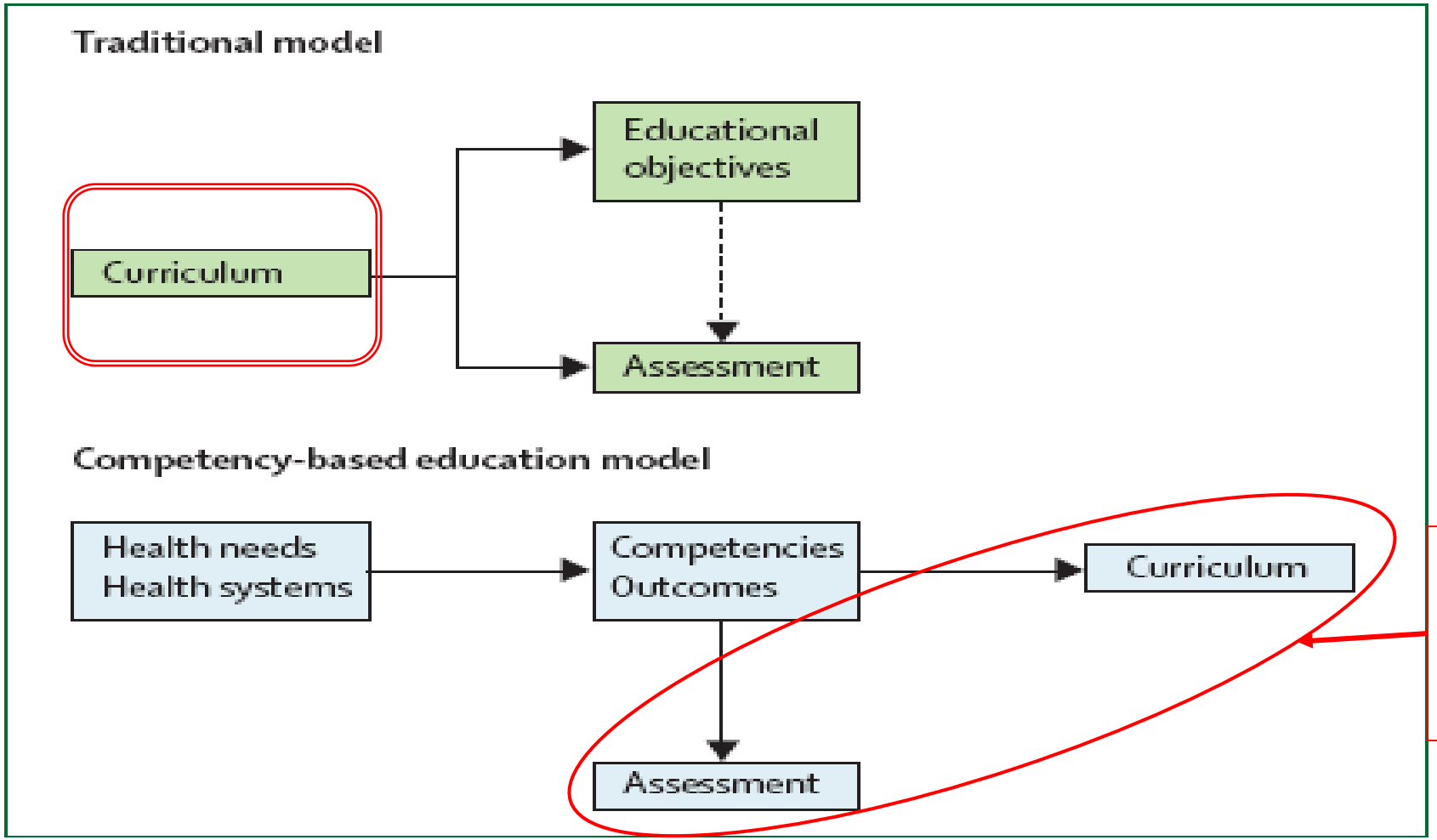
Work-based Assessment and Outcomes-based Systems

Eric Holmboe

OBME: Start with System Needs



Assessment and Curriculum must be integrated: Assessment drives learning; learning drives assessment

Frenk J, et al. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. Lancet. 2010

Early Principles: CBME

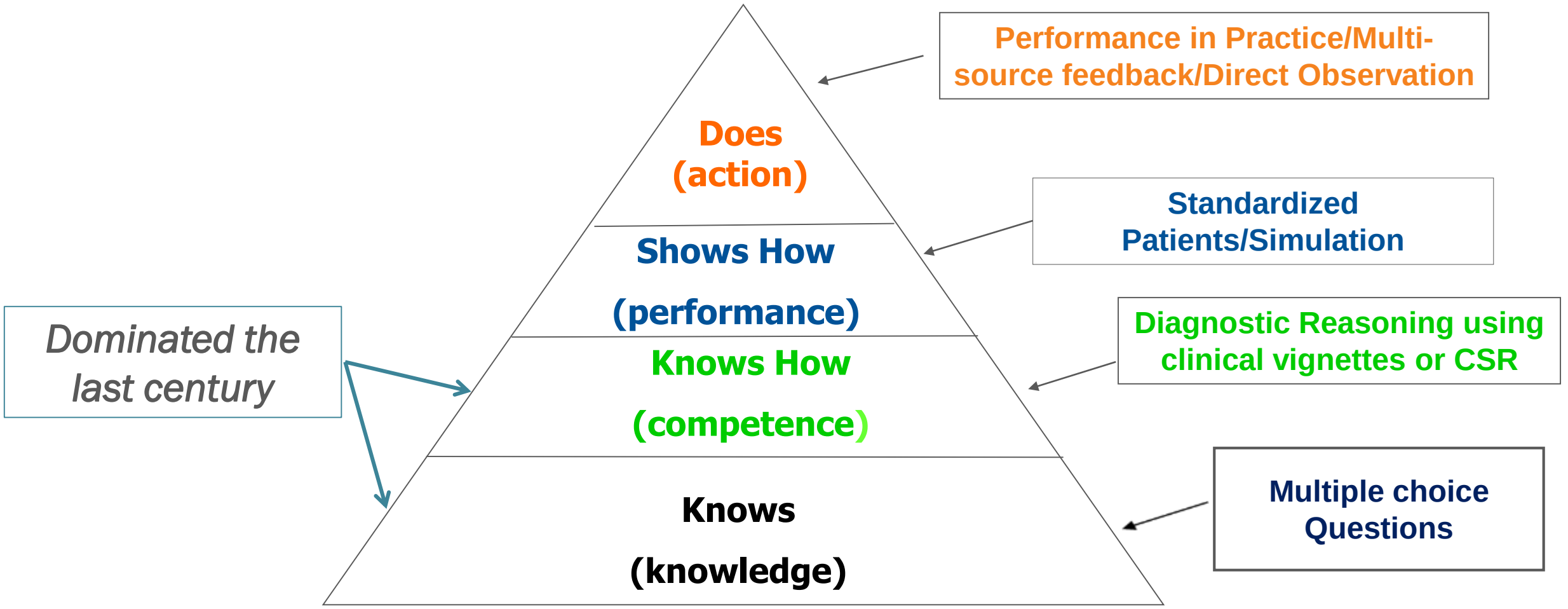
- World Health Organization (**1978**):
  - *"The intended output of a competency-based programme is a health professional who can practise medicine at a defined level of proficiency, in accord with local conditions, to meet local needs."*

McGaghie WC, ***Miller GE, Sajid AW, Telder TV. Competency-based Curriculum Development in Medical Education. World Health Organization, Switzerland, 1978.***

Work-based Assessment

- The essential element in a competency-based system
- What matters most to patients and healthcare systems
- Messy yes, but a poor excuse for not doing it and performing the research to do it better
  - Standardized testing was once a young science too
- What we've been doing for decades in assessment no longer sufficient if we are to meet societal needs

The Ubiquitous and Iconic Miller "Pyramid (1990)



AHRQ Quality Report 2013

Measure Focus	Measure Name/Description	Baseline Rate	Most Recent Rate	Aspirational Target
Aspirin Use	Outpatient visits at which adults with cardiovascular disease are prescribed/maintained on aspirin	47%[13]	53%[14]	Increase to 65% by 2017
Blood Pressure Control	Adults with hypertension who have adequately controlled blood pressure	46%[15]	53%[16]	Increase to 65% by 2017
Cholesterol Management	Adults with high cholesterol who have adequate control	33%[17]	32%[18]	Increase to 65% by 2017
Smoking Cessation	Outpatient visits at which current tobacco users received tobacco cessation counseling or cessation medications	23%[19]	22%[20]	Increase to 65% by 2017

Faculty OSCE Clinical Skills

Competency	Mean (SD)	Range	Generalizability
History Taking	65.5% (9.6%)	34% - 79%	0.80
Physical Exam	78.9% (13.6%)	36% - 100%	0.52
Counseling	77.1% (7.8%)	60% - 93%	0.33
Patient Satisfaction[1]	5.62 (0.48)	4.43 – 6.63	0.60

[1]On 7-point scale

N=44

Kogan JR. et al. Acad Med. 2010;85(10 Suppl):S25-8

The Problem of Uncertainty

From Lineberry et.al. critique of Script Concordance Testing:

The former [expert opinion] is a unique aspect of SCTs unaccounted for by classical test theory (CTT)…

We are unaware of any psychometric theory sophisticated enough to guide estimation of the adequacy or inadequacy of sampling from 'any population frequency distribution that may be observed across a 5-point scale'

Really, isn't uncertainty something MDs do everyday?!?

Clinical Questions Raised by Clinicians at the Point of Care

Systematic Review (11 articles):

- Mean frequency of questions = 0.57 per patient visit
  - GIM Physician: Roughly 90 visits/week
  - Assuming conservative estimate of 46 weeks of work/year
    - ***# of questions ≈ 2,228 per year or ≈22,000+ over 10 years***
- Physicians pursued 51% (range 36-66%) of questions
  - Found answers for 78% (range 67-88%)

No amount of testing is going to "fix this"

Del Fiol G, Workman TE, Gorman PN. Clinical Questions Raised by Clinicians at the Point of Care: A Systematic Review. JAMA Intern Med. 2014 Mar 24.

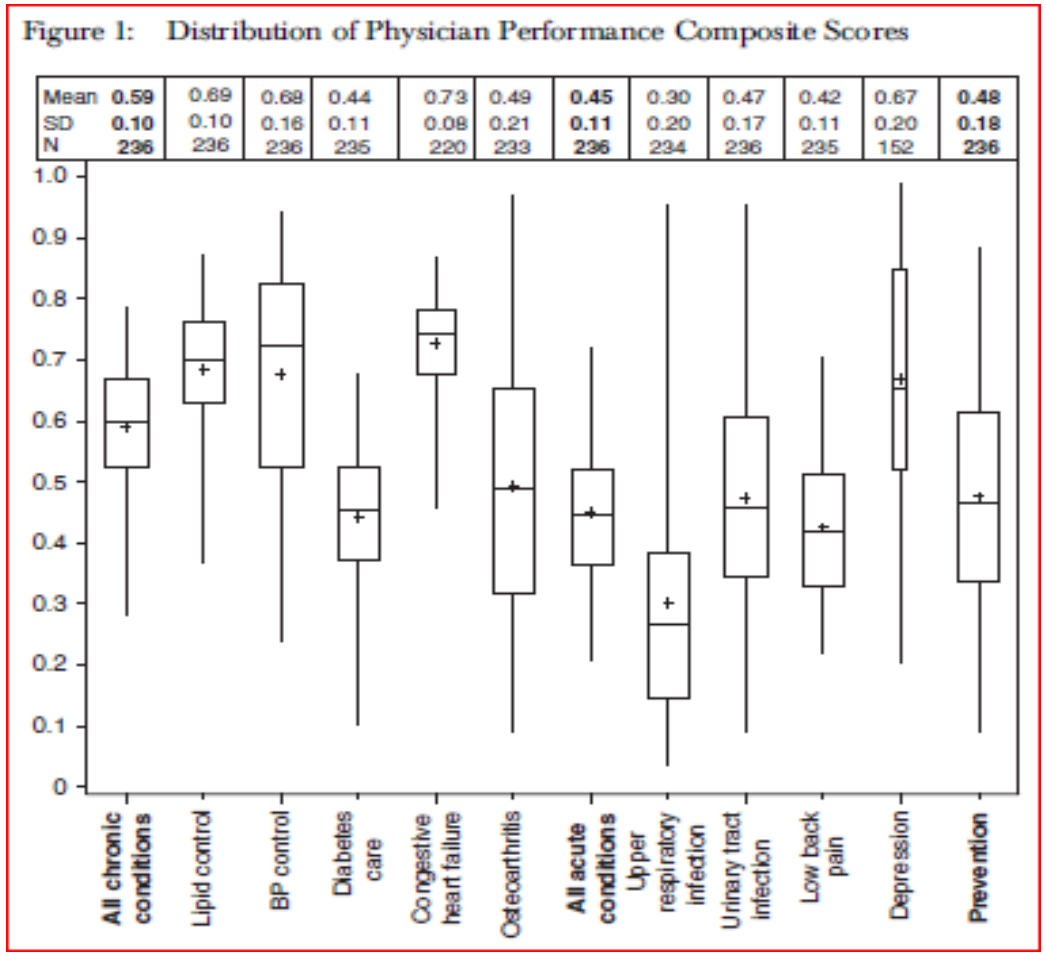
Comprehensive Care Project

Table 2: Mean Patient Characteristics Measured at the Physician Level (Chart Review of 22,526 Patients among 236 Physicians)

Characteristic	*Mean (SD)*	*Range across Physicians*
Age (years)	60 (6.0)	44–77
Patients with age ≥ 65	44% (0.16)	5–89%
Male gender	40% (0.12)	10–75%
Ethnicity		
Non-Hispanic white	37% (0.32)	0–99%
Non-Hispanic black	9% (0.18)	0–100%
Hispanic	8% (0.18)	0–100%
Unknown race	46% (0.37)	0–100%
Drug/alcohol abuse	2% (0.02)	0–13%
Psychiatric/cognitive impairment	15% (0.15)	0–60%
Problems with adherence to treatment	5% (0.08)	0–52%
Social factors limiting self-care	13% (0.18)	0–100%
Comorbidity index	1.0 (0.55)	0.02–3.5
Patients with comorbidity index >2	11% (0.09)	0–48%

- 236 GIM practices across US
- 46 total performance measures
- Patient demographics, characteristics and scope of practice varied widely

Results



- Wide range of performance across measures and composites (shown to left)
- At best, performance on the MOC exam ***only explained 4.9% of the variance*** of performance on the composite measures
- If "does" (performance) matters most to patients, then perhaps this is where we need more effort

Holmboe ES, Weng W, Arnold G, et al. The Comprehensive Care Project: Measuring Physician Performance in Ambulatory Practice. Health Services Research. 2010; 45(6 Pt 2):1912-33

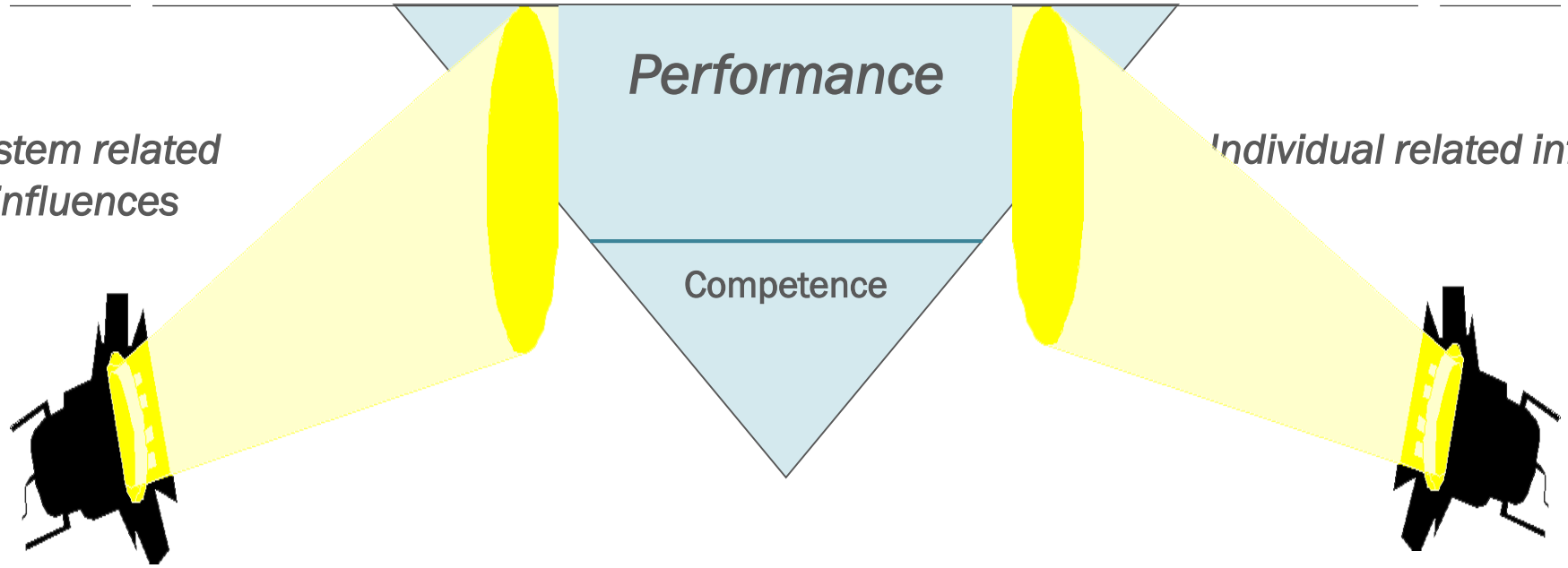
High Stakes Testing as Normal Science

Thomas Kuhn (1962):

"Normal science, the activity in which most scientists inevitably spend almost all of their time, is predicated on the assumption that the scientific community ***knows*** what the world is like. Much of the success of the enterprise derives from the community's willingness to defend that assumption, if necessary at considerable cost"

Thomas S. Kuhn. *The Structure of Scientific Revolutions*. University of Chicago Press. Chicago. 1962. Pg. 5.

Cambridge Model: “Righting” the Pyramid



Work-based assessment has to be a larger part of our future if we are to make meaningful gains in quality and safety

Rethans, Norcini, et al, 2002

Final Thoughts

"Seek out new knowledge and be willing to engage in dialogue about disconfirming data that challenges past beliefs. Learn and listen from others, most importantly from physicians and patients who sit at the core of the education and care experience. Be open to bringing the best ideas home for testing, trial and adaptation for improvement. All organizations and programs, including the regulatory entities, must embrace the obligation to investigate what they do, admit limitations, and improve for the benefit of the public."

From: HoImboe ES and BataIden PB. Achieving the Desired Transformation: Thoughts on Next Steps for Outcomes-based MedicaI Education. Acad Med. 2015; in press.