

JRMC | Journal of Regional Medical Campuses

Does empathy really decline during residency training? A longitudinal look at changes in measured empathy in a community program

William J. Crump, M.D.; Craig H. Ziegler, Ph.D.; R. Steve Fricker, M.P.A.

DOI: <https://doi.org/10.24926/jrmc.v4i4X.4206>

Journal of Regional Medical Campuses, Vol. 4, Issue 4 (2021)

z.umn.edu/JRMC

All work in JRMC is licensed under CC BY-NC



Does empathy really decline during residency training? A longitudinal look at changes in measured empathy in a community program

William J. Crump, M.D.; Craig H. Ziegler, Ph.D.; R. Steve Fricker, M.P.A.

Abstract

Introduction

Some studies, most cross sectional and urban, have shown a decline in empathy during residency training prompting medical educators to consider changes in curriculum or training environment. Our aim was to determine if there was a decline using a longitudinal, paired annual empathy measure across 3 years of a family medicine residency in a rural community hospital.

Methods

We administered the Jefferson Scale of Empathy from 2015 through 2020 and of the 116 opportunities for survey completion, 112 from 48 residents were available for scoring. We also asked our residents to rank 10 factors that affected their empathy scores. The Baptist Health Madisonville IRB approved the protocol as exempt and the authors have no conflicts of interest.

Results

With a response rate of 97%, we found no statistically significant decrease in our resident scores across the 3 years. Scores after our PG-1 year were significantly lower than 2 previous comparable studies. The longitudinal, paired study design revealed very wide ranges among individual residents even when group means indicated a statistical difference. Residents also differed widely on their rankings of factors that could affect the score, with only outlook on life showing a narrow range and high ranking.

Conclusions

The very wide range of individual paired scores as well as the broad range of factors the residents thought affected their scores indicate that empathy is a very individual concept. Some of our residents increased scores leading to resilience and others declined toward cynicism. Those seeking to make changes to curriculum or training environment to facilitate empathy during residency should consider this diversity of individual resident training experience.

Introduction

Empathy is considered a key requirement to be an effective physician, but defining this characteristic is difficult. This is important for medical educators who seek to make modifications in curriculum or environment to assist learners in developing and then

maintaining empathy for patients. Some authors have considered empathy in the affective domain similar to sympathy and others consider it as more of a cognitive process much like curiosity.¹ The differentiation we used as we designed our professional identity curriculum was that sympathy is

William J. Crump, M.D., Residency Faculty, Baptist Health Madisonville Family Medicine Residency and Associate Dean of the University of Louisville Trover Campus at Baptist Health Madisonville, Madisonville, Kentucky

Craig H. Ziegler, Ph.D., Biostatistician with the University of Louisville School of Medicine Office of Graduate/Undergraduate Medical Education, Louisville, Kentucky

R. Steve Fricker, M.P.A., Director of Rural Health/Student Affairs of the University of Louisville Trover Campus at Baptist Health Madisonville, Madisonville, Kentucky

Corresponding author: William J. Crump, M.D.

Associate Dean, University of Louisville Trover Campus at Baptist Health Madisonville 200 Clinic Drive, Third North Madisonville, KY 42071 p. 270.824.3515 f. 270.824.3590 e. bill.crump@bhsi.com



"I feel your pain," and empathy is "I understand your suffering".²

The Jefferson Scale of Empathy, the most widely used and well-validated measure, was produced by a group that views empathy as largely cognitive.³ Higher values by students on this empathy measure have been associated with positive clerkship faculty ratings of student clinical competence¹ and patients with diabetes cared for by physicians with higher scores had better outcomes.⁴⁻⁵

There is some evidence that scores on these empathy measures decrease during the clinical phase of medical school, although the absolute change is small.⁶ There have been several publications documenting a decline in self-reported empathy across the years of residency training. We summarize the most pertinent ones here and they are compared in Table 1.

Table 1: Published studies showing decline of empathy during resident training

Author	Measure	Design	N, Specialty	Response Rate	Significance
Bellini ⁷⁻⁸	IRI	Longitudinal	61, Internal Medicine	72-98%	P < 0.05
Rosen ⁹	IRI	Longitudinal Paired	47, Internal Medicine	80%	P < 0.001
Mangione ¹³	JSE	Cross Sectional	98, Internal Medicine	84%	NS
Foreback ¹⁴	JSE	Cross Sectional	45, Internal Medicine	35%	NS
Wolfshol ¹⁵	JSE	Cross Sectional	35, Emergency Medicine	92%	Scores Declined ^a
Crump ³	JSE	Longitudinal Paired	18, Family Medicine	100%	P < 0.05

^aNo statistics reported

IRI = Interpersonal Reactivity Index

JSE = Jefferson Scale of Empathy

A report studied a cohort of 61 internal medicine residents at the University of Pennsylvania from 2000 to 2003.⁷⁻⁸ The measure used was the interpersonal reactivity index (IRI) and as the residents began their internship (PG-1) year they had better baseline scores for perspective taking and empathetic concern compared to general adult and college student populations. Five months into the internship, there was a statistically significant decrease in empathetic concern measures. These residents then completed the measures through the end of their 3 years of training and the empathetic concern scores never recovered and were significantly lower than when the residents began. This was one of very few studies that were longitudinal, but scores were not paired by individual resident.

Another study conducted during 2002-2003 reported 47 interns in the internal medicine program at the University of Pennsylvania.⁹ Compared to the beginning of the year, reported depression, burnout, and chronic sleep deprivation significantly worsened across the year. Measures of empathetic concern using the IRI were more favorable than the general population at the beginning of the year, declined and approached the general population norms at the end of the year, and the difference was statistically significant. This is the only study we found that paired the results, but it was only one year duration.

The IRI is an old measure that is not designed specifically for medical trainees and defines empathy as reactions of one individual to the observed experiences of another.¹⁰ More commonly used than the IRI is the Jefferson Scale of Physician Empathy (JSE), which was designed for medical trainees and practitioners and has a very large reliability and validity literature.^{1,3-4,11-12} An early study using the JSE addressed the hypothesis that empathy decreases during residency training.¹³ It was a cross-sectional study of 98 internal medicine residents at all 3 years of training at Thomas Jefferson University Hospital, with a response rate of 84%. The residents were 43% women and, contrary to the JSE reports from medical students,¹¹ women did not have higher scores. Their conclusions were that empathy remains stable across the entire 3 years of training, even though in the cross sectional design the individuals representing each year were different. They also compared JSE scores to the American Board of Internal Medicine humanistic qualities scale scores as assigned by the program director. They concluded that there was a positive but not significant relationship between the 2 measures and they thought the humanism instrument might be measuring something that is more like being respectful and sensitive rather than being empathic. They propose the JSE as the better measure.

A report from 2018 included 45 internal medicine residents in 3 community hospital training programs in Flint Michigan.¹⁴ One of the programs was osteopathic and 2 were allopathic. This was a cross-sectional study done between May and September of 2014. A baseline was done in the first few months of the first postgraduate year and the others were done

in the summer at the end of each training year. There was a 35% response rate. They conclude that empathy increased over the first year of training and then declined across the next 2 years, but none of the changes were statistically significant.

A 2019 study addressed the association of empathy and burn out among 35 emergency medicine residents and 33 attending physicians at a tertiary referral county hospital.¹⁵ It was cross-sectional and done in winter 2018/2019. They found self-reported empathy to decrease during residency training and to be higher among attendings. Reported burnout was very high among residents and was lower in the attendings. They found a weak negative correlation between self-reported empathy and the category of burnout classified as patient-related, but not the personal-related or work-related burn out categories. Statistics were not reported.

In 2018, we published a pilot one-year study of paired JSE results before and after implementation of a 6-month professional identity curriculum with 18 family medicine residents at all 3 levels of training at a rural community hospital.¹⁶ We found stable JSE scores at the end of the curriculum and then significant decline when measured 6 months after the end of the curriculum. Residents who attended more professional identity sessions showed a non-significant smaller decline, and there were large standard deviations within each training level with some individual residents showing little change across the year. This curriculum was not continued beyond the pilot because of a change in the didactics schedule.

Our goal in this study was to measure empathy using the JSE at baseline and annually across all 3 years of the family medicine residency at our rural community hospital. We used a methodology that allowed us to match each response to the individual, providing a truly longitudinal paired measure for comparison to previously published cross-sectional reports that were based in larger cities.

Methods

Setting

Our program is based in a town of 20 000 in a rural area in the upper southeast, with 6 residents in each

year and no other residencies in town. It was begun in 1971, becoming the 85th family medicine program in the United States and the first in the state.¹⁷ The mission is focused on providing rural family physicians for our region, and at the last report 49% of graduates remain in rural practice and 41% were practicing within 100 miles of the program. The host hospital is part of a statewide system, with 312 licensed beds and 140 in operation. At last report, there were 5 189 annual discharges, 26 000 ED visits, 40 000 outpatient visits, and 640 births. The site is also host for our regional rural medical school campus, with the main campus in a metropolitan community 160 miles away.¹⁸ Although administratively separate, in such a small site students and residents share all facilities, faculty, and most other resources.

Survey Process

Beginning in 2015, 48 different residents completed the empathy survey either just as the new academic year began, or just as it ended, resulting in an annual survey for each resident. Each survey had the resident's name included for later matching, and residents placed the completed surveys into an envelope confidentially. Participants were assured that a research assistant unknown to them would then assign their ID number and subsequently no one would be able to connect their responses to their name. During the study period, one resident missed completing a baseline survey, one missed the post-PG-2, and one missed the post PG-3 surveys. We excluded one resident survey because the name was not provided on the form, and so the final response rate was 112/116 (96.6%). The results reported by year of training are from the graduating classes of 2016-2023, with the full complement of all 4 surveys reported for the classes of 2019-2020.

After the residents had completed the JSE, we asked them to rank 10 factors that may have affected how they answered the empathy scale. We developed these factors during focus groups in 2015 after the conclusion of the 6-month implementation of our professional identity curriculum.¹⁶ These included 2 items that could be considered traits, 4 that were described as daily irritants that could affect empathy if it were a state that could change frequently, one that specifically addressed confidence in the doctor

role, 2 that addressed the negative effect of the “hidden curriculum”,¹⁹ and lastly the effect of continuity of patient care. Residents were asked to rank 10 items with 1= most important and 10= least important in response to “I think the following explains the empathy score of an individual resident”.

We drew resident demographics from their applications and summarized these using frequencies and percentages. We compared JSE scores among ours and 2 previous reports at baseline, post PGY 1, post PGY 2, and post PGY 3 using one way analysis of variance at each time. If we found significant differences, we performed a Tukey post hoc test. We used a one-way analysis of variance to find any difference among the training years as well as to compare ours with the other 2 previous reports.

For our data, we created spaghetti plots showing the wide variation among individual resident scores. We used IBM SPSS Statistics for Windows (version 26.0, 2019, IBM Corporation, Armonk, NY, 877-426-6006) and the website statpages.info²⁰ to analyze the data. We created figures with the R package [ggplot2](https://ggplot2.tidyverse.org/).²¹ Statistical significance was set by convention at $p < 0.05$.

Results

As shown in Table 2, most residents were male and white, with significant proportions of Hispanic and Asian. Almost 21% were older than those who went directly through schooling with no gap years. Seventy-three percent graduated from medical schools outside of the U.S., and 28% were rural.

Table 2: Demographics of residents (N=48)

		Freq	(%)
Gender	Male	31	64.6%
Race/Ethnicity	White	22	45.8%
	Asian	17	35.4%
	Hispanic	6	12.5%
	Black	2	4.2%
	Other	1	2.1%
Age at Residency Entry	Median	28	
Age > 31 at Residency Entry		10	20.8%
Graduated from a US Medical School		13	27.1%
Rural ^{a,b}		11	27.5%

^aRural was defined as a hometown population of <30,000 and a non-metro Rural Urban Continuum Code (RUCC).²²

^bN=40. Eight residents were born outside of U.S. and RUCC could not be applied to them to determine rurality.

The second column in Table 3 shows that there were no significant changes across the 3 years of training in our population. The last column and row labeled baseline shows that at baseline one previous study had a higher score than ours as well as higher than the baseline from the other program shown, but this was not significant. Our residents showed a significant decline after the PG-1 year when compared to either of the other studies. As shown graphically in Figure 1, it is noteworthy that one of the programs showed an increase in empathy scores after the PG-1 year.

Table 3: JSE means of training by year of 3 programs

JSE	Current Data		Foreback ¹⁴		Mangione ¹³		ANOVA comparison among programs by PGY year
	N	Mean (SD)	N	Mean (SD)	N	Mean (SD)	
Baseline	30	111.65 (13.80)	15	110.4 (14.6)	40	117.5 (12.4)	F(2,82)=2.40, p=0.098
Post PGY 1	28	105.14(14.99) ^{a,b}	11	116.8 (9.8) ^a	27	114.0 (14.3) ^b	
Post PGY 2	27	106.23 (16.09)	8	108.1 (14.1)	31	113.5 (10.8)	F(2,63)=2.15, p=0.126 T=1.15, df=36, p=0.259
Post PGY 3	27	108.37 (15.47)	11	114.3 (11.4)			
	F(3,108)=1.049, p=.374 ^c		F(3,41)=0.947, p=0.427 ^d		F(2,95)=1.09, p=0.341 ^e		

^aReflect p-value of .057 between Current Data and Foreback using Tukey's post hoc test

^bReflect p-value of .057 between Current Data and Mangione Tukey's post hoc test

^cReflects one way analysis of variance for differences of means for Current Data.

^dReflects one way analysis of variance for differences of means as reported by Foreback

^eReflects one way analysis of variance for differences of means as reported by Mangione.

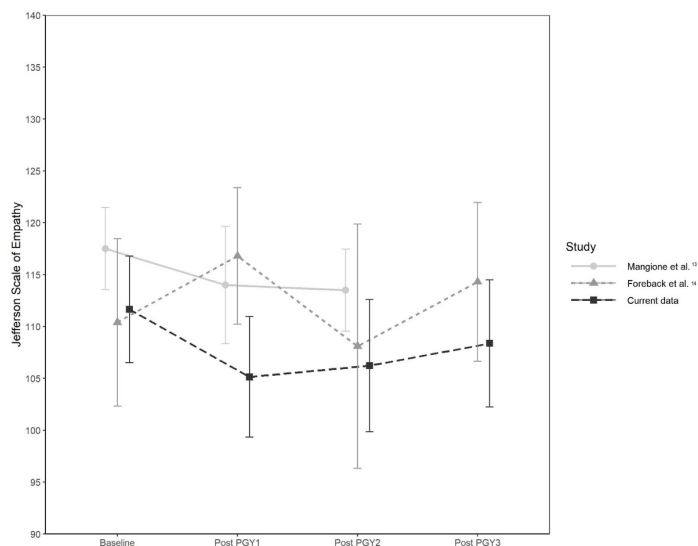


Figure 1: Changes of empathy across time for residents of selected programs

The above comparisons use group means to demonstrate differences. Given the relatively large standard deviations, Figure 2 shows the individual resident paired scores across all 4 measures, and Figure 3 shows the same for a larger group including some who began our program before we started doing a baseline survey.

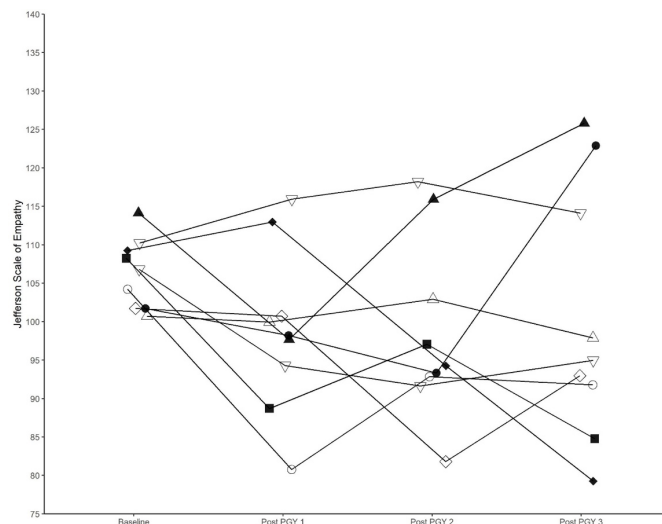


Figure 2: Individual resident JSE scores by training year (N=9)

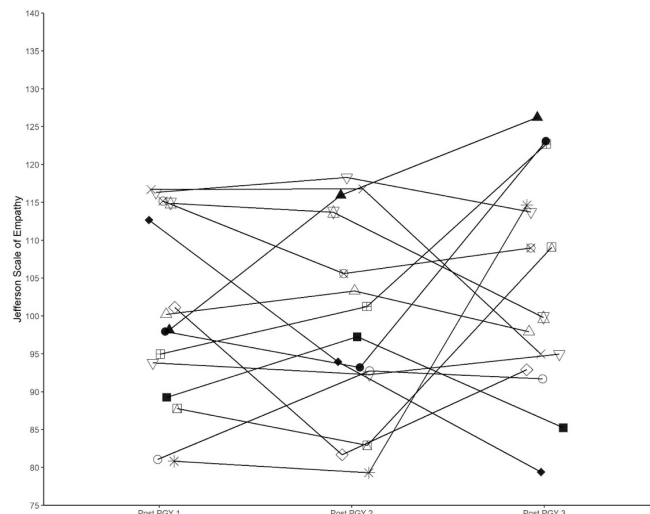


Figure 3: Individual resident JSE scores by training year (N=15)

Table 4 shows the ranking of the 10 factors that the residents thought explained their empathy scores, with general outlook on life clearly the highest rank with a narrow range. All the others showed very wide ranges, with continuity of care the lowest ranking factor.

Table 4. Ranking of empathy factors identified by residents^a (n = 41).

Factors	Average Ranking of Importance	Range of Importance Rankings
General outlook on life	2.3	1-6
Adequate sleep the night before the survey	4.0	1-10
Amount of admin work not associated with learning expected that day	4.2	1-10
Confidence in one's role as a doctor	5.5	1-10
Exposure to more cynical faculty	5.9	1-10
Exposure to more cynical residents	5.9	1-10
On a difficult rotation	6.3	1-10
Childhood experiences	6.7	1-10
Too much time in clinic	7.1	3-10
Continuity of care with patients	7.2	1-10

^aAnswers in response to "Please rank order each

factor from 1 to 10, using each number only once:

1 = most important in explaining an individual resident's empathy score;

10 = least important in explaining an individual resident's empathy score"

Discussion

For many, the evidence that empathy declines during residency is not convincing.⁶ The 3 previous publications that reported JSE score decline were cross sectional and did not find any statistically significant differences.¹³⁻¹⁵ The 2 reports that did find a significant decline during residency training were done using the IRI that is designed to be used in the general population. These were both done in the same university hospital in Philadelphia, and both were limited to internal medicine.⁷⁻⁹ The IRI did have a moderate correlation ($r=0.45$, $R^2=0.20$, $p<.01$) with the JSE in a study at another Philadelphia university hospital internal medicine program.²³ Even with this finding, 80% of the variation in the IRI was not explained by the JSE, suggesting they may not be measuring exactly the same thing.

Focusing just on studies using the JSE, there are very wide standard deviations around the means, including the results reported here. How can we explain that in our previous study there was a significant decline after our professional identity curriculum ceased and in this study that we found significantly lower scores after the PG-1 year when compared to the other 2 comparable reports? It is tempting to echo what many other reports have cited

as explanations including the dehumanizing effects of medical training and the stress and time demands of residency.^{7-8,24} There is certainly some degree of truth in those explanations, probably different among programs and perhaps changing with more recent duty hour limitations and night float routines, but it may be that empathy is such an individually defined concept that inferential statistics may not provide a view of the full truth.

There are clear advantages to using means to display inferential statistical results, but in the case of empathy, this may obscure more important findings. In our previous study, the significant decline when looking at all 18 residents after the end of our professional identity curriculum was almost entirely explained by the sharp decline in our PG-3 class, as there were no significant changes across the year in the other 2 classes.¹⁶ This prompted us to look more closely at individual scores in the current study.

Our study design allows us to match individual scores across all longitudinal JSE measures. Review of Figure 2 shows that while the mean did decrease from baseline to post PG-1, 3 of the 9 residents showed very little change and 2 actually increased. In addition, one showed a marked decrease after the PG-2 year, 2 showed a remarkable recovery after the PG 2 year, and one showed a steady decline after the PG-1 year. Figure 3 shows similar findings including a class that began before we started doing the baseline measures, thus increasing the group to 15. This mirrors what we previously reported where the means changed very little during the 6-month implementation of our professional identity curriculum and then significantly decreased 6 months after it was completed.¹⁶ Looking at the individual JSE scores from that study, only 6 of the 18 residents showed a decrease, 8 showed little change, and 4 actually increased.

These findings suggest that even using a very reliable and valid survey, measured empathy varies widely among individual residents, as they experience their training time differently. This raises the possibility that there may be individual resident characteristics or other life experiences that modulate the effect of the training period, leading some to resilience and others to cynicism. This assumes that the exposure of

each resident to the effects of the “hidden curriculum”¹⁹ that may erode empathy is similar. The original use of this term referred to the institutional policies, resource allocation, evaluation systems, and institutional “slang” that explain what a student learns beyond the formal and informal curriculum.

Assuming similar exposure to the hidden curriculum is a questionable assumption in most large programs where month to month or even week to week the resident works in different environments from those their classmates experienced during the same rotation. Our study site provides an opportunity to decrease this potential confounder significantly, as each resident works with the same small faculty of various specialties, residents at other levels, and nursing staff in the same physical facility as their classmates on each rotation, with no residents of other specialties or fellows present.

To begin to explain the wide range of JSE responses, we asked our residents first to brainstorm in a focus group all the factors that might affect a resident’s score on any given day, and then individually to rank their perceived importance. As shown in Table 4, the rankings varied widely. Some of these items are the daily irritants of being a resident and others, including the highest ranked item of “outlook on life” would be more stable day to day. Our resident responses highlight the problem with the overly simplistic “state versus trait” debate about empathy and many other interpersonal attribute measures.²⁵⁻²⁹ If our residents saw empathy as more of a trait that is stable day to day, childhood experiences would not be expected to be ranked so low. The high rankings of sleep deprivation and administrative work suggest that many of our residents see empathy as more of a state that can change as often as daily. Perhaps the most important message from Table 4 is the very wide variability among our residents, further supporting that measured empathy is a nebulous and highly individual concept.

Strengths and Limitations

As with all single site studies including all previous reports on this issue, ours is subject to selection bias and limited generalizability. Residents who match to a rural regional hospital in an upper southeastern town of 20 000 are likely very different from those who match to university hospitals in Philadelphia^{7-9,13} and

even from those at community hospitals in a town of 100 000 in Michigan.¹⁴ Since none of those previously reported sites are representative of residencies nationwide, our report may provide some balance for the evidence, and other study sites are needed.

None of the previous publications provided details of their resident demographics for comparison. As other programs with different demographics and larger resident classes study measures of empathy, it would be useful to compare these to our findings. Our results are also subject to type 2 error because of small group sizes, although ours is comparable in size to most previous studies. All but one previous report using the JSE were cross sectional, and our study used longitudinal, matched data with a loss rate of only 3%. Future studies from other sites would be more useful if a high response rate and matched individual results could be reported.

Our next research effort is to look more closely at individual factors that may be associated with an individual resident’s resilience and susceptibility to empathy decline during training. This will include demographics, measures of burnout, and individual resident opinions of the value of various methods to avoid burnout. We will also look more deeply into the overlap among faculty-perceived resident empathy and other resident characteristics such as respect, sensitivity, and curiosity.

Conclusions

In our resident population, we found some non-significant declines in mean measures of empathy and our post PG-1 mean values were significantly lower than the 2 previous comparable studies. However, using individual paired comparisons there appear to be subsets of our residents who react very differently during their training time. Some are resistant to decline or even improve measures of empathy, demonstrating resilience, and others move towards cynicism. Interventions to change residency training to facilitate empathy should consider this diversity of individual resident experience.

References

1. Hojat M, Gonnella JS, Mangione, et al. Empathy in medical students as related to academic performance, clinical competence, and gender. *Med Educ.* 2002;36:522-527. doi: 10.1046/j.1365-2923.2002.01234.x.
2. Crump, WJ. Professional identity curriculum at the University of Louisville Trover campus: reflection and meaning in medical education. *J KY Acad Fam Physicians.* 2017;Winter:88:18.
3. Hojat M, Vergare MJ, Maxwell K, Brainard G, Herrine SK, Isenberg GA, et al The devil is in the third year: A longitudinal study of erosion of empathy in medical school. *Acad Med.* 2009;84(9):1182-1191. doi: 10.1097/ACM.0b013e3181b17e55. Erratum in: *Acad Med.* 2009 Nov;84(11):1616.
4. Hojat M, Louis DZ, Markham FW, Wender R, Rabinowitz C, Gonnella JS. Physicians' empathy and clinical outcomes for diabetic patients. *Acad Med.* 2011;86(3):359-364. doi: 10.1097/ACM.0b013e3182086fe1.
5. Del Canale S, Louis DZ, Maio V, et al. The relationship between physician empathy and disease complications: An empirical study of primary care physicians and their diabetic patients in Parma, Italy. *Acad Med.* 2012;87(9):1243-1249. doi: 10.1097/ACM.0b013e3182628fbf.
6. Colliver JA, Conlee MJ, Verhulst SJ, Dorsey JK. Reports of the decline of empathy during medical education are greatly exaggerated: a reexamination of the research. *Acad Med.* 2010; 85:588-593.
7. Bellini LM, Baime M, Shea JA. Variation of mood and empathy during internship. *JAMA* 2002; 287:3143-3146.
8. Bellini LM, Shea JA. Mood change and empathy decline persist during three years of internal medicine training. *Acad Med.* 2005; 80:164-167.
9. Rosen IM, Gimotty PA, Shea JA, Bellini LM. Evolution of sleep quantity, sleep deprivation, mood disturbances, empathy, and burn out among interns. *Acad Med.* 2006; 81:82-85.
10. Davis MH. Measuring individual differences in empathy: evidence for a multidimensional approach. *J of Personality and Social Psychology.* 1983; 44:113-126. doi: 10.1037/0022-3514.44.1.113.
11. Hojat M, Gonnella JS. Eleven years of data on the Jefferson Scale of empathy-medical student version (JSE-S): Proxy norm data and tentative cutoff scores. *Med Princ Pract.* 2015;24(4):344-350. doi: 10.1159/000381954. Epub 2015 Apr 28.
12. Hojat M, Shannon SC, DeSantis J, Speicher MR, Bragan L, Calabrese LH. Does empathy decline in the clinical phase of medical education? A nationwide, multi-institutional, cross-sectional study of students at DO-granting medical schools. *Acad Med.* 2020;95(6):911-918. doi: 10.1097/ACM.0000000000003175.
13. Mangione S, Kane GC, Caruso JW, Gonnella JS, Nasca TJ, Hojat M. Assessment of empathy in different years of internal medicine training. *Medical Teacher.* 2002;24(4):370-373.
14. Foreback J, Kusz H, Lepisto BL, Pawlaczyk B. Empathy in internal medicine residents at community-based hospitals: a cross-sectional study. *J Med Educ Curric Dev.* 2018;5:1-4.
15. Wolfshol JA, Bradley K, Bell C, Bell S, et al. Association between empathy and burn out among emergency medicine physicians. *J Clin Med Res.* 2019; 11(7): 532-538.
16. Crump WJ, Ziegler CH, Fricker RS. A residency professional identity curriculum and a longitudinal measure of empathy in a community-based program. *J Reg Med Campuses.* 2018;1(4). doi: 10.24926/jrmc.v1i4.1353 .
17. Martin DA. A short history of Trover Clinic with commentary. McClanahan Publishing House: Kuttawa, Kentucky, 1989.
18. Crump WJ, Fricker RS, Ziegler CH, Wiegman DL. Increasing the rural physician workforce: a potential role for small rural mMedical school campuses. *J Rural Health.* 2016 Jun;32(3):254-9. doi: 10.1111/jrh.12156. Epub 2015 Oct 30.
19. Hafferty FW. Beyond curriculum reform: confronting medicine's hidden curriculum. *Acad Med.*1998 73(4):403-407.
20. Pezzullo JC. Analysis of variance from summary data: StatPages website. Accessed February 3, 2021. <https://statpages.info/anova1sm.html>.

21. Wickham H. (2016). *Elegant graphics for data analysis* ggplot website. New York, N.Y.: Springer-Verlag; <https://ggplot2.tidyverse.org>. Accessed February 3, 2021.
22. United States Department of Agriculture, Economic Research Service. 2003 Rural Urban Continuum Codes. <http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx> last accessed September 17, 2020.
23. Hojat M, Mangione S, Kane GC, Gonnella JS. Relationships between scores of the Jefferson Scale of Physician Empathy (JSPE) and the Interpersonal Reactivity Index (IRI). *Medical teacher*. 2005;27(7):625-628.
24. Shanafelt TD, West C, Zhao X, et al. Relationship between increased personal well-being and enhanced Empathy among Internal Medicine Residents. *J Gen Intern Med*. 2005;20(7): 177-183.
25. Costa-Drolon E, Verneuli L, Manolios E, Revah-Levy A, Sibeoni J. Medical students' perspectives on empathy: a systematic review and metasynthesis. *Acad Med*. 2021;96:142-154.
26. Schwartz ZH. Commentary on an excerpt from *daybreak*: thoughts on the prejudices of morality. *Acad Med*. 2020;95 (9):1382-1383.
27. Sulzer SH, Feinstein NW, Wendland CL. Assessing empathy development in medical education: a systematic review. *Med Educ*. 2016;50:300-310.
28. Spiro HM. Empathy: An introduction and what is empathy and can it be taught? In *empathy and the practice of medicine*. Spiro HG, Curnen MGM, Peschel E, St. James D, eds. Yale University Press, New Haven and London, 1993, pp1-14.
29. Glick SM. The empathetic physician: Nature and nurture. In *empathy and the practice of medicine*. Spiro HG, Curnen MGM, Peschel E, St. James D, eds. Yale University Press, New Haven and London, 1993, pp 85-102.