2025, Volume 57, Issue 7, 1-7, e-ISSN 1938-3800

Family Medicine

ORIGINAL ARTICLE

The Impact of Length of Training on Clinical Preparedness Among New Graduates: A Report From the Length of Training Pilot Study in Family Medicine

Patricia A. Carney, PhD, MS[a]; Steele Valenzuela, MS[a]; Annie Ericson, MA[a]; Dang H. Dinh, MS[a]; Colleen Conry, MD[b]; Lars E. Peterson, MD, Phd[c,d]; Alan Douglass, MD[b]; Stephanie Rosener, MD[e,f]; W. Perry Dickinson, MD[b]; Mark T. Nadeau, MD, MBA[g]; Karen Mitchell, MD[h]; James C. Martin, MD[g]; M. Patrice Eiff, MD[a]

AUTHOR AFFILIATIONS:

[a] Department of Family Medicine, School of Medicine, Oregon Health & Science University, Portland, OR

[b] University of Colorado School of Medicine, CU Anschutz Medical Campus, Denver, CO

[c] American Board of Family Medicine, Lexington, KY

[d] Department of Family and Community Medicine, University of Kentucky, Lexington, KY

[e] Department of Family Medicine, Mountain Area Health Education Center, Asheville, NC

[f] University of North Carolina School of Medicine, Chapel Hill, NC

[g] Family Medicine Residency Program, Long School of Medicine, University of Texas Health Science Center, San Antonio, TX

[h] Student and Resident Initiatives, American Academy of Family Physicians, Leawood, KS

CORRESPONDING AUTHOR:

Patricia A. Carney, Department of Family Medicine, School of Medicine, Oregon Health & Science University, Portland, OR, carneyp@ohsu.edu

HOW TO CITE: Carney PA, Valenzuela S, Ericson A, et al. The Impact of Length of Training on Clinical Preparedness Among New Graduates: A Report From the Length of Training Pilot Study in Family Medicine. *Fam Med.* 2025;57(7):1-7.
doi: 10.22454/FamMed.2025.801918

PUBLISHED: 5 June 2025

KEYWORDS: clinical preparedness, family medicine residency, graduate medical education, length of training

ABSTRACT

Background and Objectives: Associations between training length and clinical preparedness are unknown. We compared assessments of clinical preparedness for family medicine graduates from 3-year and 4-year training programs.

Methods: In this prospective case-control study, we compared responses from two surveys, which occurred 3 months after graduates started their first job. One survey was conducted by a supervising physician. The other was conducted by a clinic staff member who rated family medicine graduates from both 3-year and 4-year programs.

Results: Our study included 403 graduates of 3-year programs, 185 who trained in 4-year programs with 36 months of training (4YR–36) and 274 who trained in 4-year programs with 48 months of training (4YR–48). Physician assessor ratings were similar across study groups on 18 of 21 Entrustable Professional Activities. The 4YR–48 graduates were rated higher for "practicing independently" on providing a usual source of comprehensive, longitudinal medical care for people of all ages (86.5% vs 77.9%); managing prenatal (63.1% vs 41.2%); and labor, delivery, and postpartum care (41.4% vs 25.7%). For five care process areas, physician assessors were more likely to rate 4YR–48 graduates as having "no challenges" with speed/timing related health care visits (91.9%) compared to 3YR graduates (82.4%). We noted no differences according to study group for staff member assessors.

Conclusions: We found several differences in clinical preparedness according to length of training in this pilot study. Comprehensive longitudinal care, including prenatal and maternity care, were rated higher among graduates of 4YR–48 programs.

INTRODUCTION

The optimal length of training in family medicine, which varies globally between 2 and 5 years,[1] has been debated for decades.[2–5] Some argue that additional training is needed to ensure preparedness for independent practice given duty hour restrictions, expanding curriculum needs, and evolving clinical and payment models for primary care practice. Others argue that the current length (2 years in Canada and 3 years in the United States) is adequate, and residency curriculum could be revised to address emerging issues. These debates have lacked rigorous data to inform them. Current assessments of clinical preparedness have been based on resident self-reporting or

program directors' evaluations prior to graduation.[6,7] Such approaches can be affected by recall and social response bias. For example, program directors may want to avoid being perceived as not preparing graduating residents fully, and residents may inflate or deflate their assessment of themselves.[8,9]

To explore these issues, the Length of Training Pilot (LoTP) in family medicine was funded by the American Board of Family Medicine Foundation in 2013.[10] One of the study's core questions was designed to explore the extent to which training length affected preparedness for independent clinical practice. We sought to overcome weaknesses in prior preparedness assessments by developing new survey instruments designed to be completed by two independent observers of new graduates: a supervising physician and a staff member in the graduate's new practice.[11] A key feature of our assessment was the extent to which new graduates requested assistance as a marker of their preparedness for independent practice. The relationship between requesting assistance/supervision and progressive independence is well established in clinical psychology and medical education literature.[12–16] In fact, consequences have been identified when progressive independence is not well-instituted or erodes during training.[12] The purpose of this paper is to report findings on clinical preparedness of graduates with 3 versus 4 years of training in the LoTP.

METHODS

Length of Training Pilot

The LoTP was designed to explore the impact that length of training, 3-year (3YR) versus 4-year (4YR), has on several program and learner outcomes in family medicine.[10] Briefly, it is a mixed methods prospective case-control pilot study including 17 residencies that applied for and were selected to participate in the study (seven 3YR civilian programs, six 4YR civilian programs, four 4YR Navy programs). The training models varied with four 4YR programs having a required fourth year with an integrated curriculum and two 4YR programs having an optional fourth year. The Navy programs were excluded because predetermined military objectives differ from civilian programs. All evaluation activities were overseen by researchers in the Department of Family Medicine at Oregon Health & Science University (OHSU). All LoTP programs obtained Institutional Review Board approval, and the evaluation team at OHSU was granted an educational exemption (IRB # 9770). Prior published papers describe other findings derived from the LoTP.[11–14,17–19]

Instrument Development and Testing

A prior publication details the methods undertaken to develop and test the surveys used in this study.[11] Briefly, we performed a literature review of assessments for independent clinical practice, conducted key informant interviews with nine rural and urban family physicians regarding how best to assess recent graduates of family medicine residencies, and then convened leaders in family medicine, expert evaluators, and other stakeholders to take part in a Clinical Preparedness Measurement Summit in 2015 to guide decisions related to measuring preparedness for independent clinical practice.

At the Summit meeting, we defined clinical preparedness as "the extent to which graduates of family medicine residency training are independent/self-reliant in practicing core skills in the care of patients." We also determined that the settings in which the care was provided needed to be comprehensive, including outpatient, inpatient, and other care settings (eg, home, nursing homes). We used the 2015 Entrustable Professional Activities (EPAs) for Family Medicine End of Residency Training as the framework for assessment of clinical preparedness by physician assessors.[20] For the instrument designed for staff assessors, we identified nine validated variables in published literature that assessed interpersonal communication skills[21] and seven validated variables designed to measure processes of care delivery.[22]

The final draft surveys contained 61 items for the physician assessor and 36 items for the staff assessor. Two question sets on the physician assessor survey included scaled items. The first set assessed preparedness issues that we identified with our key informant surveys (eg, confidence and timing). This scale included three responses (1=no challenges, 2=some challenges, 3=many challenges). The second set focused on the EPAs and used the following scale: 1=not practicing very independently, frequently requests assistance; 2=practicing mostly independently, sometimes requests assistance; 3=practicing independently, rarely requests assistance.

Prior to administration, the question sets were extensively pilot tested using cognitive interviewing techniques[23]. Both surveys are available for review in our previously published paper.[11] Neither of these surveys collected qualitative data; thus, only quantitative data are presented here. We assessed graduates after being oriented to their new position but before significant on-the-job learning could occur (eg, 3 months after the graduates started their first position as a clinician).

Data Collection

Author A.E. worked with residency program coordinators to determine where graduating residents were entering practice after completion of training. She then contacted these practices to identify the graduates' start dates, along with the best physician and staff assessors to receive surveys after new graduates had been on the job for 3 months. We asked participants to complete the surveys within 2 weeks, and up to three reminders were sent at 2-week intervals. Data collection started in 2016 and continued through 2023, with an overall response rate of 42.1% for the physician assessor survey and 39.0% for the staff assessor survey.

Data Analyses

As in other LoTP analyses, we found that some graduates of required 4YR programs completed only 36 months of training because they graduated before the 4-year curriculum was fully implemented. Also, the optional 4YR programs had residents who graduated after 36 months of training. To accommodate these differences, we used both an intention-to-treat analysis[24] or program level analysis and an as-treated

analysis,[25] leading to three groups for analysis: (a) graduates of 3YR programs; (b) all graduates of 4YR programs with either 36 or 48 months (4YR–36/48); and (c) graduates of 4YR programs with 48 months of training (4YR–48).

We used descriptive statistics to characterize graduates' and assessors' demographic information by length of training group, including mean and standard deviation as well as frequencies and percentages, and we summarized physician and staff assessor demographic information. To assess differences between groups, we used the Fisher exact test or χ^2 test for categorical variables and the independent samples t test with unequal variances for continuous variables.

Frequencies and percentages of assessment for preparedness by physician and staff were summarized according to each response category. For response options "not done in practice" and "cannot assess at this time," we considered these a single category when deriving P values. Missing variables were excluded from analyses.

P values were derived from χ^2tests, testing differences between 3YR and 4YR–36/48 groups and between 3YR and 4YR–48 groups. To determine effect sizes beyond statistical significance, we calculated φ coefficients for dichotomous outcomes and Cramer's V for outcomes greater than two levels. φ coefficients typically range from 0 to 1 with an effect size of about 0.30 considered medium or meaningful.[26] Cramer's V is calculated based on the degrees of freedom used in analyses (based on the number of rows and columns). For our study, a meaningful effect size (medium) ranged between .15 and .21.[27] Effect sizes with negative values indicate a reverse relationship. We performed analyses using R software version 4.3.0 (R Foundation). All tests were two-tailed, and we set α at 0.05.

RESULTS

Study Participants

The study population included 403 graduates of 3YR programs and 459 graduates of 4YR programs (36- and 48-month programs combined). Of these, 274 (59.7%) received 48 months of training (Table 1). The mean age across study groups ranged from 33.1 years (SD=3.9) to 33.7 (SD=3.7). The majority of graduates were non-Hispanic White (64.3% to 70.2%), US medical school graduates (70.5% to 85.4%), who had not enrolled in or completed a fellowship (65.7% to 75.2%); though more graduates in 4YR programs reported having completed or were in fellowship training (*P*<.006).

Characteristics of physician and staff assessors were similar across study groups with one exception (Tables 2 and 3). Staff assessors of 3YR graduates tended to be medical assistants, while assessors of both 4YR–36/48 and 4YR–48 graduates tended to be registered nurses (50.8% vs 39.5% to 39.8%; *P*<.004).

Physician assessors indicated that they typically worked 7 to 10 half days per week and that the information sources they used to assess graduates included direct observations (81.7% to 86.5%) and interactions with physicians (82.4% to 85.8%; Appendix Table A). The majority of physician assessors indicated that they consulted with new graduates frequently (69.0% to 70.5%). For 3YR program graduates, 82.2% of staff indicated that they worked every day or nearly every day with their physician colleague, compared to 68.8% in the 4YR–48 group. Across all groups, direct observation and interactions with the physician were the most common sources used by staff in their assessments. Among those that reported practice setting, academic health centers (14.8% to 18.4%), federally qualified health centers (18.4% to 22.3%), and hospital/health system-owned medical practices (9.8% to 16.9%) were all represented. Cramer's V yielded a meaningful effect size for 3YR graduates compared to 4YR–36/48 and 4YR–48 graduates' comparisons of "How often do you work with this physician while caring for patients?"

Assessments of Clinical Preparedness for Independent Practice

Physician assessor ratings of EPAs were similar across study groups except for three items (Appendix Table B). First, for providing a usual source of comprehensive, longitudinal medical care for people of all ages, 86.5% of 4YR–48 graduates were "practicing independently" compared to 77.9% for 3YR program graduates (*P*=.036). Second, for managing prenatal care, 41.2% of 3YR graduates were "practicing independently" versus 63.1% for 4YR–48 graduates (*P*=.007). Of note is that 40.4% of 3YR program graduates did not provide prenatal care in practice, compared to 21.6% of 4YR–48 graduates. For managing labor, delivery, and postpartum care, both the 4YR–48 and 4YR–36/48 groups had a higher percentage of graduates "practicing independently" (4YR–36/48, 38.6%, *P*=.014; 4YR–48, 41.4%, *P*=.005) compared to 25.7% for 3YR graduates. Again, of note is that 54.4% of graduates of 3YR programs don't provide maternity care in their practice, compared to 37.8% among graduates in the 4YR–48 group. Third, when assessing specific care processes, physician assessors were more likely to rate 4YR–48 graduates as having "no challenges" with speed/timing related health care visits (91.9%) compared to 3YR graduates (82.4%, *P* =.030). Effect sizes yielded from Cramer's V for all significant P values were meaningful effect sizes.

Findings from staff assessors are shown in Appendix Table C. The only interpersonal communication characteristic noted to be different among the groups was in explaining the rationale for their care plans or actions with a rating of "always" for 82.2% of 3YR graduates compared to 70.1% or 68.9% of 4YR–36/48 or 4YR–48 graduates (*P*=.026 and *P*=.016). Processes of care characteristics rated as "among the best" were similar across all study groups with a range of 45.7% to 71.9%. For staff assessments of challenges with care process areas, graduates rated as having "many challenges" ranged from 0% to 4.3% with no significant differences among the comparison groups. Cramer's V effect sizes for all significant *P* values were meaningful.

TABLE 1. Characteristics of Graduates Included in Clinical Preparedness Analyses

Characteristic	Graduates of 3YR programs (3YR–36)*	Graduates of 4YR programs (4YR–36/48)**	Graduates of 4YR programs (4YR–48)***	3YR vs 4YR–36/48	3YR vs 4YR– 48
	N=403	N=459	N=274	*P* value	*P* value
Age (in years), mean (SD)	33.1 (3.9)	33.7 (3.7)	33.5 (2.9)	.016	.105
Gender identity, n (%)				.017	.210
Male	140 (34.7)	197 (42.9)	109 (39.8)		
Female	263 (65.3)	262 (57.1)	165 (60.2)		
Race/ethnicity, n (%)				.353	.747
Non-Hispanic White	258 (64.3)	322 (70.2)	189 (69.0)		
Hispanic	29 (7.2)	34 (7.4)	17 (6.2)		
Non-Hispanic Black	17 (4.2)	18 (3.9)	12 (4.4)		
Non-Hispanic Asian/PI	70 (17.5)	60 (13.1)	42 (15.3)		
Non-Hispanic AI/AN	0	1 (0.2)	0		
Other/multiracial	27 (6.7)	24 (5.2)	14 (5.1)		
US medical school graduate, n (%)				.085	**.039**
Yes	284 (70.5)	353 (77.0)	234 (85.4)		
No	54 (13.4)	46 (10.0)	26 (9.5)		

*36 months of training in a 3YR program
**36 or 48 months of training in a 4YR program
***48 months of training in a 4YR program
Abbreviations: YR, year; SD, standard deviation; PI, Pacific Islander; AI/AN, American Indian/Alaska Native

DISCUSSION

In this study of the impact of length of training in family medicine, we found that for three of the 21 EPAs for residency graduates, those receiving 4 years of training were more likely to be rated by physician assessors as "practicing independently" 3 months into practice compared to graduates of 3YR programs. This was true for the 4YR–36/48 and 4YR–48 groups, which suggests that something about these 4YR programs results in their graduates being better prepared. Graduates of 4YR programs were better prepared for EPAs of providing a usual source of comprehensive longitudinal medical care for people of all ages; managing prenatal care; and managing labor, delivery, and postpartum care. Possibly these graduates purposely chose a 4YR program because they wanted to provide a full scope of services to their patients. This finding is reinforced by the fact that the practice settings did not differ among graduates of 3YR and 4YR programs. Given the rising complexity of patient care and shortages of well-trained primary care clinicians, producing graduates better prepared is of value to physician employers and society as a whole.

We know that pregnancy care provided by family physicians is declining,[28,29] which is less impactful in urban areas than in rural areas, which have fewer, if any, obstetricians. Four-year residency programs are resisting this trend by offering more robust pregnancy care curricula. These graduates are greatly needed, particularly in rural and underserved areas.

We also found that speed/timing related to visits was rated by physician assessors as having fewer challenges for graduates with 48 months of training. Perhaps the additional year of training, which a prior LoTP paper found to result in approximately 1,000 more patient visits at graduation,[18] better prepared these graduates, compared to graduates of 3YR programs, to work more effectively from the beginning, perhaps by integrating complex medical issues into a single visit. Interestingly, we found no differences in staff members' assessments of new graduates according to length of training. Staff assessments focused on interpersonal behaviors and communication skills rather than assessments of clinical care, which may indicate that these areas of practice are more sufficiently developed after 3 years of training.

We found that many variables did not differ between graduates of 3YR and 4YR programs, which suggests that many areas of practice can be well-developed after 3 years of training. Caring for patients with complex and chronic medical and mental issues in multiple settings appeared not to differ across study groups; similarly, we found no differences in managing acute illnesses, performing common procedures, managing end-of-life care, using data and best science to optimize and coordinate care, and providing leadership within an interprofessional team. These last variables represent more recently emerged competencies for family physicians; 3 years of training appears to prepare graduates well in these areas.

Of note is that graduates of 4YR programs reported having undertaken fellowship training. We suspect that these respondents were considering their fourth year as a fellowship year. In addition, graduates of 3YR programs were more likely to have been assessed by a medical assistant rather than a nurse, which more likely occurred among 4YR graduates. This finding suggests a potential difference in practice staffing that affects which practices graduates choose to join post training.

TABLE 2. Characteristics of Physician Assessors of Clinical Preparedness and of Practice Setting

Characteristics	Assessors of 3YR graduates (3YR)*	Assessors of 4YR graduates (4YR–36/48)**	Assessors of 4YR graduates (4YR–48)***	3YR vs 4YR–36/48	3YR vs 4YR–48
Physician assessors	**N=142**	**N=141**	**N=112**	***P* value**	***P* value**
Age (in years), mean (SD)	46.8 (10.7)	46.1 (9.8)	45.7 (9.8)	.604	.405
Gender, n (%)				.143	.243
Female	65 (48.5)	55 (39.6)	45 (40.9)		
Male	67 (50.0)	82 (59.0)	64 (58.2)		
Gender nonconforming	2 (1.5)	2 (1.4)	1 (0.9)		
Race/ethnicity, n (%)				.692	.355
White	92 (64.8)	90 (63.8)	71 (63.4)		
Black	6 (4.2)	4 (2.8)	1 (0.9)		
Asian/Pacific Islander	14 (9.9)	15 (10.6)	12 (10.7)		
American Indian/Alaskan Native	0	0	0		
Hispanic	5 (3.5)	10 (7.1)	7 (6.2)		
Other	7 (4.9)	9 (6.4)	9 (8.0)		
Missing	18 (12.7)	13 (9.2)	12 (10.7)		
Board certifications, n (%)				.190	.060
ABFM	102 (71.8)	120 (85.1)	98 (87.5)		
ABIM	15 (10.6)	6 (4.3)	3 (2.7)		
ABP	2 (1.4)	2 (1.4)	1 (0.9)		
AOA	4 (2.8)	2 (1.4)	2 (1.8)		
Other	9 (6.3)	8 (5.7)	5 (4.5)		
Missing	10 (7.0)	3 (2.1)	3 (2.7)		
Fellowship training, n (%)	28 (19.7)	43 (30.5)	31 (27.7)	.027	.122
Other degrees besides MD, n (%)	25 (17.6)	29 (20.6)	21 (18.8)	.142	.292
Currently precepts learners,†n (%)	96 (67.6)	116 (82.3)	95 (84.8)	.011	.006
Burnout score, mean (SD)	3.9 (1.7)	3.5 (1.7)	3.5 (1.7)	.115	.128

*36 months of training in a 3YR program
**36 or 48 months of training in a 4YR program
***48 months of training in a 4YR program
†Medical students, residents, fellows
Abbreviations: YR, year; SD, standard deviation; ABFM, American Board of Family Medicine; ABIM, American Board of Internal Medicine; ABP, American Board of Pediatrics; AOA, American Osteopathic Association; IQR, interquartile range

These findings are important given the novel feature of assessing new graduates after being in their practice for only 3 months—an interval we felt would best reflect their readiness for independent practice as a result of their residency training rather than the education that occurred on the job. One of the final papers to emerge from the LoTP will address self-reported scope of practice, assessed 1 year after graduation, and will shed additional light on the impact length of training has on practice.

Strengths of this study included the effort that went into developing and testing the clinical preparedness instruments used and the inclusion of programs from geographically diverse areas and settings. Weaknesses included the less than desirable response rates of around 40%, which are not generalizable to the study groups and suggest that we could likely have biases, including recall and social response biases. Though the literature has suggested that progressive independence typically includes self-reliance,[15,16] some personality characteristics may have influenced new graduates' requests for assistance, though we have no reason to believe this would differ across study groups. Differentiation between the study groups could become more apparent with a larger sample.

The 4YR programs were selected based on interest in the LoTP and their ability to implement a 4-year curriculum. So, the programs that produced these cohorts of graduates may not be fully comparable. Another weakness is that this study was designed as a pilot and was not fully powered to conduct a rigorous assessment of hypotheses, though it did provide invaluable information about effect sizes across our study groups. We did find some differences according to who the staff assessor was (eg, registered nurse vs medical assistant), which may have caused some measurement errors. We also choose not to account for multiple comparisons to avoid Type 1 errors because that would have resulted in us setting the α level at $P<.0004$. Doing that would have increased the likelihood of

TABLE 3. **Characteristics of Clinic Staff Assessors of Clinical Preparedness and of Practice Setting**

Characteristics	Assessors of 3YR graduates (3YR)*	Assessors of 4YR graduates(4YR–36/48)**	Assessors of 4YR graduates (4YR-48)***	3YR vs 4YR–36/48	3YR vs 4YR–48
Staff assessors	**N=118**	**N=124**	**N=93**	***P* value**	***P* value**
Age (in years), mean (SD)	38.8 (12.8)	39.3 (11.3)	38.5 (11.2)	.731	.887
Gender, n (%)				.803	.915
Female	104 (92.0)	112 (91.8)	84 (91.3)		
Male	6 (5.3)	8 (6.6)	6 (6.5)		
Gender nonconforming	3 (2.7)	2 (1.6)	2 (2.2)		
Race/ethnicity, n (%)				.437	.637
American Indian/Alaska Native	0	0	0		
Asian/Pacific Islander	4 (3.4)	2 (1.6)	2 (2.2)		
Black	10 (8.5)	12 (9.7)	5 (5.4)		
Hispanic	29 (24.6)	28 (22.6)	24 (25.8)		
White	53 (44.9)	70 (56.5)	52 (55.9)		
Other	14 (11.9)	9 (7.3)	8 (8.6)		
Missing	8 (6.8)	3 (2.4)	2 (2.2)		
Profession, n (%)				<.001	.004
Medical assistant	60 (50.8)	34 (27.4)	28 (30.1)		
Licensed practical nurse	8 (6.8)	19 (15.3)	13 (14.0)		
Registered nurse	27 (22.9)	49 (39.5)	37 (39.8)		
Other	23 (19.5)	22 (17.7)	15 (16.1)		
How long at current job (inyears), Mean (SD)	5.6 (7.3)	5.3 (5.3)	5.1 (5.2)	.764	.632
Percent of time spent in direct patient care, median (IQR)	90 (35)	90 (40)	90 (26.3)	.623	.769
Work with learners†	98 (83.1)	107 (86.3)	82 (88.2)	.465	.340
Burnout score, mean (SD)	2.6 (1.5)	2.9 (1.7)	3.2 (1.8)	.266	.039

*36 months of training in a 3YR program
**36 or 48 months of training in a 4YR program
***48 months of training in a 4YR program
†Medical students, residents, fellows
Abbreviations: YR, year; SD, standard deviation

Type 2 errors (making false negative conclusions).

The EPAs that provided a framework for our survey were developed in 2015; since then, in 2022, the Accreditation Council for Graduate Medical Education Family Medicine Review Committee and the American Board of Family Medicine (ABFM) developed Core Outcomes consisting of 12 competencies measuring the foundational knowledge, skills, and activities needed at the beginning of independent practice.[30] The EPAs did not seem to find a place in residencies, with a minority of programs using them as an assessment tool.[30] Unfortunately, the Core Outcomes were not available when we were developing our instruments. Future research on clinical preparedness should use the ABFM's 12 competencies, and ideally the new ACGME AIRE program in family medicine will continue researching length of training. That programallows residencies to pursue innovation and offers freedom from program requirements with ongoing assessment of outcomes, as long as the residencies have approval from the ACGME specialty's Review Committee and ABFM.[31]

CONCLUSIONS

In conclusion, the differences we found related to a longer length of training appear to be primarily related to comprehensive, longitudinal relationships with patients; pregnancy care; and speed and timing in providing patient care, which were more discernable by the physician assessors than the staff assessors. Most competency variables did not differ according to length of training.

Financial Support

The Length of Training Pilot is sponsored by the Accreditation Council for Graduate Medical Education and is funded by the American Board of Family Medicine Foundation.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the stakeholders who attended the Clinical Preparedness Summit held in Portland, Oregon, September 2015. These include Austin Bailey, MD (Director of Primary Care, Colorado Health Medical Group); Todd Bodner, EdD (Portland State University); Freddy Chen, MD, MPH (STFM); Stan Kozakowski, MD (AAFP); Joseph Mazzola, DO (Osteopathic Representative); Mike Mazzone, MD (AFMRD); Amy McGaha, MD (AAFP Commission on Education); Tim Munzing, MD (RPS Consultant, former RC member); Thomas O'Neill, PhD (ABFM Psychometrician); Katie Patterson, MD (Stakeholder, Indianola, Mississippi); Michael Peabody, PhD (ABFM Psychometrician); Michael Rabovsky, MD (Stakeholder, Cleveland Clinic); and Russell Thomas, DO (Stakeholder, Eagle Lake, Texas).

REFERENCES

1. Tannenbaum D, Kerr J, Konkin J. Length of training in the core family medicine residency: a paper prepared by the Working Group on Postgraduate Curriculum Review. *College of Family Physicians of Canada.* 2012. https://www.cfpc.ca/CFPC/media/Resources/Education/Triple-C-Length-of-Training.pdf.
2. Raîche P. Should family medicine residency be 3 years? NO. *Can Fam Physician.* 2009;55(4):347-348.
3. Lehmann F. Should family medicine residency be 3 years? YES. *Can Fam Physician.* 2009;55(4):344.
4. Orientale E. Length of training debate in family medicine: idealism versus realism?. *J Grad Med Educ.* 2013;5(2):192-194.
5. Sairenji T, Dai M, Eden AR, Peterson LE, Mainous AG, Iii. Fellowship or further training for family medicine residents. *Fam Med.* 2017;49(8):618-621.
6. Bérubé S, Ayad T, Lavigne F, Lavigne P. Resident's preparedness for independent practice following Otorhinolaryngology-Head and Neck Surgery residency program: a cross-sectional survey. *Eur Arch Oto-Rhino-Laryngol.* 2021;278(11):551-555.
7. Smith BK, Rectenwald J, Yudkowsky R, Hirshfield LE. A Framework for understanding the association between training paradigm and trainee preparedness for independent surgical practice. *JAMA Surg.* 2021;156(6):535-540.
8. Adams AS, Soumerai SB, Lomas J, Ross-Degnan D. Evidence of self-report bias in assessing adherence to guidelines. *Int J Qual Health Care.* 1999;11(3):187-192.
9. Wang HH, Lin YH. Assessing physicians' recall bias of work hours with a mobile app: interview and app-recorded data comparison. *J Med Internet Res.* 2021;23(12):26763.
10. Length of Training Pilot Study. 2022. http://www.lotpilot.org.
11. Carney PA, Ericson A, Conry C, Martin JC, Douglass AB, Eiff MP. Measuring clinical preparedness after residency training: development of a new instrument. *Fam Med.* 2024;56(1):16-23.
12. Carney PA, Conry CM, Mitchell KB. The importance of and the complexities associated with measuring continuity of care during resident training: possible solutions do exist. *Fam Med.* 2016;48(4):286-293.
13. Eiff MP, Ericson A, Waller E. A comparison of residency applications and match performance in 3-year vs 4-year family medicine training programs. *Fam Med.* 2019;51(8):641-648.
14. Carney PA, Ericson A, Conry CM. Financial considerations associated with a fourth year of residency training in family medicine: findings from the Length of Training Pilot Study. *Fam Med.* 2021;53(4):256-266.
15. Kennedy T, Regehr G, Baker GR, Lingard LA. Progressive independence in clinical training: a tradition worth defending. *Acad Med.* 2005;80(10):106-111.
16. Franzone JM, Kennedy BC, Merritt H, Casey JT, Austin MC, Daskivich TJ. Progressive independence in clinical training: perspectives of a national, multispecialty panel of residents and fellows. *J Grad Med Educ.* 2015;7(4):700-704.
17. Carney PA, Valenzuela S, Ericson A. The association between length of training and family medicine residents' clinical knowledge: a report from the Length of Training Pilot Study. *Fam Med.* 2023;55(3):171-179.
18. Eiff MP, Ericson A, Dinh DH. Resident visit productivity and attitudes about continuity according to 3 versus 4 years of training in family medicine: a length of training study. *Fam Med.* 2023;55(4):225-232.
19. Eiff MP, Ericson A, Dinh DH, et al. Postresidency practice setting and clinical care features according to 3 versus 4 years of training in family medicine: a Length of Training Pilot Study. *Fam Med.* 2025;56(5):302-307.
20. Association of Family Medicine Residency Directors. 2015. https://www.afmrd.org/page/epa.
21. Joshi R, Ling FW, Jaeger J. Assessment of a 360-degree instrument to evaluate residents' competency in interpersonal and communication skills. *Acad Med.* 2004;79(5):458-463.
22. Lockyer J. Multisource feedback in the assessment of physician competencies. *J Contin Educ Health Prof.* 2003;23(1):4-12.
23. Beatty PC, Willis GB. Research synthesis: the practice of cognitive interviewing. *Public Opin Q.* 2007;71(2):287-311.
24. Gupta SK. Intention-to-treat concept: a review. *Perspect Clin Res.* 2011;2(3):109-112.
25. Smith VA, Coffman CJ, Hudgens MG. Interpreting the results of intention-to-treat, per-protocol, and as-treated analyses of clinical trials. *JAMA.* 2021;326(5):433-434.
26. Ialongo C. Understanding the effect size and its measures. *Biochem Med (Zagreb).* 2016;26(2):150-163.
27. Bobbitt Z. How to interpret Cramer's V (with examples). *Statology.* 2021. https://www.statology.org/interpret-cramers-v.
28. Tong ST, Makaroff LA, Xierali IM, Puffer JC, Newton WP, Bazemore AW. Family physicians in the maternity care workforce: factors influencing declining trends. *Matern Child Health J.* 2013;17(9):576-577.
29. Fashner J, Cavanagh C, Eden A. Comparison of maternity care training in family medicine residencies 2013 and 2019: a CERA program directors study. *Fam Med.* 2021;53(5):331-337.
30. Newton W, Cagno C, Hoekzema GS, Edje L. Core outcomes of residency training 2022 (provisional). *Ann Fam Med.* 2023;21(2):191-194.
31. Newton WP, Hoekzema G, Magill M, Fetter J, Hughes L. The promise of AIRE. *Ann Fam Med.* 2022;20(4):389-391.