Implementing an Anonymous Closed-Loop Feedback Process and Associations With ACGME Annual Survey Scores in a Family Medicine Residency

Jordan Knox, MD, CAQSM; Katherine T. Fortenberry, PhD; Fares Qeadan, PhD, MS, MES; Benjamin Tingey, BS; Anna Holman, MD; Kirsten Stoesser, MD; Sonja Van Hala, MD, MPH

BACKGROUND AND OBJECTIVES: The annual Accreditation Council for Graduate Medical Education (ACGME) survey evaluates numerous variables, including resident satisfaction with the training program. We postulated that an anonymous system allowing residents to regularly express and discuss concerns would result in higher ACGME survey scores in areas pertaining to program satisfaction.

METHODS: One family medicine residency program implemented a process of quarterly anonymous closed-loop resident feedback and discussion in academic year 2012-2013. Data were tracked longitudinally from the 2011-2019 annual ACGME resident surveys, using academic year 2011-2012 as a baseline control.

RESULTS: For the survey item "Satisfied that evaluations of program are confidential," years 2013-2014, 2014-2015, and 2018-2019 showed a significantly higher change from baseline. For "Satisfied that program uses evaluations to improve," year 2018-2019 had a significantly higher percentage change from baseline. For "Satisfied with process to deal with problems and concerns," year 2018-2019 showed significantly higher change. For "Residents can raise concerns without fear," years 2013-2014 and 2018-2019 saw significantly higher changes.

CONCLUSIONS: These results suggest that this feedback process is perceived by residents as both confidential and promoting a culture of safety in providing feedback. Smaller changes were seen in residents' belief that the program uses evaluations to improve, and in satisfaction with the process to deal with problems and concerns.

The Accreditation Council for Graduate Medical Education (ACGME) requires that residency programs elicit resident feedback and concerns.[1] The ACGME surveys all residency programs annually on a variety of variables, including residents' satisfaction with their program's feedback mechanisms. The ACGME provides programs with their aggregate data and benchmarked national averages. Residents' responses to these surveys are critical, as poor scores can adversely impact accreditation.

Specific guidelines for conducting feedback are absent, leaving much autonomy to individual programs. Traditional models of feedback, such as written surveys and small-group meetings,[2] can be limited in timeliness, scope, and anonymity—all aspects that are important for effective feedback.[3] Several programs have investigated anonymous feedback utilizing confidential focus groups, electronic suggestion boxes, and online surveys.[4-7] One study comparing anonymous and open (identifiable) evaluations of residency faculty

From the Department of Family and Preventive Medicine, University of Utah School of Medicine, Salt Lake City, UT.

within internal medicine showed statistically significant lower scores in anonymous feedback compared to open feedback,[8] suggesting that feedback format can influence results. Conversely, significant drawbacks to anonymous feedback include unfocused and reactionary comments that may not be representative of overall resident sentiments or constructive for meaningful change.[7,9]

In this study, we examined longitudinal changes in ACGME survey scores, before and after implementing an anonymous, closed-loop quarterly survey process. We hypothesized that an anonymous system allowing residents to regularly provide program concerns and suggestions, and discuss these openly in a group setting with administration, would correlate with higher ACGME survey scores in satisfaction and confidence with the program's feedback mechanism.

Methods

Setting
The University of Utah Family Medicine Residency Program is a 3-year program with eight residents per class, which expanded to 10 residents per class beginning in 2018. We took data from the 2011-2019 annual ACGME resident survey aggregate reports. Survey responses from 2011-2012 served as baseline data. The intervention of the feedback process began in academic year 2012-2013.

Intervention
Every quarter, (1) a web link to an anonymous online survey was emailed to all residents, providing opportunity to list concerns, suggestions, or other residency-related feedback; (2) responses were reviewed by the program director, core residency faculty, and the chief residents during regularly-scheduled meetings; and (3) survey results, as well as planned program responses, were summarized and discussed at a regularly scheduled resident business meeting. In this way, each feedback cycle was followed by closed-loop

communication to acknowledge and address the survey comments. This quarterly feedback system was a supplement to already-established feedback mechanisms, including postrotation surveys, evaluations of faculty, annual program evaluation, and didactic evaluations.

Measurements
This study evaluated four questions from the ACGME resident survey. The ACGME questions used a 5-item Likert scale (1=strongly disagree, 5=strongly agree), and were reported as an aggregate program score for each question. The ACGME survey reports aggregate national scores for program comparison. The statements evaluated are presented in Table 1.

Statistical Analysis
We calculated relative differences for each averaged response between 2011-2012 and each ensuing academic year, reported as percentage change from baseline. We used simple linear regressions to show trends. We also plotted the relative differences against academic year to show the magnitude of change compared to the first-year score. We used one-sample z test to compare each year's percentage change from baseline. We used two-sample t test to compare program average scores versus the national average, and we set significance at 5%. We performed analyses in STATA software, version 16.0 (STATA Corp, College Station, TX).

The University of Utah Institutional Review Board deemed this project nonhuman subjects research (#00130505).

Results
Fifty-six percent of the 66 residents involved in the feedback process, across the 8 years were female, and predominately identified as non-Hispanic/White (90.1%).

The plots of average scores across academic years for the four questions are shown in Figure 1, and with the exception of Question 1 (Evaluations of Program Confidential), all show a

stable or increasing trend across academic years.

Figure 2 shows the relative differences of each year's average score compared to baseline scores from 2011-2012. Scores increased for all questions, except for one instance. In no cases did scores decrease from baseline. For the statement, "Satisfied that evaluations of program are confidential," years 2013-2014, 2014-2015, and 2018-2019 showed significantly higher scores from baseline (2012-2013: 13.2% increase, P=.051; 2013-2014: 18.4% increase, P=.001; 2014-2015: 15.8% increase, P=.008; 2018-2019: 18.4% increase, P=.002). For "Satisfied that program uses evaluations to improve," only year 2018-2019 had a significantly higher percentage change (16.2% increase, P=.008). For "Satisfied with process to deal with problems and concerns," year 2012-2013 showed marginally higher change than 5%, and year 2018-2019 showed significantly higher change (2012-2013: 13.2% increase, P=.051; 2018-2019: 13.2% increase, P=.04). The last statement, "Residents can raise concerns without fear," saw significantly higher percent changes in year 2013-2014 and year 2018-2019 (2013-2014: 17.9% increase, P=.002; 2018-2019: 25.6% increase, P<.001).

Table 1 shows the comparison of program average scores to the national scores across each academic year for the four questions. For the question "Satisfied that evaluations of program are confidential," the University of Utah scored significantly lower than the nation in years 2011-2012 (baseline) and 2015-2016 (2011-2012: 3.8 vs 4.3, P=.01; 2015-2016: 3.8 vs 4.3, P=.03), but was on par in other years. No significant differences were seen in question "Satisfied that program uses evaluations to improve." The program scored marginally lower in the baseline year 2011-2012 on "Satisfied with process to deal with problems and concerns" (3.8 vs 4.2, P=.05), and on par in subsequent years. Finally, the program scored marginally higher for "Residents can raise concerns

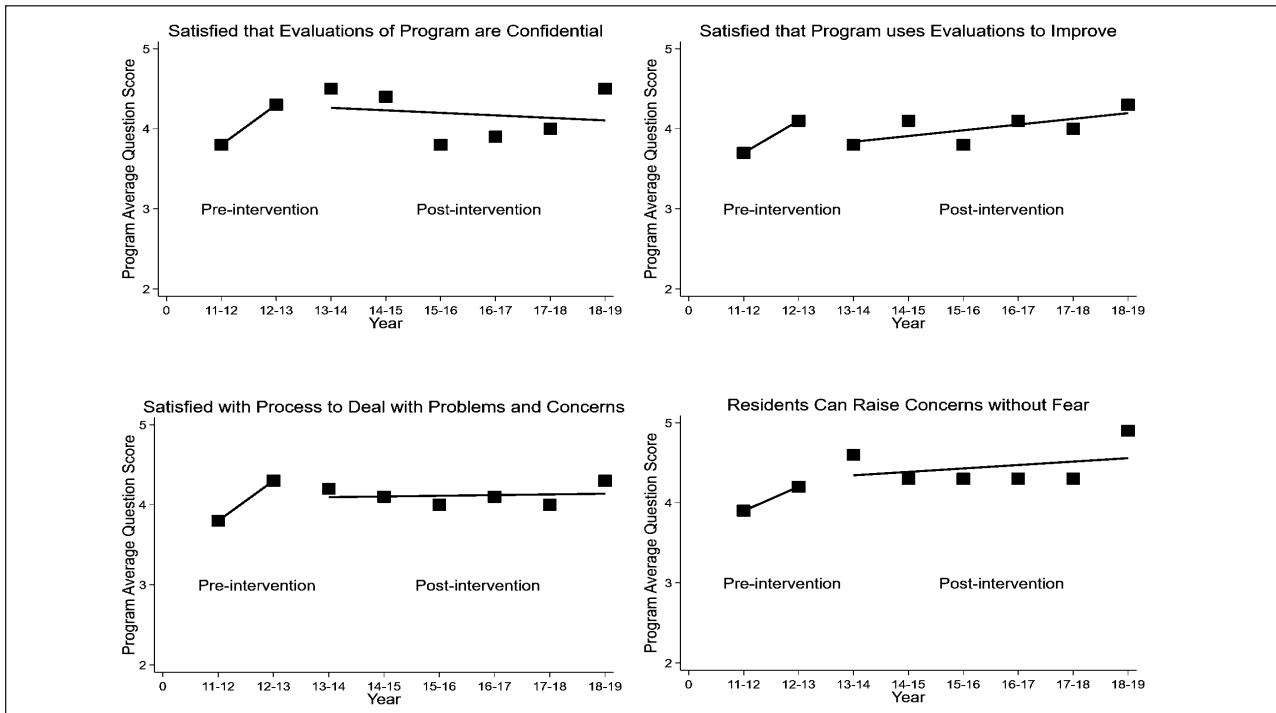**Table 1: Comparison of Program Scores to National Scores**

Academic Year	Satisfied that Evaluations of Program are Confidential (*P* Value)	Satisfied That Program Uses Evaluations to Improve (*P* Value)	Satisfied With Process to Deal with Problems and Concerns (*P* Value)	Residents Can Raise Concerns Without Fear (*P* Value)
2011-12	3.8,[a] 4.3[b] (.01)[c]	3.7, 4.0 (.14)	3.8, 4.2 (.050)	3.9, 4.2 (.14)
2012-13	4.3, 4.3 (>.99)	4.1, 4.0 (.66)	4.3, 4.2 (.66)	4.2, 4.2 (>.99)
2013-14	4.5, 4.3 (.33)	3.8, 4.0 (.33)	4.2, 4.1 (.62)	4.6, 4.2 (.050)
2014-15	4.4, 4.3 (.62)	4.1, 4.0 (.62)	4.1, 4.1 (>.99)	4.3, 4.2 (.62)
2015-16	3.8, 4.3 (.03)	3.8, 4.0 (.37)	4.0, 4.2 (.37)	4.3, 4.2 (.65)
2016-17	3.9, 4.3 (.07)	4.1, 4.1 (>.99)	4.1, 4.2 (.65)	4.3, 4.2 (.65)
2017-18	4.0, 4.4 (.07)	4.0, 4.1, (.65)	4.0, 4.2 (.37)	4.3, 4.2 (.65)
2018-19	4.5, 4.4 (.64)	4.3, 4.1 (.35)	4.3, 4.2 (.64)	4.9 ,4.3 (.005)

[a] Program mean

[b] National mean

[c] *P* value for comparisons made using two-sample *t* test for difference in means with equal variances.

Figure 1: Program Average Scores on ACGME Survey Questions (2011-2019)

Satisfied that Evaluations of Program are Confidential

Satisfied that Program uses Evaluations to Improve

Satisfied with Process to Deal with Problems and Concerns

Residents Can Raise Concerns without Fear

without fear" in 2013-2014 and significantly higher in 2018-2019 (2013-2014: 4.6 vs 4.2, *P*=.05; 2018-2019: 4.9 vs 4.3, *P*=.005).
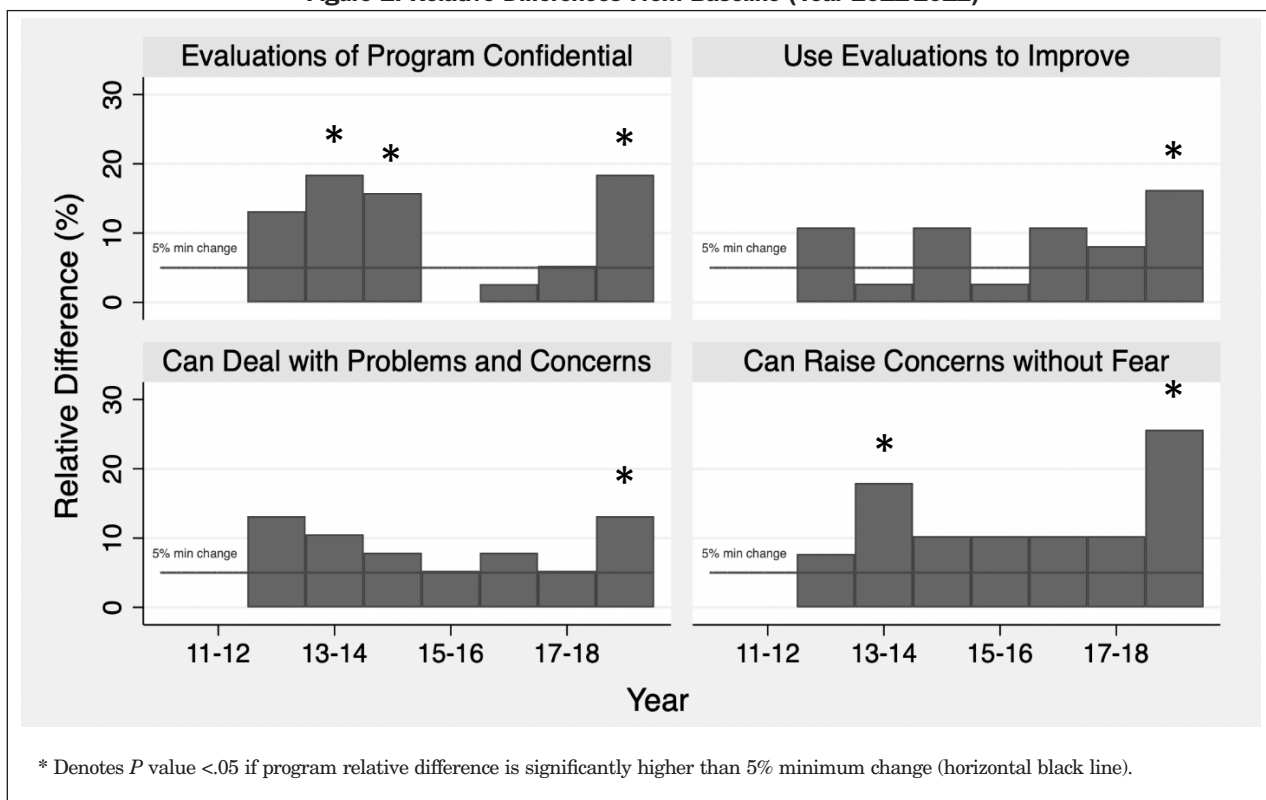
Discussion

Improvements were seen in the ACGME annual survey results following implementation of an anonymous, closed-loop feedback process. Residents perceived this process as confidential and safe. Prior research has suggested that psychological safety is an important variable in resident satisfaction and learning,[10] and anonymous feedback processes may be a useful contributor to this perception.

We observed small changes in residents' belief that the program uses evaluations to improve, and in satisfaction with the process to deal with problems and concerns. Qualitative analysis to explore this phenomenon could prove beneficial, but was beyond the scope of this project. The small improvement may relate to residents learning more about program changes that occur as a result of the closed-loop feedback, and may also reflect frustration with aspects of the program that cannot be easily changed. Prior research examining

Figure 2: Relative Differences From Baseline (Year 2011-2012)

* Denotes *P* value <.05 if program relative difference is significantly higher than 5% minimum change (horizontal black line).

the ACGME survey suggests that perceived responsiveness to resident feedback is particularly salient to many residents.[11]

Limitations of this project include the single residency program design, limiting generalizability. This study spanned 8 years of data, introducing many confounders in the work and learning environments (eg, duty hour changes, off-service rotations, shifting some elective rotations from intern year to a later year). Strengths of the project include its longitudinal nature and comparisons between the program's own data and national data. Further investigations are needed within multiple residency programs, and should incorporate a qualitative component on the feedback process.

CORRESPONDING AUTHOR: Address correspondence to Dr Jordan Knox, 375 Chipeta Way Ste A, Salt Lake City, UT 84108. 801-581-7234. Fax: 801-587-3353. Jordan.Knox@hsc.utah.edu.

References

1. Resident/Fellow and Faculty Surveys. Accreditation Council for Graduate Medical Education. Accessed October 15, 2020. https://www.acgme.org/Data-Collection-Systems/Resident-Fellow-and-Faculty-Surveys/

2. Frye AW, Hemmer PA, Frye ANNW, Hemmer PA. Program evaluation models and related theories. AMEE Guide No. 67. 2012;(67). doi:10.3109/0142159X.2012.668637

3. Myerholtz L, Reid A, Baker HM, Rollins L, Page CP. Residency faculty teaching evaluation: what do faculty, residents, and program directors want? Fam Med. 2019;51(6):509-515. doi:10.22454/FamMed.2019.168353

4. Saddawi-Konefka D, Scott-Vernaglia SE. Establishing psychological safety to obtain feedback for training programs: a novel cross-specialty focus group exchange. J Grad Med Educ. 2019;11(4):454-459. doi:10.4300/JGME-D-19-00038.1

5. Benson NM, Vestal HS, Puckett JA, et al. Continuous quality improvement for psychiatry residency didactic curricula. Acad Psychiatry. 2019;43(1):110-113. doi:10.1007/s40596-018-0908-4

6. Schwartz AR, Siegel MD, Lee AI. A novel approach to the program evaluation committee. BMC Med Educ. 2019;19(1):465. doi:10.1186/s12909-019-1899-x

7. Boazak M, Cotes RO, Ward MC, Schwartz AC. Explorations with a residency-wide, online, anonymous suggestion box: a roller coaster ride. Acad Psychiatry. 2019;43(6):627-630. doi:10.1007/s40596-019-01084-0

8. Afonso NM, Cardozo LJ, Mascarenhas OAJ, Aranha ANF, Shah C. Are anonymous evaluations a better assessment of faculty teaching performance? A comparative analysis of open and anonymous evaluation processes. Fam Med. 2005;37(1):43-47.

9. Schwartz AC, Crowell A, Stern M, Cotes RO. Revisiting an anonymous suggestion box. [published online ahead of print, 2021 Jan 12]. Acad Psychiatry. 2021;45(2):244-245. doi:10.1007/s40596-020-01391-x

10. Torralba KD, Loo LK, Byrne JM, et al. Does psychological safety impact the clinical learning environment for resident physicians? Results from the VA's Learners' Perceptions Survey. J Grad Med Educ. 2016;8(5):699-707. doi:10.4300/JGME-D-15-00719.1

11. Caniano DA, Yamazaki K, Yaghmour N, Philibert I, Hamstra SJ. Resident and faculty perceptions of program strengths and opportunities for improvement: comparison of site visit reports and ACGME resident survey data in 5 surgical specialties. J Grad Med Educ. 2016;8(2):291-296. doi:10.4300/JGME-08-02-39