Tracking the Longitudinal Stability of Medical Students' Perceptions Using the AAMC Graduation Questionnaire and Serial Evaluation Surveys

Michele P. Pugnaire, MD
Mary L. Zanetti, EdD
Urip Purwono, PhD
Michele Carlin, BSc

University of Massachusetts Medical School

Background:

The student satisfaction ratings: as a measure of student perception of their medical education

- ❖ **Clerkship evaluation** surveys: Student satisfaction ratings widely used to assess education quality in clinical years
- ❖ The **AAMC GQ** survey: “universally” used to measure students’ satisfaction with their medical education upon graduation
- ❖ **Alumni** surveys: “often” used to assess long term satisfaction after graduation

Background:

Student satisfaction surveys in UME **What we know:**

- ❖ All three types of surveys are commonly used by medical schools to assess their educational program
- ❖ These provide useful measures for educational Q/A

Research Question

Student satisfaction surveys in UME **What we DON'T know:**

- ❖ How stable are students' perceptions (satisfaction ratings) of their educational experience over time ?
- ❖ How can we best interpret and apply student satisfaction ratings in assessing the quality of our education programs

Study design

- ❖ Examines the longitudinal stability of student perceptions of their medical education over time
- ❖ Compares ratings to similar items in 3 sequential surveys at 3 points in time:

**End of Clerkship Evals
EOC**

**AAMC
GQ**

**Post grad
PGY1**

Preclerkship	Clerkship	Year 4	PGY1
--------------	-----------	--------	------

Study design

- ❖ Examines the longitudinal stability of student perceptions of their medical education over time
- ❖ Compares ratings to similar items in 3 sequential surveys at 3 points in time:

Methods

For classes of '00(N=100) and '01(N=93):
Compiled/compared ratings on similar items in

EOC and GQ surveys
and
GQ and PGY1 surveys

AAMC GQ Survey

❖ Graduation Requirement for Classes '00 & '01

❖ Response Rate :100%

Class of '00: N= 100

17% excluded (declined data release)

Class of '01: N= 93

14% excluded (declined data release)

❖ Study: IRB approved, exempt status !

EOC Surveys

- ❖ “Required” and administered @ end of each yr 3 clerkship
- ❖ Year 3 clerkships: Medicine, Pediatrics, Surgery, Ob-Gyn, Psychiatry, Family Med
- ❖ Response rate: 86-100% , classes of '00, '01

End of Clerkship Survey questions adapted from GQ:

- ❖ EOC question: “Overall, rate this clerkship”
- ❖ GQ question: “Rate the educational quality of each clerkship”
- ❖ Both surveys used 4 point Likert rating:
‘excellent’/ ‘good’/ ‘fair’/ ‘poor’

PGY1 Survey

- ❖ Three-page scantron questionnaire with cover letter sent to all PGY1 alums in June/July
- ❖ Incentive for returned surveys (lottery gift)
- ❖ Response rates: '00=61%, '01=66%
- ❖ **Three Items adapted from the AAMC GQ in PGY1 survey**
 - “Hot topics” question
 - Overall satisfaction with medical education
 - Confidence that residency skills have been acquired

PGY 1 Survey questions adapted from GQ '00, '01: Hot topics question

- ❖ GQ Question: “Do you believe that the time devoted to your instruction in the following areas was ‘inadequate’, ‘appropriate’, or ‘excessive’?”
- ❖ PGY1 Question : **“Now that you have been out of medical school for a year**, do you believe that the time devoted to your instruction in each of the following areas was ‘inadequate’, ‘appropriate’, or ‘excessive’?”
- ❖ **29** of 40+ GQ hot topics items: listed verbatim on PGY1 survey

29 GQ “Hot Topic” items from GQ included “verbatim” in PGY1 Survey

29 “Hot Topics” included “verbatim” in PGY1 Survey

- ❖ “Care of hospitalized patient”
- ❖ “Drug and alcohol abuse”
- ❖ “Law and Medicine”
- ❖ “Managed care”
- ❖ “Patient follow up”
- ❖ “Primary Care”

Other “Hot Topics” excluded in PGY1 Survey

- ❖ Communication skills”
- ❖ “Interpretation of clinical data and research reports”
- ❖ “Literature reviews/critiques”
- ❖ “Interpretation of laboratory results”

Rationale for Exclusion of GQ Items

⊕ Brevity:

“Cultural diversity” as surrogate for “cultural differences and health related behaviors/ customs” & “culturally appropriate care for diverse populations”

⊕ Redundancy:

“medical interviewing”	included
“communications skills”	excluded

⊕ Covered elsewhere

“diagnostic planning” included elsewhere in survey

Other PGY 1 Survey questions adapted from GQ:

- ❖ GQ question : “Overall, I am satisfied with the quality of my medical education”
- ❖ PGY1 question : “Overall, I am satisfied with the quality of my medical **school** education”
- ❖ 4 point Likert Scale for both :
‘strongly agree’/ ‘agree’/ ‘disagree’/ ‘strongly disagree’

Other PGY 1 Survey questions adapted from GQ:

- ❖ GQ question: “I am confident that I have acquired the clinical skills required to begin a residency program”
- ❖ PGY1 question: “**At the time that I graduated, I had** acquired the clinical skills required to begin a residency program”
- ❖ 4 point Likert Scale for both:
‘strongly agree’/ ‘agree’/ ‘disagree’/ ‘strongly disagree’

Analysis Methods

- ❖ Matched responses for similar items were compared: EOC vs GQ and PGY1 and GQ
- ❖ The proportion of most positive ratings ('excellent', 'strongly agree', 'adequate') was calculated for each item by combining all other ratings to a **dichotomous rating scale**
- ❖ **Rationale:** EOC, GQ, PYG-1 responses were negatively skewed with more than 80% of responses for all 3 surveys in highest 2 categories ('excellent'/'good'; 'strongly agree'/'agree')
- ❖ This bivariate analysis tested a null hypothesis that the difference between specific proportions within EOC & GQ=0 and GQ & PGY1 =0

Analysis Methods (con't)

- ❖ Comparison analysis of :
EOC vs GQ
and
PGY1 vs GQ
- ❖ Performed by computing the significance of differences in the proportion of the most positive rating categories (**z scores**)

Results ...

End of Clerkship (EOC) vs. Graduation Questionnaire (GQ) Comparisons

Class of 2000					Class of 2001			
EOC (%) [*]	GQ (%) [*]	Z (Absolute)	<i>p</i>	Clerkship	EOC (%) [*]	GQ (%) [*]	Z (Absolute)	<i>p</i>
66	55	1.35	.18	A	67	59	1.05	.29
60	68	1.01	1.69	B	51	60	1.02	1.69
20	24	.61	1.46	C	27	39	1.47	1.86
41	36	.67	.50	D	39	33	.71	.48
34	41	.84	1.60	E	29	39	1.26	1.79
52	67	1.77	1.92	F	52	68	1.88	1.94

^{*} Percentages reflect "Excellent" responses.

N= 66-77, matched

N= 67-71, matched

N

Ratings of “Appropriate” instructional time

GQ 2000 vs. PGY1 2001 (N=58, Matched)

Ratings of “Appropriate” instructional time

GQ 2000 vs. PGY1 2001 (N=58, Matched)

Ratings of “Appropriate” instructional time

GQ 2001 vs. PGY1 2002 (N=47, Matched)

Ratings of “Appropriate” instructional time GQ 2001 vs. PGY1 2002 (N=47, Matched)

“Overall, I am satisfied with the quality of my medical education”

N=58 Matched, Class of 2000

[†] Z=.63, p=.53

[‡] Z=1.30, p=.19

[§] Z=1.75, p=.08

■ % Strongly Agree

▨ % Agree

Percent

“Overall, I am satisfied with the quality of my medical education”

N=47 Matched, Class of 2001

[†]Z=.51, p=.61

[‡]Z=.77, p=.44

[§]Z=.41, p=.68

■ %Strongly Agree

▨ % Agree

“At the time I graduated, I had acquired the clinical skills required to begin a residency program”

N=58 Matched, Class of 2000

†Z=2.26, p=.02

‡Z=1.55, p=.02

§Z=1.68, p=.09

■ %Strongly Agree
▨ % Agree

“At the time I graduated, I had acquired the clinical skills required to begin a residency program”

N=47 Matched, Class of 2001

†Z=3.11, p=.00

‡Z=2.85, p=.00

§Z=.79, p=.43

■ %Strongly Agree
▨ % Agree

Summary of results

- ❖ From year 3 to graduation, remarkable stability of students' satisfaction of their clerkships (EOC vs GQ)
- ❖ From graduation to internship across 2 cohorts, significant changes in a total of only 7 of 31 ratings (GQ vs PGY1)
- ❖ Large majority of changes observed were positive, with more favorable ratings in post PGY1 survey
- ❖ Overall, UMMS satisfaction ratings: strong in medical school and maintained post PGY1
- ❖ Trends were reproducible across 2 different sequential cohorts.

Summary of results

One major exception:

Residency Preparedness: GQ-PGY1

- ❖ Both cohorts show a large and comparable shift towards more favorable ratings post internship.

Explanation:

- ❖ Undervaluing of skill set acquired in med school
- ❖ Over-rating of the challenges of internship at graduation
- ❖ Under-confident of their own readiness

Conclusions

- ❖ Student perceptions of their medical education as measured by GQ and other “satisfaction” surveys are notably stable over time: yr 3, yr 4, internship.
- ❖ Supports the usefulness of the AAMC GQ in programmatic assessment for our educational programs, retrospectively and prospectively
- ❖ Reinforces the general value of student satisfaction surveys as Q/A measures of our educational programs.

Study Limitations

- ❖ Possible response bias (66% return rate for PGY1 Survey)
- ❖ Exclusion of GQ ratings from students declining data release:
 - This group disproportionately represents less ‘favorable’ ratings (Hodgson:AcadMed 2002,10supp)
 - May skew the ‘favorability’ of ratings, but would not necessarily affect the stability of ratings

Study Limitations

❖ Generalizability concerns:

- At UMMS : ? Other graduating classes
??Other survey items
- At other Schools: ???

Application/implications

- ❖ Feedback to course/clerkship directors and curriculum committees
- ❖ Long term outcomes measurement for accreditation (NEASC, LCME)
- ❖ Strengthens continued use of PGY1 survey at UMMS for Classes of '02 and beyond
- ❖ “GQ-Adapted Post Graduate Surveys” linked to AAMC GQ may be of use to other schools

The great unknown

The great unknown

How will the decrease in 2004 GQ respondents affect our outcomes?