Using Machine Learning to Detect Noncredible Cognitive Test Performance

John-Christopher A. Finley, Anthony D. Robinson, Jason R. Soble, & Violeta J. Rodriguez

Doctoral Candidate, Northwestern University Feinberg School of Medicine

Predoctoral Intern, Emory University School of Medicine

Agenda

- No Disclosures
- Existing Research
- Scope of Study
- Key Findings
- Interpretation of Findings
- Limitations & Future Directions

Validity Assessment Standards

Use multiple performance validity tests

Apply appropriately adjusted cutoffs

Avoid redundancy

Consider contextual, diagnostic, intrapersonal factors

Challenges with Current Standards

But accounting for all these factors can be time-consuming and burdensome

….this in turn may limit the frequency and quality of validity assessment

Algorithms May Help Address These Challenges

Algorithmic methods may improve the assessment process

Unsupervised machine learning is one such method that shows promise

Why Unsupervised Machine Learning?

Efficient	Rapid, automated pattern recognition
Personalized	Synthesizes complex, messy, multidimensional data
Objective	Data-driven, does not require ground truth

….these aspects may help with validity assessment challenges

Current Study

Unsupervised machine learning can help in many ways

But first, we wanted to know –
can unsupervised machine learning render similar conclusions as the standard validity approach?

Method

Obtained Neuropsychological Test Data from Clinical Patients

- 308 outpatient adults with mixed diagnoses
- Most in their 40s, 55% women, 13 educational years, diverse racial identities

Rated their Data as Valid (81%) or Invalid (19%) Using 2-Failure Rule

- Invalid = ≥2 failures, Valid = ≤1 failures
- TOMM Trial 1; MSVT (CNS, DR, IR); Dot Counting Test; Word Choice Test; Reliable Digit Span; RAVLT Forced Choice

Gave Unsupervised Machine Learning Model Various Data to Analyze

- Performance validity tests scores + total failures
- Contextual factors (disability or litigation seeking)
- Diagnostic factors (12 diagnoses)
- Number of diagnoses most associated with invalid performance

Did not Tell Model How to Group Data & Compared to Validity Ratings

- Used SidClustering Technique

Results:

Model Found Two Groups with 92% accuracy (95% CI: 88-97%)

	Group 1: Valid (81%)	Group 2: Invalid (19%)
Contextual		
Disability / Compensation Seeking	27 (11%)	17 (28%)
Litigation Seeking	8 (3%)	1 (2%)
Diagnostic		
Internalizing Psychiatric Disorder	40 (16%)	15 (25%)
Attention-Deficit/Hyperactivity Disorder	28 (11%)	12 (20%)
Somatic Symptom and Related Disorder	25 (10%)	11 (18%)
Remote Concussion	26 (10%)	8 (13%)
Neurologic Disorder	112 (45%)	8 (13%)
Acquired Brain Injury (Excluding Concussion)	22 (9%)	7 (12%)
No Diagnosis	25 (10%)	6 (10%)
Posttraumatic Stress Disorder	16 (6%)	3 (5%)
Systemic Disease with Neurologic Sequalae	32 (13%)	2 (3%)
Primary Seizure Disorder	20 (8%)	2 (3%)
Neuroinfectious Disease	6 (2%)	1 (2%)
Substance Use Disorder	9 (4%)	1 (2%)
High-Risk of Invalid Performance Diagnoses	M = 0.54	M = 0.82
Performance Validity Tests		
TOMM Trial 1	M = 46.33	M = 33.00
MSVT Consistency	M = 95.00	M = 74.33
MSVT Immediate Recognition	M = 97.68	M = 82.25
MSVT Delayed Recognition	M = 95.58	M = 75.25
Dot Counting Test E-Score	M = 10.77	M = 20.30
Word Choice Test	M = 46.31	M = 45.72
Reliable Digit Span	M = 9.21	M = 5.75
RAVLT Forced Choice	M = 14.56	M = 11.62
Total Number of Failures	M = 0.61	M = 4.28

What do These Findings Suggest?

Proof of concept – unsupervised machine learning can synthesize various data to render similar conclusions as the standard validity approach

But it can do so without human intervention

Where Can We go From Here?

Leverage machine learning to reduce the cumbersome nature of validity assessment

May be able to do this by optimizing:
Efficiency
Objectivity
Personalization

….which in turn may increase the frequency and quality of validity assessment

Study Limitations

Needs larger sample size and replication

Include participants with greater cognitive impairment

Use other methods and validity tests to determine validity criterion ratings

Incorporate more variables relevant to validity assessment

Thank You!

Co-authors: Anthony Robinson, Jason Soble, & Violeta Rodriguez

The patients who made this possible

TCN/AACN for sponsoring this presentation