Milestones then Tombstones

The Death of the Likert Scale

Wendy Warren, MD (Associate Program Director)
Holly Montjoy, MD (Faculty)
Brittany Thoma (Residency Program Coordinator)
4/1/2015

AMERICAN ACADEMY OF FAMILY PHYSICIANS

AAFP Disclosure

It is the policy of the AAFP that all individuals in a position to control content disclose any relationships with commercial interests upon nomination/invitation of participation. Disclosure documents are reviewed for potential conflicts of interest and, if identified, conflicts are resolved prior to confirmation of participation. Only those participants who had no conflict of interest or who agreed to an identified resolution process prior to their participation were involved in this CME activity.

All individuals in a position to control content for this session have indicated they have no relevant financial relationships to disclose.

Proposed Key Practice Changes

Make the Milestones process easier and more meaningful for all by:

1) Changing the language of residency evaluations to reflect the Milestones language
2) Changing the format of evaluations to match sub-competencies to appropriate core rotations
3) Changing the scale of evaluations to reduce assessment variability

Proposed Key Practice Changes

Ultimately the goal is to make the ACGME Milestones evaluations:

- More accurately represent the residents' performance
- A seamless, less timely process for faculty

Poll Question

Per the Milestones Review, a _____is designed to be the resident graduation target:

A. Level 2
B. Level 3
C. Level 4
D. Level 5

Milestones Levels

- Level 1: The resident demonstrates milestones expected of a resident who has had some education in family medicine.
- Level 2: The resident is advancing and demonstrating additional milestones.
- Level 3: The resident continues to advance and demonstrate additional milestones; the resident consistently demonstrates the majority of milestones targeted for residency.
- Level 4: The resident has advanced so that he or she now substantially demonstrates the milestones targeted for residency. This level is designed as the graduation target.
- Level 5: The resident has advanced beyond performance targets set for residency and is demonstrating “aspirational” goals which might describe the performance of someone who has been in practice for several years. It is expected that only a few exceptional residents will reach this level.

Key Points - Milestones Working Group

- Level 4 is indicated as the target for graduation… but the determination of an individual's readiness for graduation is at the discretion of the program director
- Target vs Requirement

Our Predicament

- Core rotation evaluations contain primarily numeric data via a Likert-Type scale
  - Difficult to extrapolate information
  - Very little supporting data provided
  - Low compliance rate
  - Residents want high quality feedback and MORE of it!

Things to Consider

- Variance in faculty ratings of an individuals performance
- Learner centric assessment
  - Promotes higher performance,
  - Integrates education and experience
  - Provides constructive feedback that motivates the learner
  - Likert scale difficult to individualize
- Moving away from the RIME (reporter, interpreter, manager, educator) model
  - Transitioning to a continuum

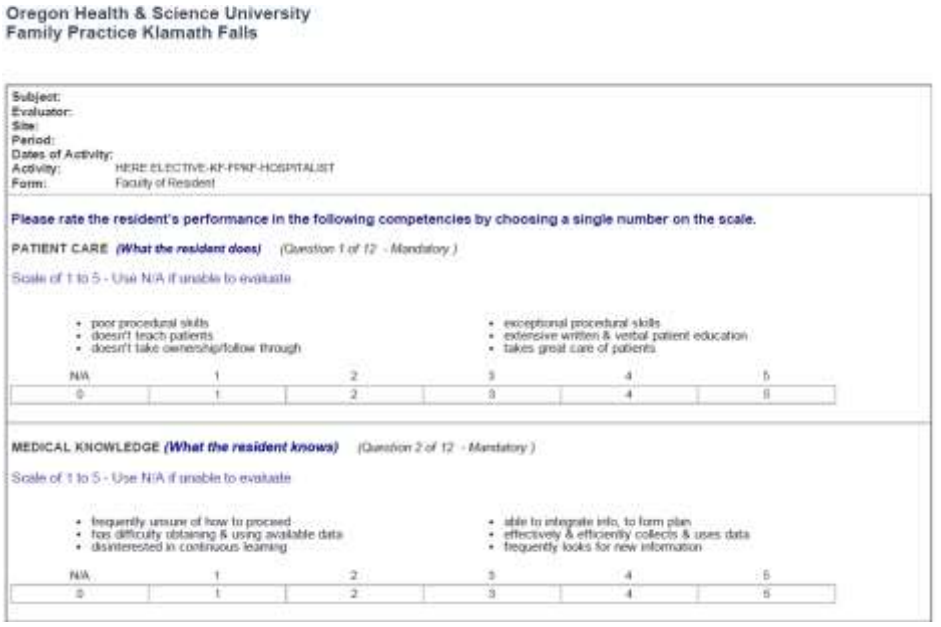
Things to Examine

- The purpose and function of assessment processes
  - Will the assessment result in useful and defendable performance appraisals?
- The amount and quality of evidence
- Feedback loops within our programs
- Summative evaluations (end of rotation) vs Focused assessments (informal and frequent)
  - More emphasis needed on the latter

Comments versus Numeric Scores

- Analyzing written comments can improve the sensitivity of the evaluations
  - Especially in identifying learners who may benefit from remediation than would been identified by scores alone
- Comments are more likely to reveal issues in professionalism

OUR ORIGINAL EXAMPLE:



Oregon Health & Science University
Family Practice Klamath Falls

Subject:
Evaluator:
Site:
Period:
Dates of Activity:
Activity: HERE ELECTIVE-KF-FPKF-HOSPITALIST
Form: Faculty of Resident

Please rate the resident's performance in the following competencies by choosing a single number on the scale.

PATIENT CARE *(What the resident does)* *(Question 1 of 12 - Mandatory)*

Scale of 1 to 5 - Use N/A if unable to evaluate

- poor procedural skills
- doesn't teach patients
- doesn't take ownership/follow through

- exceptional procedural skills
- extensive written & verbal patient education
- takes great care of patients

N/A	1	2	3	4	5
0	1	2	3	4	5

MEDICAL KNOWLEDGE *(What the resident knows)* *(Question 2 of 12 - Mandatory)*

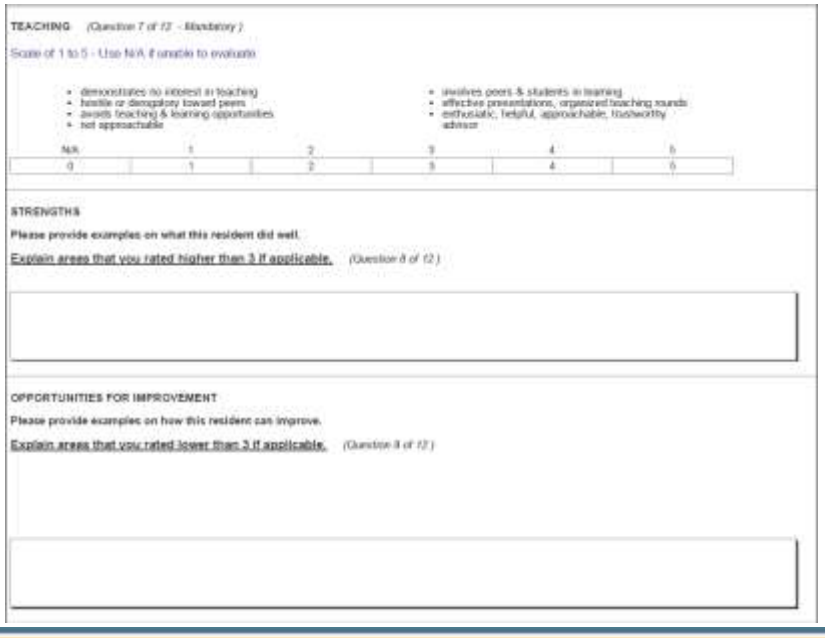
Scale of 1 to 5 - Use N/A if unable to evaluate

- frequently unsure of how to proceed
- has difficulty obtaining & using available data
- disinterested in continuous learning

- able to integrate info, to form plan
- effectively & efficiently collects & uses data
- frequently looks for new information

N/A	1	2	3	4	5
0	1	2	3	4	5

OUR ORIGINAL EXAMPLE:



TEACHING (Question 7 of 12 - Mandatory)

Scale of 1 to 5 - Use N/A if unable to evaluate

- demonstrates no interest in teaching
- hostile or derogatory toward peers
- avoids teaching & learning opportunities
- not approachable
- involves peers & students in learning
- effective presentations, organized teaching rounds
- enthusiastic, helpful, approachable, trustworthy advisor

N/A	1	2	3	4	5
0	1	2	3	4	5

STRENGTHS

Please provide examples on what this resident did well.

Explain areas that you rated higher than 3 if applicable. (Question 8 of 12)

OPPORTUNITIES FOR IMPROVEMENT

Please provide examples on how this resident can improve.

Explain areas that you rated lower than 3 if applicable. (Question 9 of 12)

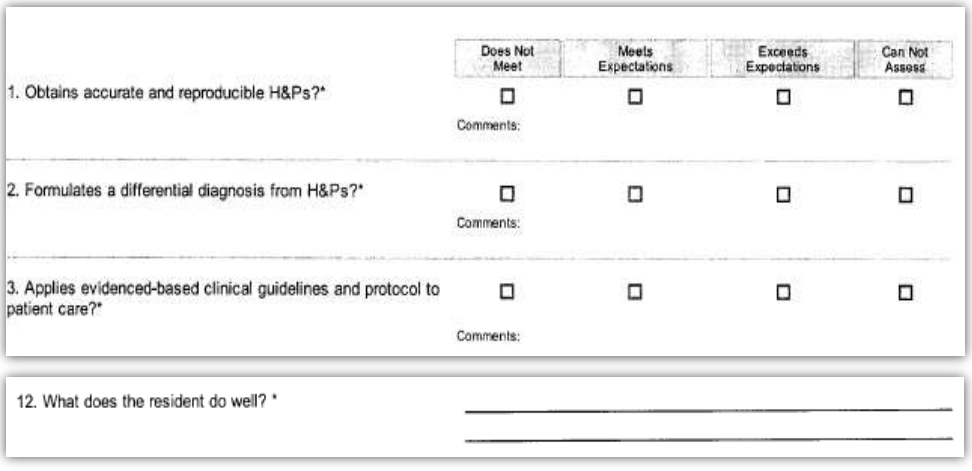
Our Process

- Program coordinator and faculty member met bi-monthly for six months
- Identified which rotations align with sub-competencies
  - Example: PC-1 relevant to: Inpatient,OB, ER, Pediatrics
- Evaluated curriculum for gaps in content
  - Are we effectively teaching practice-based learning? (PBLI)
- Began building PGY1 evaluation questions

Our Process

- Decided to use Meets/Exceeds/Does Not Meet scale (based on faculty feedback)
- Tailored adjunct faculty evaluations to have more of a professionalism focus
- Made comments REQUIRED
- Conveyed vision to MedHub to map answers
- Began building PGY2 and PGY3 evaluations

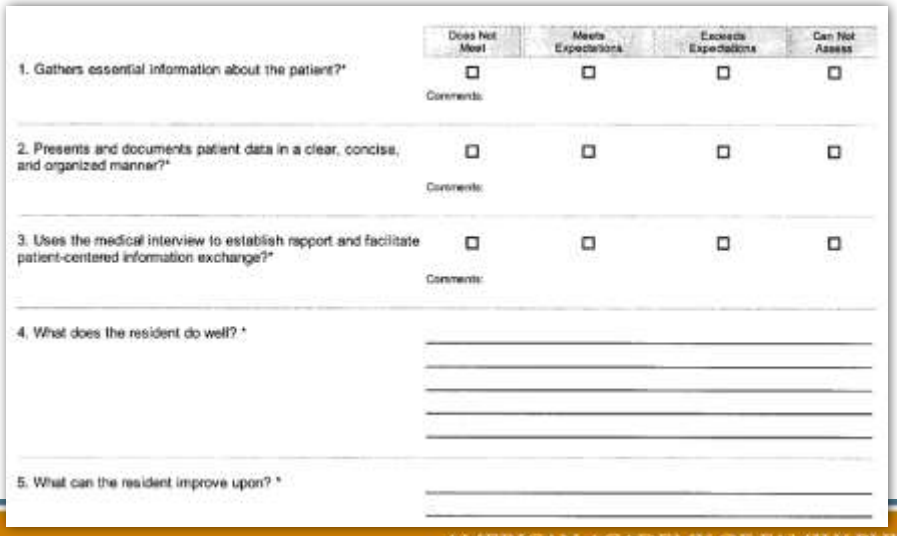
Our New Evaluations- Inpatient



	Does Not Meet	Meets Expectations	Exceeds Expectations	Can Not Assess
1. Obtains accurate and reproducible H&Ps?*	☐	☐	☐	☐
Comments:				
2. Formulates a differential diagnosis from H&Ps?*	☐	☐	☐	☐
Comments:				
3. Applies evidenced-based clinical guidelines and protocol to patient care?*	☐	☐	☐	☐
Comments:				

12. What does the resident do well? *

Our New Evaluations- Outpatient



	Does Not Meet	Meets Expectations	Exceeds Expectations	Can Not Assess
1. Gathers essential information about the patient?*	☐	☐	☐	☐
Comments:				
2. Presents and documents patient data in a clear, concise, and organized manner?*	☐	☐	☐	☐
Comments:				
3. Uses the medical interview to establish rapport and facilitate patient-centered information exchange?*	☐	☐	☐	☐
Comments:				
4. What does the resident do well? *				
5. What can the resident improve upon? *				

Our Process

- December 2015… SHOWTIME!
  - Provided advisors with printed Med Hub evaluations and Milestones maps to aid advisor in completing Milestones packet
  - Met as a group of (devoted, collegial) faculty to discuss Milestones data on each resident
- What map do we speak of?

Milestones Mapping via MedHub

Poll Question

The ACGME Milestones for Family Medicine consist of _____ core competencies and _____ sub-competencies.

A. 8 and 24
B. 8 and 21
C. 7 and 30
D. 6 and 22

Our Review Process

- Faculty asked to complete surveys immediately following the group Milestones review
  - Response rate: 100%+ (core faculty +others comments)
- Residents (PGY-2/3) asked to complete surveys immediately after advisor meeting dedicated to Milestones
  - Response rate: 53%

Our Review Process

- Lower response rate with residents:
  - Difficult to schedule advisor meetings while on inpatient medicine and maternal/child health
  - More scattered process (meetings occurring at various times, different days, etc)

Faculty Survey

Milestones then Tombstones: The Death of the Likert Scale- Faculty

1. Quality feedback is typically described as being timely, specific, and relevant. Do you think changing the core rotation evaluations to reflect Milestones language has changed the quality of the feedback you provide about resident performance?

   Yes **No** **Neutral**

2. Did changing the evaluations to reflect milestones language help increase your efficiency when filling out the semi-annual milestones for your advisees?

   Yes **No** **Neutral**

3. If yes, please estimate much time was saved.

   a.

4. Does having the core resident rotation evaluations in Milestones language help eliminate the translation dilemma between evaluator tools? (ie. Do you feel it is easier to extrapolate a score from the rotation evaluations to the semi-annual ACMGE milestones tool now versus when previously using Likert Data?) Note: previous Likert rotation template provided.

   Yes **No** **Neutral**

5. Please offer any additional comments on the two forms and their effectiveness in your own experience.

Survey Results- Faculty

	Yes	No	Neutral
Changes in Quality	82%	9%	9%
Increased Efficiency	40%	40%	20%
Decreased Translation Dilemma	67%	0%	33%

Faculty Comments- Positive

"Gives very specific suggestions and things to discuss with the resident about their performance and areas for improvement"

"This system is far more objective and concentrated- allowing for evidence to flow easier across evaluations"

"Milestones are rooted more in observables rather than attitudes…It is far more accurate to get committee consensus than the old likert-scaling"

Faculty Comments- Critical

"We haven't been using the Milestones long enough to make the onerous language quicker to navigate"

"Without comments, can be non-specific"

Faculty Perspective- Mixed

- "I think the new evaluations can make things more complicated because it's difficult to assess each characteristic in milestone. However, I so think the evals are more accurate, timely, and easier to complete."
- I found that we have trouble describing the right level for residents who are markedly higher or lower than their PGY level. With the anchoring language being from the milestones we are not likely to be able to mark a yes/exceeds and show that a first year meets a level 4 on milestones. We change from a pass/fail mindset to a developmental process and I like that.

Survey Results- Faculty

	Yes	No	Neutral
Changes in Quality	82%	9%	9%
Increased Efficiency	40%	40%	20%
Decreased Translation Dilemma	67%	0%	33%

Resident Survey

Milestones then Tombstones: The Death of the Likert Scale- Residents

1. Quality feedback is typically described as being timely, specific, and relevant. Do you think changing the core rotation evaluations to reflect Milestones language has changed the quality of the feedback you receive about your performance?

   Yes **No** **Neutral**

2. If yes, please designate if the changes have made your feedback:

   Better or **Worse**

3. Please identify two specific differences have you noticed regarding the ACGME milestone evaluation process now versus when previously using the Likert Scale? (Previous Likert evaluation template provided.)

   a.

   b.

4. Please complete. The current process of evaluation, using Milestones language in core rotation evaluations, _____ accurately reflects my performance compared to the previous Likert Scale Evaluations.

   More **Less** **If no difference noted please circle: Neutral**

5. Please offer any additional comments on the two forms and their effectiveness in your own experience.

Survey Results- Residents

	Yes	No	Neutral
Changes in Quality	11%	44%	44%
Increased Accuracy	44%	***	44%

*** One person did not answer

Resident Comments- Positive

"Changes with advancement in training"

"More comprehensive in terms of content/subject matter"

"More specific goals – less room for personality-type feedback which improves objectivity across evaluations"

"The new scale is easier to interpret (meets expectations) had a pretty clear meaning whereas 3.5 or 4 on the likert scale is open to interpretation"

Resident Comments- Critical

"I haven't noticed differences"

Survey Results- Residents

	Yes	No	Neutral
Changes in Quality	11%	44%	44%
Increased Accuracy	44%	***	44%

*** One person did not answer

Poll Question

In how many sub-competencies do you feel you've reached a Level 5 (consistently "demonstrating aspirational goals")?

A. 0
B. 1-8
C. 9-15
D. 16-22

In Summary…

What Worked- PEARLS

- Prompted our program to re-visit both inpatient and outpatient evaluations
  - Developed three-question evaluations for clinic with preceptors evaluating one resident in each half-day of precepting
- Relying on core faculty to provide more feedback
- Able to schedule time for program coordinator and faculty to work on the project

What Didn't Work- PITFALLS

- Faculty Milestones Fatigue (FMF Syndrome)
  - Disgruntled faculty completing surveys
- Community preceptors still don't evaluate residents regularly (perhaps even less feedback because of changing to a different system)

What we learned!

- Faculty like the new evaluation system and most felt it:
  - Increased the quality of resident evaluation data
  - Decreased the burden of translating data between varying scales
- Residents were mixed:
  - They did not feel it changed the quality of their feedback
  - Half of respondents felt the new system more accurately represented their performance on core rotations
- It takes a lot of work…and you can't please everyone ☺

Remember: Milestone Working Group

"The Milestone tables were **NOT** designed to be used as evaluation forms for specific rotations or experiences…Utilizing language from the Milestones may be helpful as part of a mapping exercise to determine what competencies are best covered in specific rotation and curricular experiences…It is **imperative** that programs remember that the Milestones are not inclusive of the broader curriculum."

The Road Ahead…

During the break…

- Discuss / think about how you might implement the information you just heard.
- Fill out a session evaluation.

References

1. Colbert CY, Ogden PE, Ownby AR, Bowe C. Systems-based practice in graduate medical education: systems thinking as a foundational construct. Teach Learn Med. 2011;23(2):179–185.

2. Dannefer EF. Beyond assessment of learning toward assessment for learning: educating tomorrow's physicians. Med Teach. 2013;35(7):560–563.

3. Colbert CY, Dannefer EF, French J. Clinical Competency Committees and Assessment: Changing the Conversation in Graduate Medical Education. J Grad Med Ed. 2015 June:162-165

4. Boateng B, Bass L, Blaszak R, Farrar H. he Development of a Competency-Based Assessment Rubric to Measure Resident Milestones. J Grad Med Ed. 2009 Sept(11):45-48

5. Paulito A.R., Donelly M.B., Plymale M., Mentzer R.M. What do faculty observe of medical students' clinical performance? Teach Learn Med. 2006;18(2):99-104

6. Pangaro L. A new vocabulary and other innovations for improving descriptive in-training evaluations. Acad Med. 1999;749110:1203-1207

References

7. Hauer KE, Chesluk B, Iobst W, Holmboe E, Baron RB, Boscardin CK, et al. Reviewing residents' competence: a qualitative study of the role of Clinical Competency Committee in performance assessment. *Acad Med*. 2015;90(8):1084–1092.

8. Kogan JR, Holmboe ES. Preparing residents for practice in new system of care by preparing their teachers. *Acad Med*. 2014;89(11):1436–1437.

9. Moonen–van Loon, Joyce M.W. PhD; Overeem, Karlijn MD, PhD; Govaerts, Marjan J.B. MD, PhD; Verhoeven, Bas H. MD, PhD; van der Vleuten, Cees P.M. PhD; Driessen, Erik W. PhD. The Reliability of Multisource Feedback in Competency-Based Assessment Programs: The Effects of Multiple Occasions and Assessor Groups. Acad Med. 2015;90(8):1093-1099

10. Ginsburg S, Gold W, Cavalcanti R, Kurabi B, McDonald-Blumer H. Competencies "Plus": The Nature of Written Comments on Internal Medicine Residents' Evaluation Forms. Acad Med. 2011;86:S30–S34

11. Ginsburg S, Eva K, Regehr G. Do In-Training Evaluation Reports Deserve Their Bad Reputations? A Study of the Reliability and Predictive Ability of ITER Scores and Narrative Comments. Acad Med. 2013;88:1539–1544

AMERICAN ACADEMY OF FAMILY PHYSICIANS

STRONG MEDICINE FOR AMERICA