

ORIGINAL ARTICLE

Measuring Clinical Preparedness After Residency Training: Development of a New Instrument

Patricia A. Carney, PhD, MS^a; Annie Ericson, MA^b; Colleen Conry, MD^c; James Martin, MD^d; Alan Douglass, MD^c; M. Patrice Eiff, MD^a

AUTHOR AFFILIATIONS:

^aSchool of Medicine, Oregon Health & Science University, Portland, OR

^bDepartment of Family Medicine, School of Medicine, Oregon Health & Science University, Portland, OR

^cUniversity of Colorado, Denver, CO

^dLong School of Medicine, University of Texas Health Science Center at San Antonio, San Antonio, TX

CORRESPONDING AUTHOR:

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

HOW TO CITE: Carney PA, Ericson A, Conry C, Martin J, Douglass A, Eiff MP. Measuring Clinical Preparedness After Residency Training: Development of a New Instrument. *Fam Med.* 2024;56(1):16–23. doi: [10.22454/FamMed.2023.973082](https://doi.org/10.22454/FamMed.2023.973082)

PUBLISHED: 13 September 2023

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

© Society of Teachers of Family Medicine

ABSTRACT

Background and Objectives: Research on preparedness for independent clinical practice typically uses surveys of residents and program directors near graduation, which can be affected by several biases. We developed a novel approach to assess new graduates more objectively using physician and staff member assessors 3 months after graduates started their first job.

Methods: We conducted a literature review and key informant interviews with physicians from varying practice types and geographic regions in the United States to identify features that indicate a lack of preparedness for independent clinical practice. We then held a Clinical Preparedness Measurement Summit, engaging measurement experts and family medicine education leaders, to build consensus on key indicators of readiness for independent clinical practice and survey development strategies. The 2015 entrustable professional activities for family medicine end-of-residency training provided the framework for assessment of clinical preparedness by physician assessors. Sixteen published variables assessing interpersonal communication skills and processes of care delivery were identified for staff assessors. We assessed frequencies and compared survey findings between physician and staff assessors in 2016 to assist with survey validation.

Results: The assessment of frequencies demonstrated a range of responses, supporting the instrument's ability to distinguish readiness for independent practice of recent graduate hires. No statistical differences occurred between the physician and staff assessors for the same physician they were evaluating, indicating internal consistency.

Conclusions: To learn about the possible impact of length of training, we developed a novel approach to assess preparedness for independent clinical practice of family medicine residency graduates.

INTRODUCTION

Residency training is designed to prepare physicians for independent clinical practice. However, studies have indicated that not all residents feel prepared to practice independently.^{1,2} A 1998 national survey study assessed 2,626 residents at the time of training completion in internal medicine, pediatrics, family practice, and other specialties.¹ More than 10% of residents in each specialty reported feeling unprepared to undertake one or more tasks common in their disciplines. Much has happened in graduate medical education since these early studies, including the development and implementation of the Accreditation Council of Graduate Medical Education (ACGME) general competencies,³ ACGME milestones, now in its second iteration,⁴ and a recent movement toward competency-based

graduate medical education.⁵

Recent studies have revisited preparedness for independent clinical practice in various disciplines, including family medicine.^{6–10} Collectively, these studies continue to report that graduates of residency training appear not fully prepared for independent clinical practice. A few caveats deserve mention. Some of these studies were conducted outside the United States.^{9,10} Also, the vast majority of these studies used resident self-reported or program directors' survey findings close to the time of residency graduation.^{6,7,9,10} Survey studies use subjective measures, often affected by different types of biases. Program directors may be affected by social response bias, to avoid being perceived as graduating residents not fully prepared. Several survey studies of physicians have revealed that

self-reporting or recall bias affects findings, where inflation or deflation of their perceived performance exists.^{11,12} What is needed are more objective measures of assessment to assess preparedness more accurately for independent practice.

In family medicine, questions about the length of training have been debated for more than 2 decades,^{13,14} leading the American Board of Family Medicine Foundation to fund the Length of Training Pilot (LoTP).¹⁵ A core question this pilot study was designed to explore was the extent to which length of training affects preparedness for independent clinical practice. The purpose of this paper is to report on the development and pilot testing process for two new instruments and to publish versions that may be beneficial to other graduate medical education researchers.

METHODS

Brief Overview of the Length of Training Pilot

The LoTP (2013–2023) was a mixed methods prospective case-control pilot study designed to explore several learner outcomes related to the length of training, including scope of practice, preparedness for independent practice, and clinical knowledge.¹⁵ A number of published papers related to this study can provide additional background.^{16–20} Briefly, 17 residency programs in good standing with ACGME and who agreed to participate in required evaluation activities were selected to participate. All evaluation activities were overseen by researchers in the Department of Family Medicine at Oregon Health & Science University (OHSU). All LoTP programs obtained local institutional review board (IRB) approval, and OHSU's IRB granted an educational exemption to obtain data from study sites (IRB #9770).

Instrument Development Planning

Figure 1 illustrates the timeline for instrument development and testing. The LoTP evaluation team began working on instrument development in the Spring of 2015, initially conducting a literature review on existing studies that assessed preparedness for independent clinical practice. To avoid use of resident self-reporting or program directors' assessments, we sought to develop more objective measures that would be feasible to deploy and sensitive enough to measure differences in clinical preparedness between 3-year and 4-year training models. 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 the care was to be provided in needed to be comprehensive, including outpatient, inpatient, and other (home, long-term care facilities, specialty care facilities). This definition was used to introduce the survey to expert observer respondents.

We conducted eight key informant interviews with nine rural and urban family physicians between May and June of 2015 regarding how best to assess recent graduates of family medicine residency training. One interviewee was from a solo practice; two were from small family medicine clinics, one of which was from a multigroup health system; three were

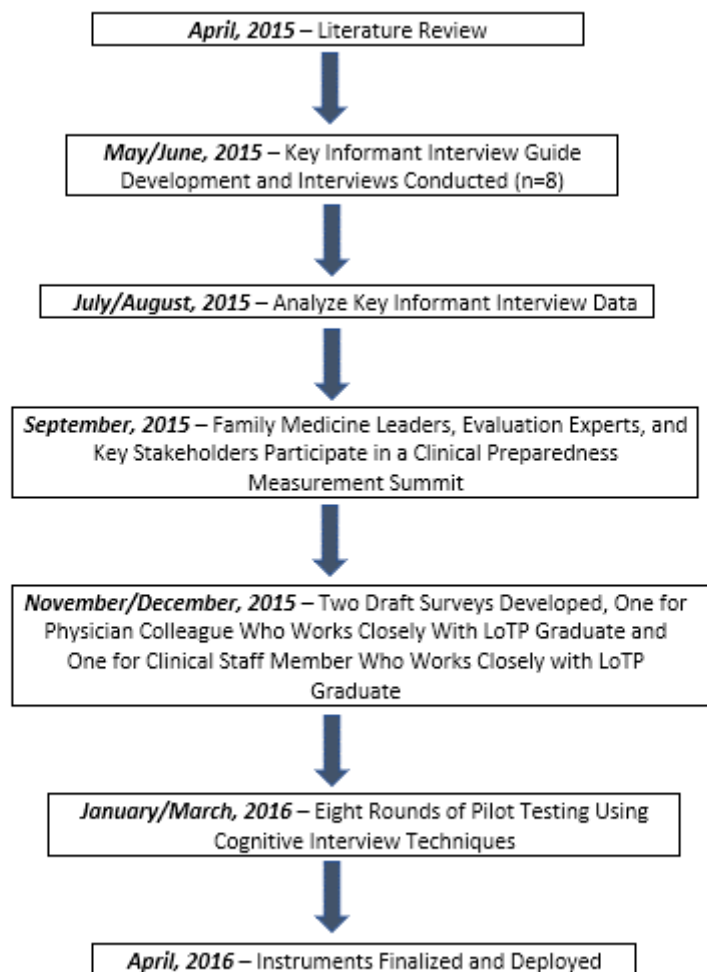
from federally qualified health centers (FQHCs); and two were private practices, one of which was self-owned and the second was not. Table 1 shows the key interview questions along with key findings. In summary, we learned that significant variability exists in the onboarding of new physicians joining a practice. Typically, new hires, including new graduates, have no supervising physician. Direct observation is inconsistent, and indirect observations are more likely to be the primary source of information about preparedness.

While findings from the interviews were helpful, we decided to convene leaders in family medicine, expert evaluators, and other stakeholders to take part in a Clinical Preparedness Measurement Summit, held in Portland, Oregon in September of 2015, to guide decisions related to measuring preparedness for independent clinical practice. The 1-day meeting included representatives from the LoTP Executive Committee, the American Board of Family Medicine (ABFM), the Society of Teachers of Family Medicine (STFM), the American Academy of Family Physicians (AAFP; including a representative from the AAFP Commission on Education), the Association of Family Medicine Residency Directors (AFMRD), the Residency Review Committee (RRC) members, both allopathic and osteopathic physician representatives, and both rural and urban residency training program directors. Seventeen stakeholders and seven members of the LoTP evaluation team from OHSU attended.

The goals of the summit were (a) to engage measurement experts, leaders in family medicine education, and key stakeholders in defining key areas of clinical preparedness for independent practice; and (b) to decide on key measurement and analytic approaches related to specified measurement issues associated with length of training. The summit provided background on the LoTP, current state of the literature on measuring clinical preparedness, results from the key informant interviews, proposed outcome variables and analytic covariates, information on how to account for threshold setting among expert observers (eg, easy scorers vs hard scorers), and pilot testing plans.

We decided to use the then recently published (2015) entrustable professional activities (EPAs) for family medicine end-of-residency training as the framework for assessment of clinical preparedness by physician assessors.²¹ For the instrument designed for staff assessors, we identified a set of nine validated variables in published literature that assessed interpersonal communication skills²² and a set of seven validated variables designed to measure processes of care delivery²³ by staff members working with residents during patient care.

To address the issue of threshold setting, or the tendency to assign ratings of performance that are different from the ratings that the performance warrants, which has been well-documented in several disciplines,²⁴ we developed a series of clinical scenarios for both the physician assessor and the staff assessor to determine how they set their threshold (eg, are they hard or easy raters) when determining readiness for independent clinical practice. We specifically developed and

FIGURE 1. Preparedness for Independent Clinical Practice Instrument: Development Activities and Timeline

then discussed the scenarios during the summit and found during pilot testing that participants produced the desired range in thresholds that would allow us to account for this issue in analyses. The final draft surveys contained 61 items for the physician assessor and 36 items for the staff assessor in the categories described in [Table 2](#).

Finally, we decided on the best timing to administer these two surveys. Recognizing the reality of on-the-job training, we chose to assess the graduates when they were fully oriented to their new positions but before they began to learn on the job. Thus, after an in-depth discussion, we came to consensus to administer the surveys 3 months after the new graduates started their first position as an independent clinician.

Instrument Testing and Refinement

After constructing the final draft surveys, we recruited five clinics in a variety of both rural and urban settings because the LoTP sites were located in both urban and rural settings. Fur-

ther, the pilot sites included a large health system clinic, local health department clinic, residency training continuity clinic, a clinic that does not train residents, and an FQHC clinic. We visited these clinics, typically during lunch, which we provided, and asked volunteer physicians and staff members to complete the surveys about the last physician who joined their practice. We assessed how long these volunteers took to complete the surveys and then used cognitive interviewing techniques²⁵ to assess whether they were responding to the questions as we intended and to determine whether the order of questions influenced their responses. After each session, the surveys were revised and the process was repeated. By the final pilot test session, no further revisions were needed. Both surveys, along with the respective scales used to assess variables, are included in the appendixes (Appendix A, physician assessor; Appendix B, staff member assessor).

After the first year of data capture (2016), we assessed frequencies of all variables to determine the extent that the

TABLE 1. Key Informant Interview Questions and Central Findings for Instrument Development and Administration

Interview questions	Central findings
1 How often do you hire new residency graduates from residency training?	Hiring new graduates is variable, from annually to once in 15 years.
2 When you are recruiting for a family physician, how does a new residency graduate compare with a more established (experienced) physician?	Compared to physicians with lots of clinical experience, new residency graduates tend to have issues with <ul style="list-style-type: none"> • confidence • speed/timing • managing multiple settings simultaneously or back-to-back • referrals (overreferring) • work week unrestricted by duty hour requirements
3 How do supervising physicians determine whether and when a new residency graduate is prepared to work completely independently?	Vastly different models exist for hiring and observing new physicians' entry/establishment in the practice (eg, not all clinics have medical directors), but for the most part, new graduates are treated the same as any new physician hire.
4 How do we determine who the best person is to assess the clinical preparedness of a new resident?	<ul style="list-style-type: none"> • Clinic manager was consistently identified as the person who will know the best person to send our survey to; that is, "Who is likely to best know the patient care skills of Dr X." • Sometimes check-ins occur if routine chart reviews identify an area for improvement.
5 How do you determine how well prepared a new residency graduate is to undertake independent clinical practice?	<ul style="list-style-type: none"> • Takes about half a day to figure out that a resident is not very good; takes about 2 to 3 days to determine that a resident is really good. • Word-of-mouth (rumor mill) is a common source of information on preparedness, but this appears to vary greatly in operational features. Word-of-mouth occurs when something doesn't go well, either in the hospital with a patient care issue or in the clinic. • Both staff and other physicians play a role in word-of-mouth information transfer.

TABLE 2. Clinical Preparedness Survey Components According to Type of Assessor

Number of items	Physician assessor components*	Number of items	Staff assessor components*
3	Demographic characteristics	3	Demographic characteristics
18	Clinical trainin and practice characteristics	7	Clinical training and practice characteristics
5	Threshold setting scenarios	4	Threshold setting scenarios
9	Characteristics of the position the graduate was hired to fill	0	Characteristics of the position the graduate was hired to fill
25	Assessment of clinical preparedness**	21	Assessment of clinical preparedness***
1	Sources of information used to assess preparedness	1	Sources of information used to assess preparedness
61	Total	36	Total

* Designed to take approximately 10 minutes to complete

** Uses 20 FM EPAs and 5 additional preparedness questions

*** Assessments of interpersonal communication (n=9 items), processes of care (n=7 items), and the same 5 additional preparedness questions used on the physician assessor survey

Abbreviations: FM, family medicine; EPA, entrustable professional activities

survey was capturing a range in physician performance and to be sure that the survey could discriminate readiness for independent clinical practice between 3 and 4 years of training. This item analysis revealed that for the majority of items, the full range in the scales were used. Lastly, we used a single global readiness score to assess the correlation between physician and staff assessors' readiness scores for the specific physicians they were evaluating. Because the data were not normally distributed, we used the nonparametric Spearman rank correlation coefficient to conduct this analysis (SPSS version 29 [IBM]). Ultimately, we plan to compare each item as well as summary scores when we publish outcome data that compares 3 versus 4 years of training from the LoTP and will generate separate scores for the physician and staff assessors.

RESULTS

Frequencies from the 2016 physician assessor survey are shown in Table 3. Scores in the highest performance category ("practicing Independently, rarely requests assistance") ranged from a low of 32.8% for managing inpatient care, discharge planning, and transitions of care to a high of 93.8% for providing preventive care that improves wellness, modifies risk factors for illness/injury, and detects illness in early treatable stages. Note that several variables, especially those related to maternity care (48.4% prenatal care, 62.6% manage labor/delivery) and end-of-life care (57.8%) were not done in practice and could not be assessed. The lowest performance category ("not practicing very independently/frequently needing assistance") was used for only one variable, "provide leadership within an interprofessional team," in the assessment of just two

individuals. Scores in the middle category (“practicing mostly independently, sometimes requests assistance”) ranged from 1.6% to 20.3%, indicating a satisfactory range for assessment.

Frequencies from the staff assessor survey are shown in Table 4. The range in scores for interprofessional communication variables for the highest category (“always”) was 32.3% for “apologizing to you for inappropriate behavior” to 95.2% for “shows respect for you as a team member.” The category least used by the assessors was “rarely” for interpersonal communication variables, with “never,” “sometimes,” and “frequently” used more often. The range in scores for the processes of care variables for the highest category (“among the best”) was 41.9% for “handles transfers of care effectively” to 67.7% for “is courteous to coworkers.” None of the staff assessors used the “among the worst category,” while “below average,” “average,” and “above average” were used more often.

Fifty-four out of 64 surveys (84.4%) could be included in the assessment of correlations between the two instruments. The mean summary score for physician assessors was 14.37 (SD=0.98), and the mean for the staff assessors was 18.35 (SD=3.03). The Spearman rank correlation coefficient was 0.107 with a *P* value of .22, indicating no significant correlations existed between the two different assessors (data not shown).

DISCUSSION

We undertook a particularly rigorous process in designing and pilot testing these two surveys to assess clinical preparedness for independent practice. We already had developed a graduate survey, which was administered 1 year after LoTP residents graduated from training. Because we wanted more than one metric measuring clinical preparedness, we developed surveys designed to be completed by both physician and practice staff assessors 3 months after graduates started their first post-training position. We considered this timing optimal because graduates have become familiar with the logistics related to providing care in a new setting but may not yet have started on-the-job clinical learning. In addition, the 3-month window provides physician and staff assessors adequate time to assess the graduate’s preparation. We do not know, however, whether this time interval is valid because a similar tool does not exist. We used a consensus process with key stakeholders to come to agreement on this interval, so it was not chosen arbitrarily, but we do not know whether ratings of clinical preparedness would have been better or worse if a different time interval was chosen. In addition, our correlational assessment revealed no statistical correlation between physician and staff assessor scores. We believe that this occurred because the assessors were evaluating different measures. The physician assessors were evaluating clinical care, which staff assessors do not have the training to do accurately. The staff assessors were assessing processes of care performance, so we did not believe they would necessarily be correlated. As a result, we planned to present outcome data separately for each assessor.

Additional issues related to validity and reliability are important to discuss. The EPAs that form the basis of the physician assessment were developed by leaders in the discipline, thus have face and content validity. These are used to assess resident progression during training, and we are unaware that other validity testing has been done with them. The items on process of care that we included on the staff assessor survey have been validated previously, and we plan to recheck these when we publish our outcome data. Because a study like this has not been done before, we do not have other available instruments to compare our findings to for additional validation. We understand the challenges of rater variation, which is why we added the threshold setting scenarios to each survey. Lastly, because assessing physician performance is dynamic and not a static phenomenon, conducting test-retest approaches would not be a valid approach for testing. We are hoping that these instruments may be used in other studies, which would allow for future comparisons.

Multifocal assessments would be beneficial, but measuring performance is complex. For example, measuring efficiency using time spent on the electronic health record may be affected by patient complexity, as can late chart completion and the number of consultations a physician makes. Assessments of diagnostic errors also can involve a series of events and several individuals and are not free of measurement error. The inclusion of self, patient, and supervisor assessments is subjective; and while knowledge assessments have been well-validated, they may not illustrate the application of knowledge. Nevertheless, as physician training moves closer to being competency-based, discussion of valid, reliable, and effective measurement should and will continue.

Decisions related to length of training in family medicine residency are important. Leaders in the discipline as well as residency directors have strong, and often diverging, opinions.³¹ We intentionally involved many key stakeholders in the Measurement Summit and in the initial key informant interviews. Lastly, the survey pilot testing period was critical to creating a widely accepted and robust measure. A key question often raised is whether a 4-year graduate performs any differently or competently than a 3-year graduate after 1 year in practice. A description of the development and testing of these surveys is important to present in a freestanding paper; so when findings from these surveys are published, this paper may be cited as a reference. To our knowledge, we are the first to develop and rigorously test surveys designed to be completed by clinical practice assessors. Most studies on clinical preparedness survey either the residents themselves or their program directors.^{6–10} We are hopeful that this work will allow us to assess differences in clinical preparedness in their first practice after residency among residents who underwent 3 years of training compared to 4. We purposefully delayed reporting on survey development until data capture on clinical preparedness in 3- and 4-year residency programs was completed because we wanted to avoid any influence publishing the survey might have on physician and staff assessors.

TABLE 3. Frequencies Generated by Physician Assessors for Entrustable Professional Activities for Family Medicine End-of-Residency Training, 2016 (N=64)

EPA #	EPA description	Not done in practice/cannot assess, n (%)	Not practicing very independently, frequently needed assistance, n (%)	Practicing mostly independently, sometimes requests assistance, n (%)	Practicing independently, rarely requests assistance, n (%)
1	Provide a usual source of comprehensive, longitudinal medical care for people of all ages.	5 (7.8)	0	9 (14.1)	50 (78.1)
2	Care for patients and families in multiple settings.	10 (15.7)	0	5 (7.1)	49 (76.6)
3	Provide first-contact access to care for health issues and medical problems.	3 (4.7)	0	4 (6.3)	57 (89.1)
4	Provide preventive care that improves wellness, modifies risk factors for illness and injury, and detects illness in early, treatable stages.	3 (4.7)	0	1 (1.6)	60 (93.8)
5	Provide care that speeds recovery from illness and improves function.	3 (4.7)	0	6 (9.4)	55 (85.9)
6	Evaluate and manage undifferentiated symptoms and complex conditions.	4 (6.2)	0	17 (26.6)	43 (67.2)
7	Diagnose and manage chronic medical conditions and multiple comorbidities.	1 (1.4)	0	9 (14.1)	54 (84.4)
8	Diagnose and manage mental health conditions.	8 (12.5)	0	11 (17.2)	45 (70.3)
9	Diagnose and manage acute illness and injury.	4 (6.2)	0	3 (4.7)	57 (89.1)
10	Perform common procedures in the outpatient or inpatient setting.	11 (17.2)	0	7 (10.9)	46 (71.9)
11a	Manage prenatal care.	31 (48.4)	0	5 (7.8)	28 (43.8)
11b	Manage labor, delivery, and postpartum care.	40 (62.6)	0	5 (7.8)	19 (29.7)
12	Manage end-of-life and palliative care.	37 (57.8)	0	3 (4.7)	34 (37.5)
13	Manage inpatient care, discharge planning, and transitions of care.	27 (42.2)	0	4 (6.3)	33 (32.8)
14	Manage care for patients with medical emergencies.	16 (25.0)	0	11 (17.2)	37 (57.8)
15	Develop trusting relationships and sustained partnerships with patients, families, and communities.	3 (4.7)	0	5 (7.8)	56 (87.5)
16	Use data to optimize the care of individuals, families, and populations.	5 (7.9)	0	5 (7.8)	54 (84.4)
17	In the context of culture and health beliefs of patients and families, use the best science to set mutual health goals and provide services most likely to benefit health.	6 (9.4)	0	3 (4.7)	55 (85.9)
18	Advocate for patients, families, and communities to optimize health care equity and minimize health outcome disparities.	7 (11.0)	0	3 (4.7)	54 (84.4)
19	Provide leadership within interprofessional health care teams.*	11 (17.2)	2 (3.2)	13 (20.3)	38 (59.4)
20	Coordinate care and evaluate specialty consultation as the condition of patients requires.	2 (3.2)	0	8 (12.5)	54 (84.4)

* Two individuals (3.2%) were assessed in the category “not practicing very independently, frequently needed assistance” for this variable.

* Two individuals (3.2%) were assessed in the category “not practicing very independently, frequently needed assistance” for this variable.

Sources: (1) The 7 essential functions of primary care that improve outcomes and access, and lower costs²⁶; (2) the joint principles of the patient-centered medical home²⁷; (3) the benefits of implementing a patient-centered medical home²⁸; (4) the clinical epidemiology definition of the three levels of prevention²⁹; and (5) the role definition of the family physician in the future of family medicine³⁰

Abbreviation: EPA, entrustable professional activities

TABLE 4. Frequencies Generated by Staff Assessors for Interpersonal Communication and Processes of Care, 2016 (N=62)

Interpersonal communication Dr X:	Cannot assess at this time n (%)	Never n (%)	Rarely n (%)	Some-times n (%)	Frequently n (%)	Always n (%)
Listens attentively to you during interactions/ conversations, especially involving patients	0	1 (1.6)	0	0	5 (8.1)	56 (90.3)
Interrupts you during conversations	0	41 (66.1)	16 (25.8)	3 (4.8)	0	0
Is courteous and polite when called or answering page	2 (3.2)	1 (1.6)	0	0	3 (4.8)	56 (90.3)
Shows respect to you as a team member	0	1 (1.6)	0	1 (1.6)	1 (1.6)	59 (95.2)
Gives consideration to your views, suggestions, and opinions	0	1 (1.6)	0	3 (4.8)	9 (14.5)	49 (79.0)
Explains rationale for his/her care plans or actions	3 (4.8)	1 (1.6)	0	1 (1.6)	14 (22.6)	43 (69.4)
Responds to your doubts or answers your questions politely	1 (1.6)	2 (3/0.2)	0	1 (1.6)	4 (6.5)	54 (87.1)
Communicates orders clearly to you either verbally or in writing	0	1 (1.6)	0	3 (4.8)	6 (9.7)	52 (83.9)
Apologizes to you for inappropriate behavior on his/her part	33 (53.2)	3 (4.8)	1 (1.6)	1 (1.6)	4 (6.5)	20 (32.3)
Processes of care Compared with other physicians I know, Dr X:	Cannot assess at this time	Among the worst	Below average	Average	Above average	Among the best
Communicates effectively with patients	2 (3.2)	0	1 (1.6)	4 (6.5)	21 (33.9)	34 (54.8)
Is available for consultation about mutual patients	4 (6.5)	0	1 (1.6)	8 (12.9)	18 (29.0)	31 (50.0)
Is courteous to coworkers	1 (1.6)	0	2 (3.2)	6 (9.7)	11 (17.7)	42 (67.7)
Documents patient care in a timely manner	3 (4.8)	0	1 (1.6)	10 (16.1)	19 (30.6)	29 (46.8)
Communicates effectively with other health care professionals	1 (1.6)	0	1 (1.6)	9 (14.5)	21 (33.9)	30 (48.4)
Handles transfers of care effectively	8 (12.9)	0	2 (3.2)	7 (11.3)	19 (30.6)	26 (41.9)
Contributes to clinic projects or initiatives	7 (11.3)	0	1 (1.6)	9 (14.5)	19 (30.6)	28 (45.2)

The LoTP is, by intention, a pilot study. As such, it is not fully powered to test its hypotheses but rather to explore them with an eye toward identifying effect sizes that may inform larger studies with more rigorous study designs than the case-control design in place for the pilot. The assessment of survey frequencies indicated an appropriate range in responses, such that when generating a summary performance score, will likely be able to discriminate readiness for independent clinical practice according to length of training. We did change one response category in the survey for the variable “apologizing to you for inappropriate behavior,” where we dropped “rarely” because it could indicate either that they rarely apologize or that a need to apologize is rare because no inappropriate behavior is observed. This is the only revision we made to the survey.

The strengths of our approach include the involvement of a diverse set of leaders in family medicine, evaluation experts, and other stakeholders from rural and urban communities as well as from academic and nonacademic clinical settings in the development of our clinical preparedness measure. Rigorous pilot testing resulted in a feasible and understandable survey that generated a satisfactory range of assessor ratings. Getting this initial input will provide confidence in both the interpretation of findings from LoTP and how these findings

can be used to advance our understanding for family medicine residency 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 Measurement Summit held in Portland, Oregon in 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, MS); Michael Peabody, PhD (ABFM Psychometrician); Lars Peterson, MD, PhD (ABFM Senior Physician Scientist); Michael Rabovsky, MD (Stakeholder, Cleveland Clinic); and Russell Thomas, DO (Stakeholder, Eagle Lake, TX).

REFERENCES

1. Blumenthal D, Gokhale M, Campbell EG, Weissman JS. Preparedness for clinical practice: reports of graduating residents at academic health centers. *JAMA*. 2001;286(9):34.
2. Cantor JC, Baker LC, Hughes RG. Preparedness for practice. Young physicians' views of their professional education. *JAMA*. 1993;270(9):40.
3. Batalden P, Leach D, Swing S, Dreyfus H, Dreyfus S. General competencies and accreditation in graduate medical education. *Health Aff (Millwood)*. 2002;21(5):103–111.
4. Nasca TJ, Philibert I, Brigham T, Flynn TC. The next GME accreditation system—rationale and benefits. *N Engl J Med*. 2012;366(11):51–52.
5. Misra S, Iobst WF, Hauer KE, Holmboe ES. The importance of competency-based programmatic assessment in graduate medical education. *J Grad Med Educ*. 2021;13:113–119.
6. Bérubé S, Ayad T, Lavigne F, Lavigne P. Resident's preparedness for independent practice following otorhinolaryngologyhead-and neck surgery residency program: a cross-sectional survey. *Eur Arch Otorhinolaryngol*. 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. Patwardhan VR, Feuerstein JD, Sengupta N. Fellowship colonoscopy training and preparedness for independent gastroenterology practice. *J Clin Gastroenterol*. 2016;50(1):45–51.
9. Dijkstra IS, Pols J, Remmelts P, Brand PL. Preparedness for practice: a systematic cross-specialty evaluation of the alignment between postgraduate medical education and independent practice. *Med Teach*. 2015;37(2):153–161.
10. Jewella K, Newtona K, Dharamsia C, S. Length of family medicine training and readiness for independent practice: residents' perspectives at one Canadian university. *UBCMJ*. 2015;6(2):15–19.
11. 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.
12. 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):26763–26763.
13. Orientale E. Length of training debate in family medicine: idealism versus realism?. *J Grad Med Educ*. 2013;5(2):192–194.
14. 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.
15. *Length of Training Pilot Study*. 2022. <https://fmresearch.ohsu.edu/lotpilot.org>.
16. 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.
17. Eiff MP, Ericson A, Waller E. A comparison of residency applications and match performance according to 3 years versus 4 years of training in family medicine. *Fam Med*. 2019;51(8):641–648.
18. 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.
19. 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.
20. 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.
21. Twenty entrustable professional activities for family medicine. *Association of Family Medicine Residency Directors*. 2015. <https://www.afmrd.org/page/epa>.
22. 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.
23. Lockyer J. Multisource feedback in the assessment of physician competencies. *J Contin Educ Health Prof*. 2003;23(1):4–12.
24. Wind SA, Guo W. Exploring the combined effects of rater misfit and differential rater functioning in performance assessments. *Educ Psychol Meas*. 2019;79(5):962–987.
25. Beatty PC, Willis GB. Research synthesis: the practice of cognitive interviewing. *Public Opin Q*. 2007;71(2):287–311.
26. Starfield B, Shi L, Macinko J. Contribution of primary care to health systems and health. *Milbank Q*. 2005;83(3):457–502.
27. Kellerman R, Kirk L. Principles of the patient-centered medical home. *Am Fam Physician*. 2007;76(6):774–775.
28. Nielsen M, Langner B, Zema C, Hacker T, Grundy P. *Benefits of implementing the primary care medical home: a review of cost & quality results, 2012. Patient-Centered Primary Care Collaborative*. 2012. <https://thepcc.org/resource/benefits-implementing-primary-care-medical-home>.
29. *Picture of America: Prevention. Centers for Disease Control*. 2023. https://www.cdc.gov/pictureofamerica/pdfs/picture_of_america_prevention.pdf.
30. Doohan NC, Harrison DM, Lesko B, Devoe S, E J. The future of family medicine version 2.0: reflections from Pisacano scholars. *J Am Board Fam Med*. 2014;27(1):142–150.
31. Carek PJ. The length of training pilot: does anyone really know what time it takes. *Fam Med*. 2013;45(3):171–172.