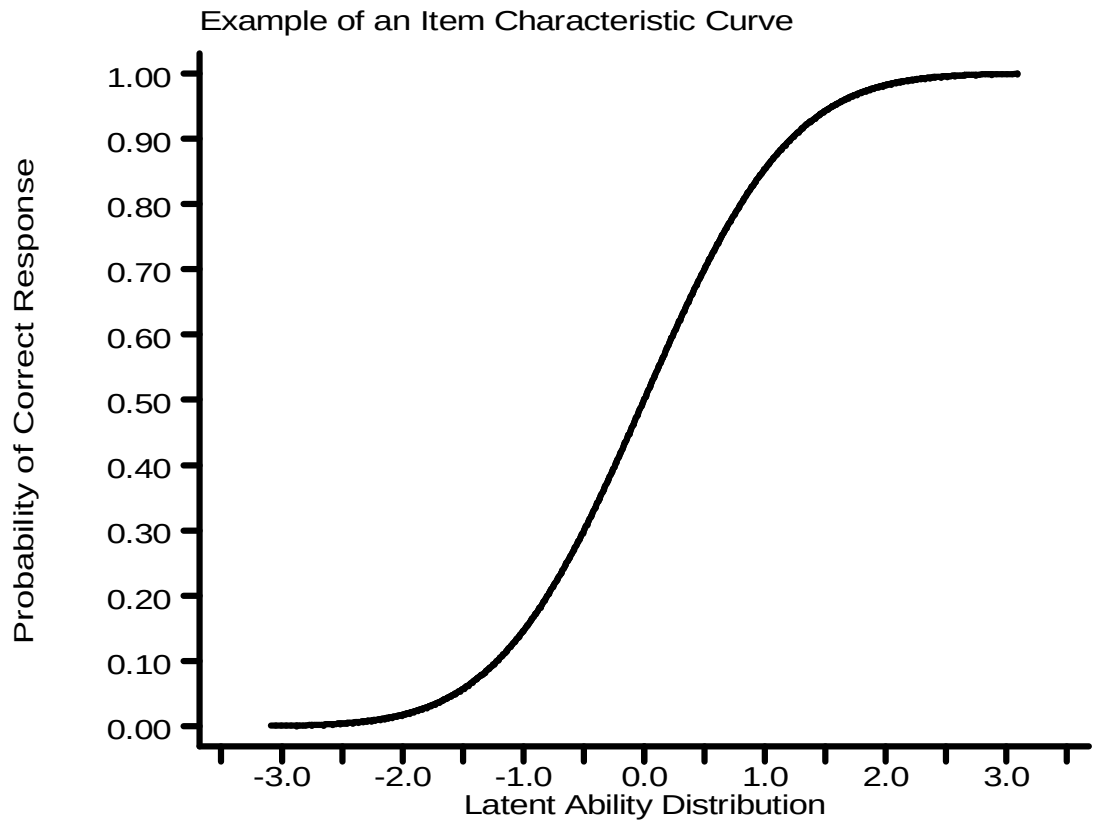
- **Use IRT to identify optimal screening items for CAM ratings of delirium**
  - **From among MMSE, MDAS, etc.**
- **Limitations**
  - **No validation**
  - **No instrument developed**

Why is IRT important?

- Can be used to harmonize different tools
- Can be used to generate new tests that *simultaneously* optimize
  - measurement precision and
  - respondent burden
- Outcome measures with interval measurement properties (i.e., study change)
- Can be extended to include models for mixtures of clinical populations
  - Separate the merely demented from the delirious

Item Response Function

$$P(y_{ij}=1|\theta_i) = F[a_j(\theta_i - b_j)]$$



Bayesian estimates of underlying trait

Bayes modal estimates of latent ability (θ)
(modal a posteriori [MAP] estimates)

likelihood function for response pattern U given ability θ:

$$g(U|\theta) = \prod_{i}^{p} P_i^{Y_i} Q_i^{1-Y_i}$$

a posteriori likelihood function of η given pattern U:

$$g(\theta|U) = \frac{\phi(\theta) g(U|\theta)}{g(U)}$$

Image: www.ouedkniss.com

How is IRT relevant to delirium research?

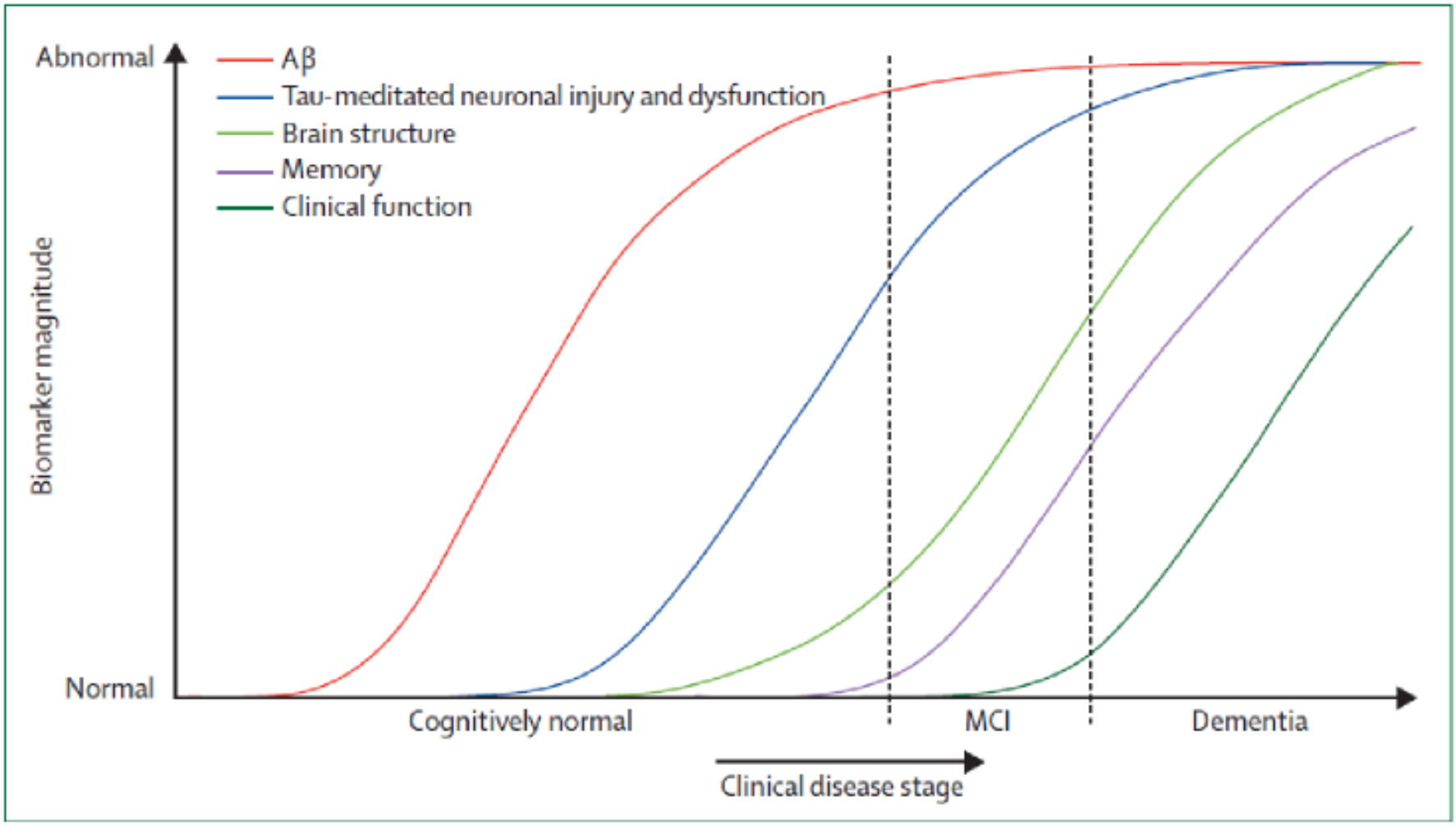


Figure 2: **Dynamic biomarkers of the Alzheimer's pathological cascade**
Aβ is identified by CSF $A\beta_{42}$ or PET amyloid imaging. Tau-mediated neuronal injury and dysfunction is identified by CSF tau or fluorodeoxyglucose-PET. Brain structure is measured by use of structural MRI. Aβ=β-amyloid. MCI=mild cognitive impairment.

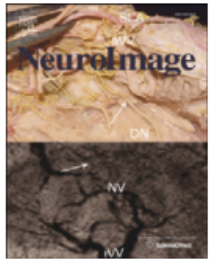
Jack *et al. (2010) Lancet Neurology* ***9:119-28.***

NeuroImage 63 (2012) 1478–1486

Contents lists available at SciVerse ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg



A computational neurodegenerative disease progression score: Method and results with the Alzheimer's disease neuroimaging initiative cohort

Bruno M. Jedynak [a,b,*], Andrew Lang [c], Bo Liu [a], Elyse Katz [d], Yanwei Zhang [d], Bradley T. Wyman [d], David Raunig [d,1], C. Pierre Jedynak [e], Brian Caffo [f], Jerry L. Prince [c]
for the Alzheimer's Disease Neuroimaging Initiative [2]

[a] *Department of Applied Math and Statistics, Johns Hopkins University, Baltimore, MD 21218, USA*

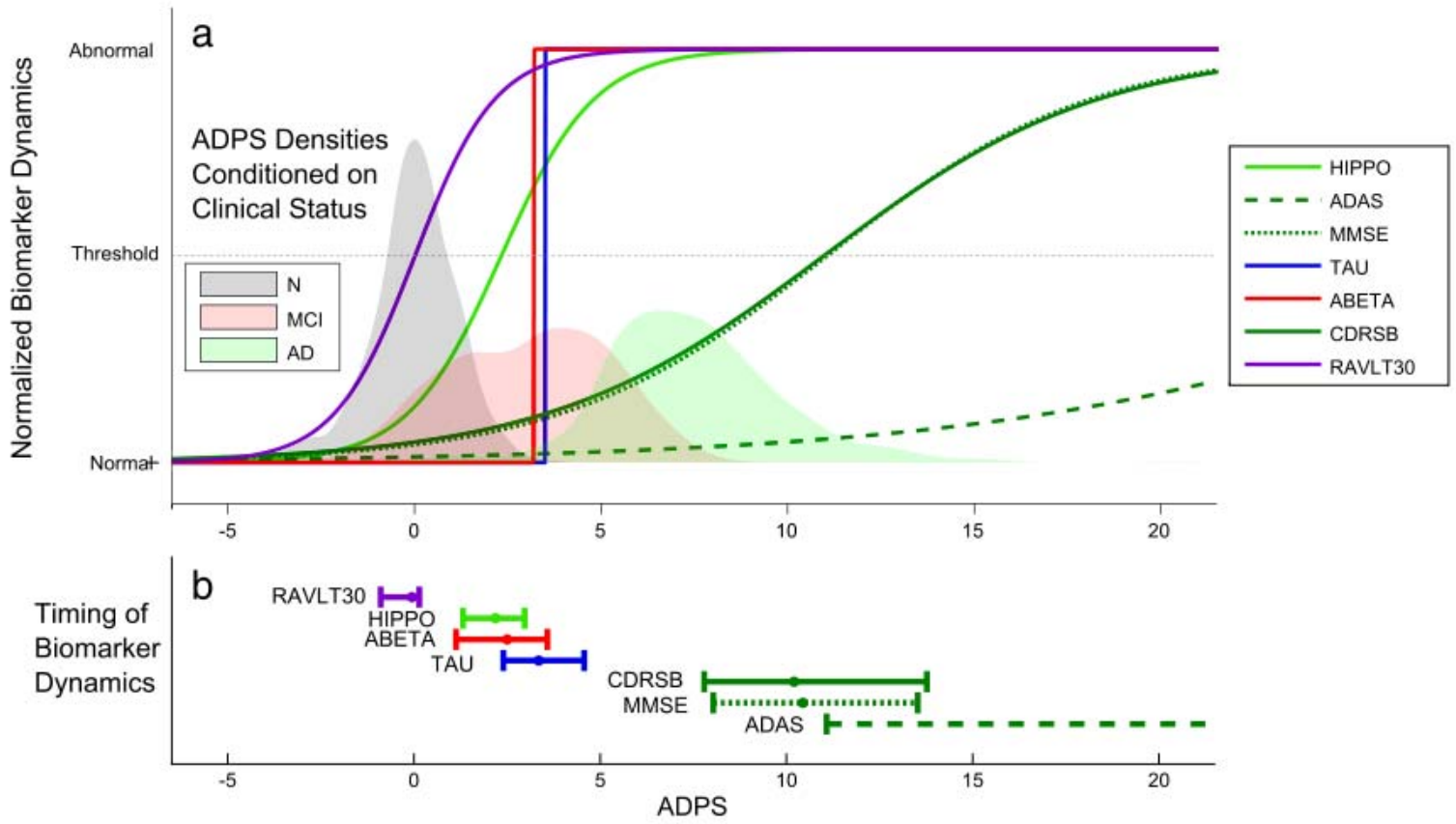


Fig. 5. (a) Estimated biomarker dynamics as a function of the normalized ADPS. Estimation of the normalized ADPS for all ADNI subjects was carried out, and common biomarker dynamics represented by sigmoidal functions were simultaneously fitted as part of the ADPS normalization algorithm. Each sigmoidal function was scaled and flipped in order to fit on a scale going from -1 representing "Normal" to 1 representing "Abnormal". The positions of vertical lines representing progression from Normal to MCI and MCI to AD were fitted as optimal separating thresholds between the clinical diagnoses provided in the ADNI database. (b) 90% confidence intervals for the inflection point of each biomarker.

Jedynak *et al. (2012) Neuroimage.*

Future innovations in delirium measurement

Measurement Models for Nominal Indicators

Opportunities

- Improved and harmonized measures
- More power
- Using measurement to test theories
  - etiology and pathogenesis
  - Mechanisms of risk and protective factors

Conclusion

- Measurement is important and complicated
- Progress requires collaboration
  - Clinical experts
  - Measurement experts

OPRPHAN SLIDES

Delirium: Class vs Continuum

- Clinical experts argue delirium is a spectrum disorder
- But, sx measurements are sometimes nominal
  - absent
  - present not fluctuating not acute
  - present and (fluctuating and/or acute)

Delirium as Subtype of Impairmemt

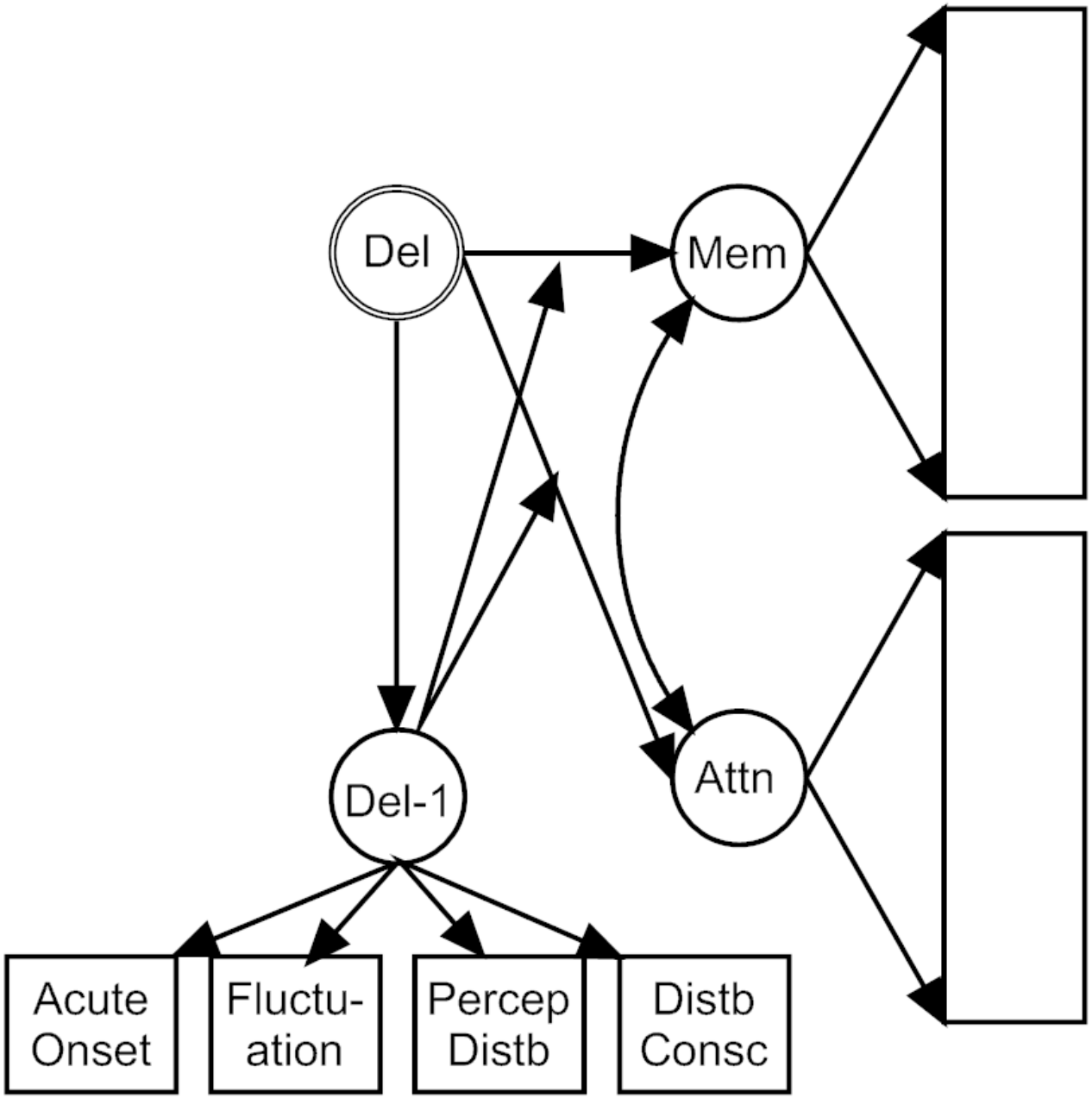
- Cognition (attention, memory) is a continuum (high fxn←→sev imp)
- Delirious are a subpopulation of those impaired
- Subpopulation indicators:
  - Disturbance of Consciousness (awareness of surroundings)
  - Sx show fluctuation (each? any?)
  - Sx show acute onset (each? any?)
  - Presence of perceptual disturbance
  - Level of impairment

Can this be modeled?

- Linear Latent Variable Models
  - Mixture measurement model (IRT Mixture Model)
  - Assumes population traits (attention, memory) are continuous normal
- Bayesian modeling
  - More flexibility in latent trait distributions
- Examples
  - Lubke and Muthen, 2005 *Psychol Methods* ***10:21-39***
  - Lubke and Neale, 2006 *Multivariate Behav Res* ***41:499***
  - Muthén and Asparouhov, 2006 *Addict Behav* ***31:1050-66***

Corollary

- Better measurement in science will lead to work that is
  - More objective
  - More reproducible
  - More validity
  - Accelerate progress


Del
Mem
Del-1
Attn
Acute Onset
Fluctu-ation
Percep Distb
Distb Consc

Why would we do this?

- Theory-Data-Model Fit
  - Are we forcing delirium into linear factor analysis (item response theory) box?
  - If yes, should we?
- What would be the benefit of such modeling exercises?
  - Models contribute to understanding of disorder (endo-) phenotypes
    - Clarify delirium sub-types
      - Yang et al., 2008 *Psychosomatics* ***50:248-54***
    - Probe meaning of 'severity' of delirium
  - Sufficiently accurate models can be used as screening tools (provide weights for criteria)
    - Move beyond "count and cut" screening