

Do student evaluations influence the teaching skills of clerkship clinical faculty?

A. Chandrasekhar, J. A. McNulty, A. Janabi, R. Durazo-Arvizu
Loyola University Stritch School of Medicine, Maywood, IL 60153



PURPOSE

The goal of this retrospective study was to determine if clinical faculty who received less than optimal student evaluations improved their teaching skills in subsequent years.

DISTRIBUTION OF STUDENT EVALUATIONS

Table 1. Number and percentage of student evaluation records for each of the 1-5 rankings.

Teaching Skill Rating	Number of student ratings of faculty	%
1	373	2
2	685	3
3	2,248	11
4	4,175	21
5	12,400	62
Total number of records	19,881	100

DISTRIBUTION OF STUDENT EVALUATIONS

Table 2. Number of faculty with optimal (>3.5) and sub-optimal (<3.5) ratings of teaching skills exhibiting year-to-year trends that were "worse", "same", and "improved".

	Optimal rating	Sub-optimal rating
Worse	9	2
Same	686	43
Improved	54	0

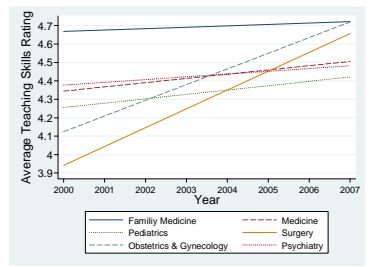
MATERIALS AND METHODS

Undergraduate medical students (120-140 per year) anonymously submitted faculty evaluations through secure web-based applications. Evaluation data for eight (8) academic years were analyzed for this report. The evaluation forms used a criteria referenced Likert rating scale of 1-5 (5 the best; 1 the worst), which, in addition to "teaching skills", included other attributes (i.e., knowledge, role modeling) and space for comments.

Descriptive statistics for faculty "teaching skills" were calculated, including mean, median, standard deviations, and proportions. A paired-sample t-test was applied to initially compare rating scores, followed by a model-based approach that accounted for repeated measures within individuals. Precisely, linear regression with random intercept was fit to the data to account for within attending correlations. This model was then applied to investigate trends over time of rating scores and differences in rating scores. Random intercept models are within the class of mixed-effects models used to study complex phenomena, in which subject responses may be correlated [6]. Graphical representations are used to better understand the results obtained with more complicated statistical models.

YEARLY TRENDS BY CLERKSHIP

Fig. 1. Regression lines showing trends for average faculty teaching skill ratings over the sample period for each of the clinical clerkships.

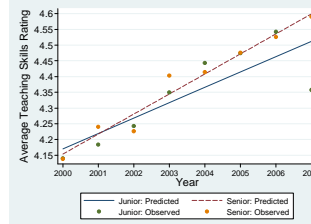


RESULTS

1. A total of 794 faculty were evaluated in more than one year allowing analysis of trends in student evaluations of their teaching skills.
2. The evaluation forms included several attributes (e.g., knowledge, role modeling) all of which showed highly significant ($p < 0.001$) correlations for individual faculty (data not shown).
3. By the criterion referenced option, 5% of faculty received a suboptimal teaching skill score (Table 1).
4. Of 45 faculty who had suboptimal evaluations, none improved, 96% received similar evaluations in subsequent years, and 4% received even worse evaluations in subsequent years (Table 2).
5. There was a trend for improved teaching over the sample period, especially in the Surgery and Ob/Gyn clerkships (Fig. 1).
6. There were not differences in the trends for improved teaching between junior (assistant professors) and senior (assoc/full professors) faculty (Fig. 2).

ASSOCIATIONS WITH GRADE

Fig. 2. Regression lines for junior and senior faculty during the sample period.



CONCLUSIONS

- 1) Clinical faculty who received sub-optimal student ratings for their teaching skills did not seem to be influenced to improve their skills.
- 2) In contrast, a large number ($n=54$; 7%) of faculty who were already viewed as having optimal teaching skills improved those skills even more, which is noteworthy considering that the average rating of faculty was already high (4.4).
- 3) One interpretation of these results is that good teachers are motivated to improve, whereas those less interested in teaching do not have this motivation.