

Evaluation of Educational Innovations for Students and Residents: Principles to Enhance Medical Education Research

Judy A. Shea, Ph.D.

Associate Dean Medical Education Research

Professor of Medicine

Division of General Internal Medicine

Penn Medicine

Overview

Mission of medical education – produce clinically competent physicians

Many players:

- » Learners - Administrators
- » Educators - Evaluators

Value similar endpoints

- » Have different foci, priorities, agendas

Objectives

- ❖ To present a framework for thinking about medical education research
- ❖ To review four evaluation/research principles to help enhance the quality of medical education research: study design, data collection methods, data quantity, data interpretation

Medical Education Research Framework

| Unit of Analyses | Focus of Research | |
|-------------------------------------|-------------------|---------|
| | Process | Outcome |
| Person (student, resident, faculty) | X | X |
| Program (course, residency) | X | X |

General research principles

Ask a good question

- » Literature/theory based

Design a good study

Select optimal data collection methods

Gather enough data

- » Psychometrics/reproducibility/power

Interpret the data appropriately

General research principles

Ask a good question

- » Literature/theory based

Design a good study

Select optimal data collection methods

Gather enough data

- » Psychometrics/statistics/reproducibility

Interpret the data appropriately

Principle #1:
Design a Good Study
“Planning is Better Than Not Planning”

What does this imply?

have a [good] research question

have thought about how to answer it

have measurable processes/outcomes

Keywords: a priori/planning/prospective

Designs

Correlational/cohort studies

observational

descriptive

Experimental studies

intervention

post-intervention observation

Non-experimental

Single group - post test only

G1: -----|-----○

Single group – pre-post test

G1: ○-----|-----○

Extremely common in medical education

Multiple threats to validity

Quasi experimental

Have a nonrandom control group

G1: _____ UC _____ O

G2: _____ I _____ O

Probably underutilized in medical education

Experimental

Subjects are randomly assigned to treatment groups

Some level(s) of blinding

RG1: _____ UC _____ O

RG2: _____ I _____ O

Example for Principle #1

What if we taught students to give oral presentations and had them practice?
Would they be better with practice than without?

Kim S, Cohen JR, Shea JA. A randomized controlled study of encounter cards to improve oral case presentation skills of medical students. Journal of General Internal Medicine, 2005;20:743-747.


Medical Education Research Framework

| Unit of Analyses | Focus of Research | |
|------------------|-------------------|---------|
| | Process | Outcome |
| Person | X | X |
| Program | X | X |

RATIONALE: Communication skills are an essential component of professional competence. Little research has focused on medical student's oral case presentation (OCP) skills.

OBJECTIVE: To develop OCP encounter cards to enable ratings of students' OCP skills and examine the impact of an intervention on students' presentation skills.

DESIGN: Medicine clerkship has 4 12-week blocks; each has a 6-week inpatient session; randomized based on whether begin with inpatient session within block

Block 1: 

Block 2: 

Block 3: 

Block 4: 

All (approximately 20) students within session get same treatment

TREATMENT: Everyone gets curriculum.

Treatment Group: complete 9 OCP cards:

3 from inpatient attendings

3 from inpatient residents

3 from outpatient attendings

OCP Cards:

9-point scale

rate 7 content dimensions (e.g., Hx of present illness, assessment, plan), organization, overall speaking ability

OUTCOME: Make oral presentation to faculty

- » Randomized to type of case
- » Examiner blinded to treatment status

ANALYSES:

Completion rate, mean OCP card ratings, completion times,
and satisfaction ratings

Differences in ratings for different types of evaluators

Reliability, homogeneity

Treatment effect: did the intervention work

KEY FINDINGS:

Performance on OCP cards correlated with:
inpatient clinical evaluations ($r=.58$)
ratings of presentation skills ($r=.43$)
final grades ($r=.40$)

BUT

Final summative OCP performance was worse
for intervention than control group

Principle #1:
Design a Good Study
“Planning is Better Than Not Planning”

TAKE HOME MESSAGE:

With some planning it is possible to implement an experimental design.

Principle #2: Select Optimal Data Collection Methods “Do the Right Thing”

What does this mean?

There are lots of ways to collect data

Quantitative/qualitative debate

How am I going to answer my question?

ask people questions

ask people about other people

watch people

Example for Principle #2:

What if I do a study on mentoring of residents?
Everyone says it is important but what do we
really know about it?

Castiglioni A, Bellini LM, Shea JA. Program directors' views of the importance and prevalence of mentoring in internal medicine residency. Journal of General Internal Medicine, 2004;19:779-782.

Medical Education Research Framework

| Unit of Analyses | Focus of Research | |
|------------------|-------------------|---------|
| | Process | Outcome |
| Person | X | X |
| Program | X | X |

RATIONALE: Mentoring is important. Most of what we know about mentoring is for faculty. The benefits might be expected to extend to residents in need of career and personal counseling.

RATIONALE: Mentoring is important. Most of what we know about mentoring is for faculty. The benefits might be expected to extend to residents in need of career and personal counseling.

OBJECTIVE: Assess program directors' attitudes about mentoring, estimate prevalence of formal mentoring programs, describe characteristics of formal mentoring programs.

METHODS: Send mail survey to internal medicine program directors. Page 1 assesses program demographics and attitudes. Page 2 asks about structural features of formal mentoring programs.

KEY FINDINGS: Attitudes regarding mentoring were very favorable. Half had a formal program. Programs are largely unstructured, loosely monitored, and under evaluated.

| Attitude | Agree |
|---|-------|
| Mentorship is an important tool for career/professional development | 91% |
| Program Directors have a responsibility to encourage faculty to mentor residents | 92% |
| It is important for a resident to have a mentor during training. | 79% |
| Program Directors have a responsibility to encourage residents to identify a mentor | 73% |
| Program Directors have a responsibility to identify mentors for residents residents | 66% |
| Residency Programs should have structured mentoring programs for residents. | 61% |

| Features of a Mentoring Program: | Ideal % | Real % |
|----------------------------------|---------|--------|
| Individual mentoring | 89 | 96 |
| Group mentoring | 23 | 18 |
| Peer mentoring | 35 | 19 |
| Regularly scheduled meetings | 57 | 31 |
| Evaluation by residents | 36 | 52 |
| Evaluation by mentors | 30 | 28 |
| A structured curriculum | 14 | 11 |

Principle #2: Select Optimal Data Collection Methods “Do the Right Thing”

TAKE HOME MESSAGE:

Process assessment is useful for learning about features of programs - what they do/offer (caution: what people say they do may be different from what they actually do).
Triangulation is helpful.

Principle #3: Gather Enough Data

“More Is Better Than Less”

What does this mean?

- Need an adequate number of observations

- Case specificity

- Sampling/bias

- Capture more signal than noise

How do we evaluate?

| | Knowledge | Skills |
|---------------------|-----------|--------|
| MCQ test | ++ | - |
| SP based test | -- | ++ |
| Teachers/attendings | + | + |
| Chart reviews | ? | ? |

Where is the sampling error?

| | |
|---------------------|--------------------------------------|
| MCQ test | Wrong content, not enough items |
| SP based test | SP training/specificity |
| Teachers/attendings | Different ideas/ “hawks” and “doves” |
| Chart reviews | A lot does not get recorded |

Need a lot of observations

Example for Principle 3:

Kogan JR, Bellini LM, Shea JA. Feasibility, reliability and validity of the mini-clinical evaluation exercise (mCEX) in a medicine core clerkship. Academic Medicine, 2003;78:s33-35.

Example for Principle 3:

Everyone knows students are not being watched and given feedback. What if we designed a process to help students get feedback about their history and physical examination skills?

Kogan JR, Bellini LM, Shea JA. Feasibility, reliability and validity of the mini-clinical evaluation exercise (mCEX) in a medicine core clerkship. Academic Medicine, 2003;78:s33-35.

Medical Education Research Framework

| | Focus of Research | |
|------------------|-------------------|---------|
| | Process | Outcome |
| Unit of Analyses | | |
| Person | X | X |
| Program | X | X |

RATIONALE: Medical students must acquire core clinical skills. Many students complete training without being observed performing a history or physical exam. A feasible tool that promotes observation is needed.

OBJECTIVE: To determine the feasibility, reliability and validity of the mCEX when used in a medicine core clerkship.

1. Medical Interviewing Skills (○ Not observed)
1 2 3 4 5 6 7 8 9
UNSATISFACTORY SATISFACTORY SUPERIOR

2. Physical Examination Skills (○ Not observed)
1 2 3 4 5 6 7 8 9
UNSATISFACTORY SATISFACTORY SUPERIOR

3. Humanistic Qualities/Professionalism (○ Not observed)
1 2 3 4 5 6 7 8 9
UNSATISFACTORY SATISFACTORY SUPERIOR

4. Clinical Judgment (○ Not observed)
1 2 3 4 5 6 7 8 9
UNSATISFACTORY SATISFACTORY SUPERIOR

5. Counseling Skills (○ Not observed)
1 2 3 4 5 6 7 8 9
UNSATISFACTORY SATISFACTORY SUPERIOR

6. Organization/Efficiency (○ Not observed)
1 2 3 4 5 6 7 8 9
UNSATISFACTORY SATISFACTORY SUPERIOR

7. Overall Clinical Competence (○ Not observed)
1 2 3 4 5 6 7 8 9
UNSATISFACTORY SATISFACTORY SUPERIOR

Mini-CEX Time: Observing: ____ min Feedback: ____ min

Evaluator Satisfaction with Mini-CEX

LOW 1 2 3 4 5 6 7 8 9 HIGH

Student Satisfaction with Mini-CEX

LOW 1 2 3 4 5 6 7 8 9 HIGH

METHODS:

Participants: 121 medicine clerkship students (1/02-9/02)

Requirement: complete 9 mCEX during clerkship

mCEX booklets: ratings in 7 domains on 9-point scale

- » Record time and satisfaction with encounter

ANALYSES:

- » Feasibility: number of forms, time and satisfaction
- » Reproducibility: student x item x rater D-study
- » Validity: Pearson correlations and ANOVAs

KEY FINDINGS:

Feasibility

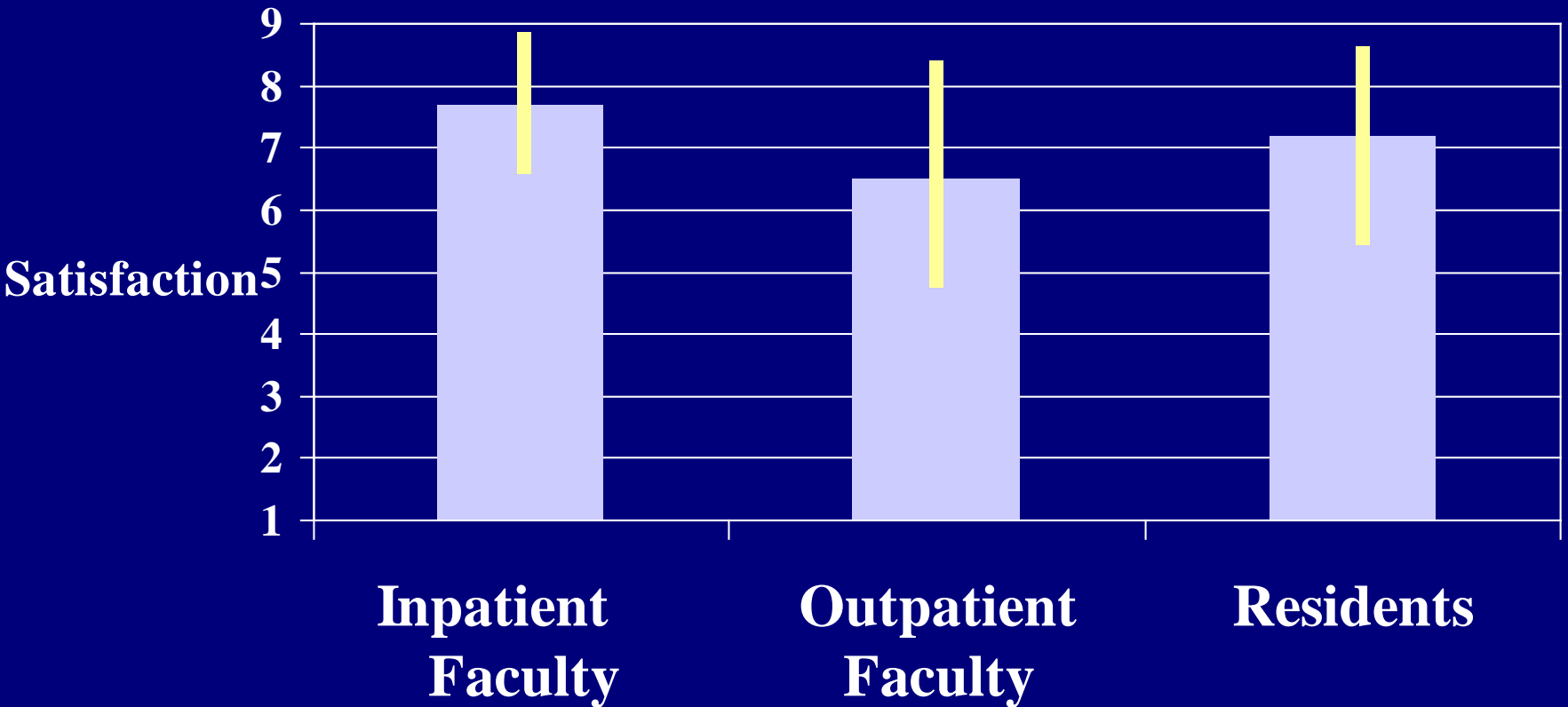
Completion rates: 89% of all forms (n=1,297)

Mean # forms/student: 7.9 (range 2-10, median 8)

Median observation time: 15 minutes

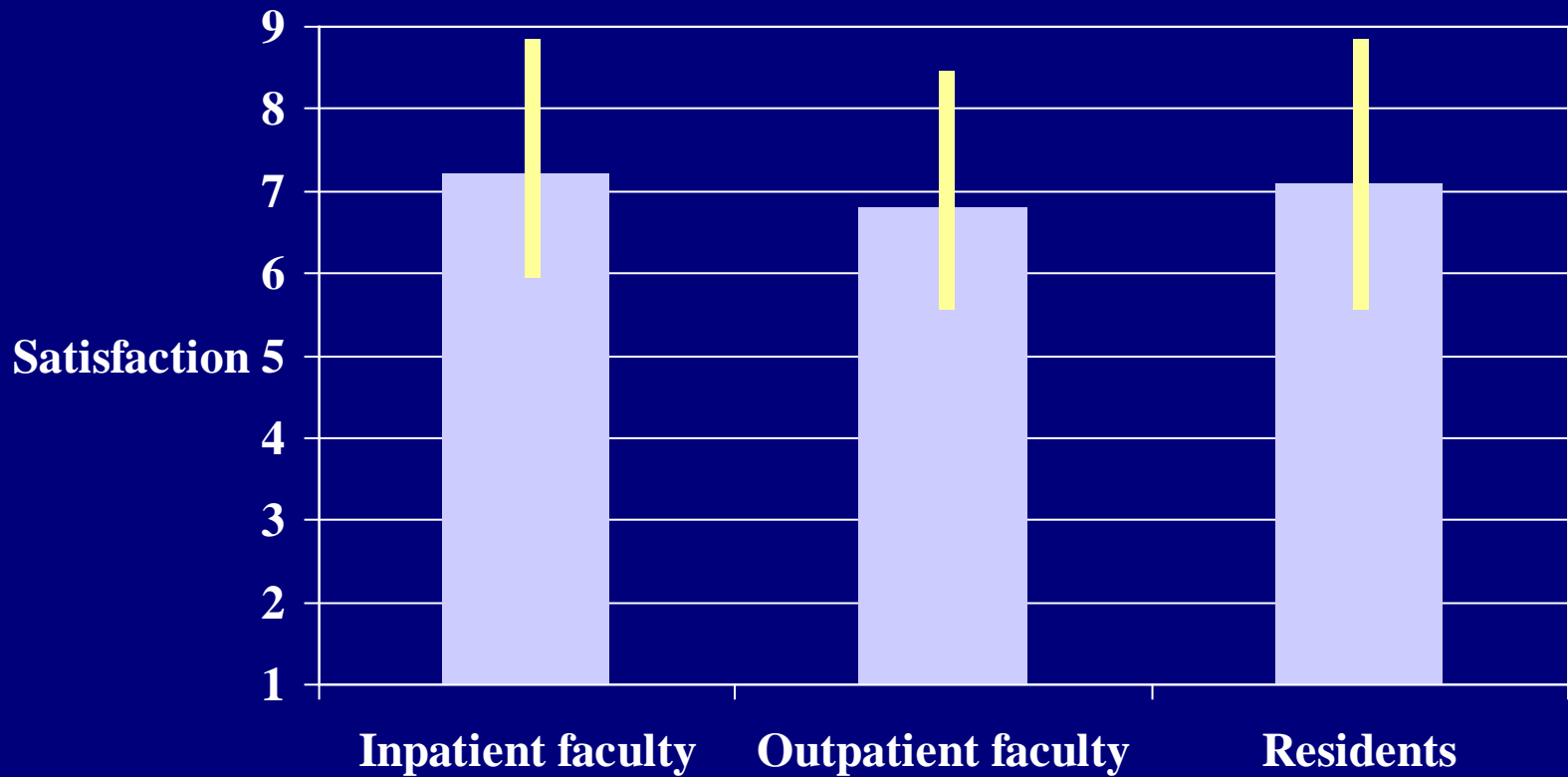
Median feedback time: 5 minutes

Evaluator Satisfaction



$p < .0001$

Student Satisfaction



$p = .03$

Reproducibility: How Many Do We Need?

| Number of Forms | Reproducibility Estimate |
|-----------------|--------------------------|
|-----------------|--------------------------|

| | |
|---|-----|
| 4 | .58 |
|---|-----|

| | |
|---|-----|
| 6 | .68 |
|---|-----|

| | |
|---|-----|
| 8 | .74 |
|---|-----|

Principle #3: Gather Enough Data “More Is Better Than Less”

TAKE HOME MESSAGE:

Multiple observations are needed to get a reliable estimate of performance.

Principle #4: Interpret the Data Appropriately “What Does it Really Mean”

What does this mean?

make a correct interpretation

consider alternative explanations

know the limits of your data

How to Study Validity

From the books....

content

concurrent

construct

consequential

Do the data behave as expected? (confirmatory)

Can we make sense of the results? (exploratory)

Example Principle 4:

Do residents and students agree on who is a good teacher? What if we used our existing evaluation data and 'looked back' to see?

Shea JA, Bellini LM. Evaluations of clinical faculty: The impact of level of learner and time of year. Teaching and Learning in Medicine: An International Journal, 2002;14:87-91.

Medical Education Research Framework

| Unit of Analyses | Focus of Research | |
|------------------|-------------------|----------|
| | Process | Outcome |
| Person | X | X |
| Program | X | X |

RATIONALE: Medical students and residents routinely evaluate teachers and attendings. The data have consequences for retention, promotion, and salary.

OBJECTIVE: To examine differences in students' and residents' ratings of the same clinical faculty and see how scores varied over the course of a year.

METHODS: Retrieved data from evaluation database for all residents and students in medicine clerkships, electives and residency.

KEY FINDINGS: Students gave more favorable ratings than residents.

Over the year, students became less critical and residents became more critical.

Winners of teaching awards had better ratings.

Example of Quarter Data

| | Q1 | Q2 | Q3 | Q4 |
|-----------|-----|-----|-----|-----|
| Students | 2.7 | 2.6 | 2.1 | 1.7 |
| | Q1 | Q2 | Q3 | Q4 |
| Residents | 1.8 | 2.1 | 2.4 | 2.7 |

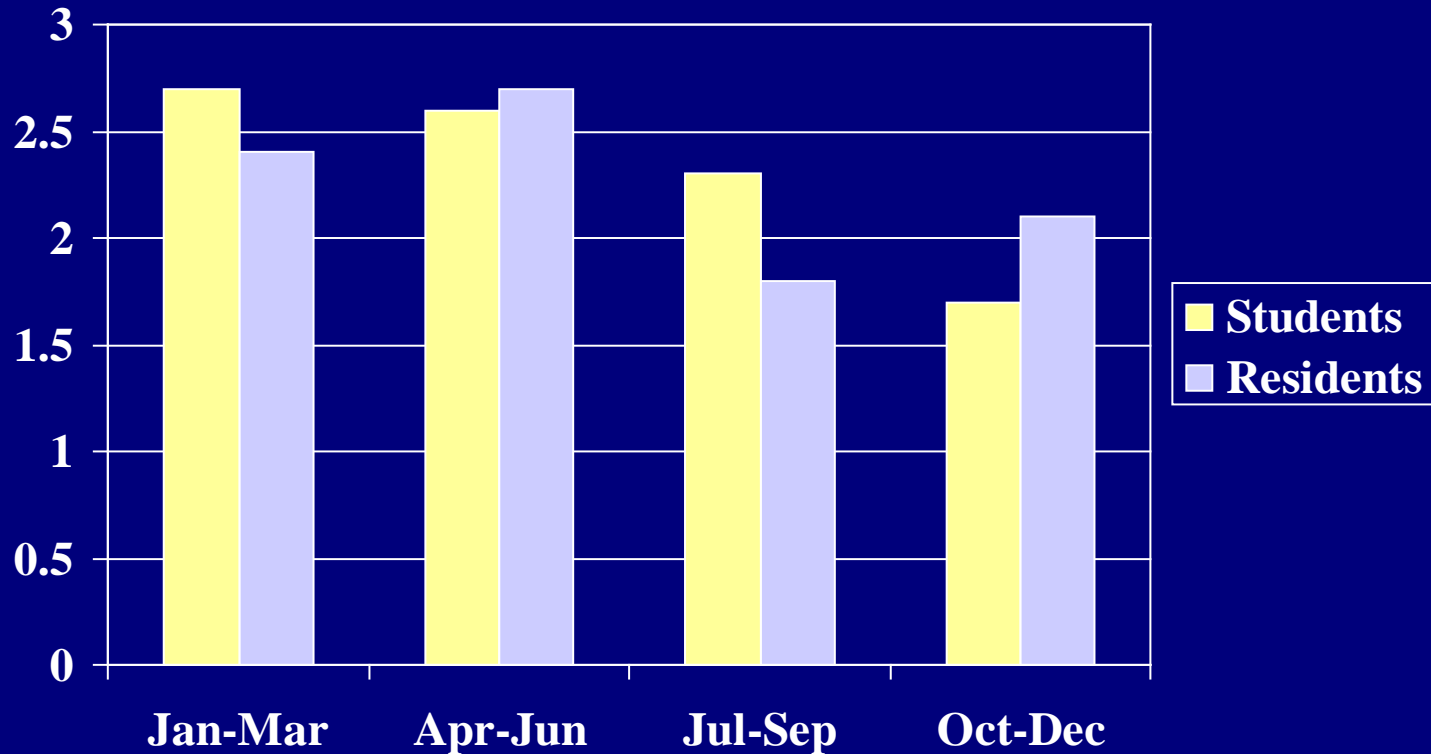
On a scale where 1 = almost always and 4 = never

But How Do the Data Really Line Up?

| | Jan-Mar | Apr-Jun | Jul-Sep | Oct-Dec |
|-----------|---------|---------|---------|---------|
| Students | 2.7 | 2.6 | 2.1 | 1.7 |
| | Jan-Mar | Apr-Jun | Jul-Sep | Oct-Dec |
| Residents | 2.4 | 2.7 | 1.8 | 2.1 |

On a scale where 1 = almost always and 4 = never

What are the Consequences?



Principle #4:
Interpret the Data Appropriately
“What Does it Really Mean”

TAKE HOME MESSAGE:

Don't just analyze the data - think about the consequences of using the data.

Summary

Medical education framework

unit of analysis

process and/or outcomes

Four research principles

enhance quality of medical education

research and in turn education

Things I Did Not Talk About

Qualitative studies

Unique challenges:

- real deadlines and timelines

- small samples

- many “competing” curricula

Abundant opportunity and need to conduct
medical education research