

Medical Education Journal Club

ACCREDITATION STATEMENT

Florida International University Herbert Wertheim College of Medicine is accredited by the Accreditation Council for Continuing Medical Education to provide continuing medical education for physicians.

CREDIT DESIGNATION STATEMENT

Florida International University Herbert Wertheim College of Medicine designates this live activity for a maximum of 1.0 *AMA PRA Category 1 Credit(s)*[™]. Physicians should claim only the credit commensurate with the extent of their participation in the activity.



Disclosure Information

Activity Directors / Planners / Reviewers / Faculty

Name	Role	Disclosure / Resolution
Carla S. Lupi, MD	Activity Director/Planner/ Speaker	Dr. Lupi reports no relevant financial relationships.
Vivian Obeso, MD	Planner/Speaker	Dr. Obeso reports no relevant financial relationships.
Christian Castro	Planner	Mr. Castro reports no relevant financial relationships.
Melissa Ward-Peterson, MPH	Planner	Ms. Ward-Peterson reports no relevant financial relationships.



FIU

**Herbert Wertheim
College of Medicine**

Learning Objectives

- Be able to list the search results for one medical education database consulted in the design of a course or teaching session.
- Conduct a critical appraisal of an article in the medical education research.
- Identify the applicability of research results to one's own course or teaching session.



FIU

**Herbert Wertheim
College of Medicine**

Medical Education Journal Club

- Establish a forum for faculty to share and discuss recent literature in medical education
- Use best evidence in medical education literature to evaluate and advance current practices in our educational program
- Establish a culture that promotes curricular innovation and change in an evidence-based manner
- Stimulate educational scholarship



FIU

**Herbert Wertheim
College of Medicine**

Objectives for Today's Session

- Describe the most recent modifications to the USMLE Step 2 CS examination
- List several measures of reliability
- Identify sources of error
- Describe Messick's framework as an approach in gathering evidence of validity
- Use the information gained from this study to evaluate and advance current practices in our educational program



FIU

**Herbert Wertheim
College of Medicine**

Validity Evidence For a Patient Note Scoring Rubric Based on the New Patient Note Format of the USMLE Step 2 CS

Yoon Soo Park, PhD, Matthew Lineberry, PhD, et al.
Academic Medicine, vol. 88, no.10, October 2013
Research in Medical Education (Rime), AAMC



FIU

**Herbert Wertheim
College of Medicine**

Problem Statement:

Lack of valid and reliable assessment tool/scoring rubric for the
USMLE Patient Note Format

USMLE Step 2 CS

- Recent Modifications
- Patient Note Format
- Assessment Drives Teaching and Learning
- USMLE has not disclosed scoring details



FIU

**Herbert Wertheim
College of Medicine**

About the USMLE Step 2 CS

3 Components

- Communication and Interpersonal Skills (CIS)
- Spoken English Proficiency (SIP)
- Integrated Clinical Encounter (ICE)
- P/F
- Students must pass all 3 components



FIU

**Herbert Wertheim
College of Medicine**

USMLE Step 2 CS

- 12 stations: 15 minute station, 10 minute note

Domain	Standardized Patient	Physician Rater
Communication and Interpersonal Skills	✓	
History	✓	
Physical Exam	✓	
Patient Note		✓

- Integrated Clinical Encounter = data gathering PE (SP) + PT note (physician rater)



FIU

Herbert Wertheim
College of Medicine

CLINICAL SKILLS EVALUATION

PATIENT NOTE

HISTORY: Describe the history you just obtained from this patient. Include only information (pertinent positives and negatives) relevant to this patient's problem(s).

PHYSICAL EXAMINATION: Describe any positive and negative findings relevant to this patient's problem(s). Be careful to include *only* those parts of examination you performed in *this* encounter.

DATA INTERPRETATION: Based on what you have learned from the history and physical examination, list up to 3 diagnoses that might explain this patient's complaint(s). List your diagnoses from most to least likely. For some cases, fewer than 3 diagnoses will be appropriate. Then, enter the positive or negative findings from the history and the physical examination (if present) that support each diagnosis. Lastly, list initial *diagnostic* studies (if any) you would order for each listed diagnosis (e.g. restricted physical exam maneuvers, laboratory tests, imaging, ECG, etc.).

DIAGNOSIS #1:

HISTORY FINDING(S)	PHYSICAL EXAM FINDING(S)

(+) Click to add row(s)

DIAGNOSIS #2:

HISTORY FINDING(S)	PHYSICAL EXAM FINDING(S)

(+) Click to add row(s)

DIAGNOSIS #3:

HISTORY FINDING(S)	PHYSICAL EXAM FINDING(S)

(+) Click to add row(s)

DIAGNOSTIC STUDIES

(+) Click to add row(s)



**Herbert Wertheim
College of Medicine**

Relevance

- Assessment drives Teaching and Learning
- Awareness of Expectations
- Writing patient notes is a core skill, 1st - 4th
- Faculty grading notes
- Reinforce 3rd year
- “Practice making permanence”
- Feedback
- Not only to pass the test - effort to improve process of clinical reasoning



FIU

**Herbert Wertheim
College of Medicine**

Reference to Literature

Process

High correlation - Patient note + data gathering - Clauser. Acad Med 2006

Performance

Videotaped - only 4% of notes matched with performance - "Do students do what they write and write what they do?" - Szauter. Acad Med 2006

Case Specificity

Mixed finding in case specificity - researchers have argued that students may have greater knowledge about some case presentations than others

Rater Reliability

Recent study - importance rater reliability/double scoring - Inconsistent performance on the part of raters makes a greater contribution to measurement error than case specificity - Clauser. Acad Med 2008



FIU

**Herbert Wertheim
College of Medicine**

Purpose

- Develop an assessment tool/scoring rubric that could assess students using the USMLE new patient note format
- Gather validity evidence for the note scoring rubric developed to assess three dimensions:
 1. Documentation
 2. DDx
 3. Workup



FIU

**Herbert Wertheim
College of Medicine**

Gaining Validity

Messick's Validity Framework:

- Content
- Internal structure
- Relationships to other variables
- Consequences
- Response process



FIU

**Herbert Wertheim
College of Medicine**

Design, Methods and Data Collection

- Retrospective, Quantitative Study
- Assessment tool developed - formed by expert committee
- 170 4th year at UIC COM - May 2012
- GCE - Graduating Competency Exam
- 5 SP encounters
- (SP checklist history/PE and Communications and Interpersonal Skills)
- Patient Note - 10 minutes
- Graded online by faculty
- Faculty trained to rubric
- One Faculty per case



FIU

**Herbert Wertheim
College of Medicine**

Content

Scoring rubric developed by faculty/committee

- Documentation, DDX, Workup
- 4 point Likert scale



FIU

**Herbert Wertheim
College of Medicine**

Table 1
Patient Note Scoring Rubric*

Dimension with description (maximum points)	Score [†]	Anchor
Documentation		
Documentation of findings in history and physical examination (30 points)	1	Key history and physical examination findings are missing or incorrect
	2	Most key positive findings present but poorly documented or disorganized or missing pertinent negatives
	3	Most key positive findings well documented and organized, may miss a few pertinent negatives
	4	All key information present, concise and well organized with little irrelevant information
DDX[‡]		
Justification of differential diagnosis (60 points)	1	Unreasonable differential diagnosis
	2	Appropriate differential diagnosis weakly supported, or several incorrect links between findings and diagnosis
	3	Appropriate differential diagnosis well supported, may have a few missing or incorrect attributions that would not impact diagnosis
	4	Excellent differential diagnosis well supported, links to diagnoses are correct and complete
Workup		
Plan for immediate diagnostic workup (10 points)	1	Diagnostic workup places patient in unnecessary risk or danger
	2	Ineffective plan for diagnostic workup, essential tests missed, irrelevant tests included
	3	Reasonable plan for diagnostic workup, may have some unnecessary tests
	4	Plan for diagnostic workup is effective and efficient, includes all essential tests, and few or no unnecessary tests



**Herbert Wertheim
College of Medicine**

Internal Structure

“Outcomes that yield reliable data is essential ... Reliable data are the foundation needed for educators to reach valid decisions, judgments about trainees.”

McGahie W, Issenberg B. A Critical Review of Simulation Based Medical Education Research: 2003-2009. Medical Education, 2010.

- **Various reliability measures exist**
- **Classical Theory** - consider only a **single source of error**
 - Test -Retest, (timing)
 - Parallel Forms (forms)
 - Internal Consistency (specific items)
 - Intra/Inter rater reliability
- **Generalizability Theory**



FIU

**Herbert Wertheim
College of Medicine**

G Theory

- **Generalizability Theory** - allows multiple sources of error in combination or by themselves
- to be estimated within a unified framework -
- more emphasis on the magnitude of the error from different sources



FIU

**Herbert Wertheim
College of Medicine**

Reliability

- **Generalizability study (G study)-**
 - Students (153) x cases (5) x dimensions rubric (4 fixed)
 - G and phi coefficients 0.47 and 0.43
- **Table 3**
 - Students 5.5%
 - Student x case - 19.5%
 - Case x dimension - 10.2%
- Estimates of a variance from a **GS** can be used to plan a **DS**
To help produce measurements that have the desired reliability
- **D study** - - 15 cases (.70%)



FIU

**Herbert Wertheim
College of Medicine**

Table 3

Variance Components of the Patient Note Scoring Rubric Generalizability Study (G Study)

Effect*	Degrees of freedom	Variance component (standard error)	% Variance component
<i>p</i>	152	0.035 (0.011)	5.5
<i>c</i>	4	0.013 (0.023)	2.1
<i>d</i>	2	0.001 (0.012)	0.1
<i>p</i> × <i>c</i>	608	0.125 (0.015)	19.5
<i>p</i> × <i>d</i>	304	0.025 (0.009)	3.9
<i>c</i> × <i>d</i>	8	0.065 (0.030)	10.2
<i>p</i> × <i>c</i> × <i>d</i>	1216	0.375 (0.015)	58.7

* *P* indicates persons (students); *c*, cases; *d*, dimensions of the rubric. The G study used the *p* (students) × *c* (cases) × *d* (dimensions of the rubric) design.



**Herbert Wertheim
College of Medicine**

Correlation: Pairwise Correlation Between Cases

- Total Patient Note Scores: 0.1 - 0.24
- Documentation and DDX = 0.44 ($p < .001$)
- DDX and Workup = 0.41 ($p < .001$)
- Documentation and Work-up = 0.33 ($p < .001$)



FIU

**Herbert Wertheim
College of Medicine**

Relationship to Other Variables

It should correlate strongly with other indicators of the same construct

- Overall documentation score and SP encounter checklist, 0.47 ($p < .001$)
- Total note score and SP checklist, 0.38 ($p < .001$)
- Total Note Scores and Comm, 0.2 ($p < 0.05$)



FIU

**Herbert Wertheim
College of Medicine**

Consequences

Explores evidence related to the intended or unintended consequences -

Effect of scoring

Impact on learning and teaching

- Compared pass/fail rates with previous GRS
- No meaningful difference

- New rubric: 1.3%
- Old rubric: 0%



FIU

**Herbert Wertheim
College of Medicine**

Response Process

- Search for data analyzing the relationship b/w the construct and the thought process and response action of examinees
- Rater survey
- Based on rater opinion
- 5- 7 min. score
- Thoroughness vs. concise
- Diff. dx. - clarity students instructions quest.- supporting and refuting findings in their justification
- Favored pertinent positives



FIU

**Herbert Wertheim
College of Medicine**

Authors' Conclusions

- First attempt to validate a scoring rubric based on USMLE new patient note format
- Gathered Validity evidence for 3 dimensions - Documentation, DDX, Workup
- Person - Case interaction = 20% total variance and low pairwise correlation b/w note scores
 - =Need for large number of cases
- Pairwise association between dimension scores suggest a link
 - =Good documentation = good ddx = good workup skills



FIU

**Herbert Wertheim
College of Medicine**

Authors' Conclusions

- Moderate correlation between SP scores and note scores
= are we measuring different skills ?
- Rater response underscored the need for additional rater training
- Rater survey addressed need for improved instruction and/or teaching documentation skills



FIU

**Herbert Wertheim
College of Medicine**

Limitations

- Single institution
- Only 5 cases
- Moderate reliability
- Unable to assess rater reliability facet
- Lack of detail regarding
- Case Development
- Rubric- lack of well defined anchors

"all key information, most, appropriate, ineffective, reasonable"

- Rater training - per case
- SP training?



FIU

**Herbert Wertheim
College of Medicine**

Questions?



FIU

**Herbert Wertheim
College of Medicine**

References

- Cook DA, Beckman TJ. Current concepts in validity and reliability for psychometric instruments; theory and application. *Am J Med* 2006;119:166e7-166.e16.
- McGahie W, Issenberg B. A Critical Review of Simulation Based Medical Education Research: 2003-2009. *Medical Education*, 2010



FIU

**Herbert Wertheim
College of Medicine**

Please complete the CME survey to receive credit for attendance.



FIU

**Herbert Wertheim
College of Medicine**