

# Reference letters for subspecialty medicine residency positions: Are they valuable for decision-making? Results from a Canadian study.

**Deepti Chopra**

Queen's University Department of Medicine

**Mala Joneja**

Queen's University Department of Medicine

**Gurjit Sandhu**

University of Michigan Departments of Surgery & Learning Health Sciences

**Christopher A. Smith**

Queen's University Department of Medicine

**Catherine M. Spagnuolo**

Queen's University Department of Medicine <https://orcid.org/0000-0001-9953-4923>

**Lawrence Hookey** (✉ [lawrence.hookey@kingstonhsc.ca](mailto:lawrence.hookey@kingstonhsc.ca))

<https://orcid.org/0000-0002-3693-1398>

---

## Research article

**Keywords:** Reference letter, letter of recommendation, residents, medicine subspecialty match

**Posted Date:** October 1st, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-52475/v2>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Version of Record:** A version of this preprint was published on October 7th, 2020. See the published version at <https://doi.org/10.1186/s12909-020-02270-7>.

# Abstract

**Background:** The letter of recommendation is currently an integral part of applicant selection for residency programs. Internal medicine residents will spend much time and expense completing subspecialty away electives to obtain a letter of recommendation. The purpose of this study was 1) to examine a large sample of reference letters in order to define essential components of a high-quality letter, and 2) to elucidate the relationship between quality of reference letter and the letter writer.

**Methods:** We conducted a two-phase study. In phase one, a large sample of letters of recommendation was examined using an audit tool as a coding framework. A 5-point composite endpoint of high-quality letter components was subsequently developed. In phase two, program director letters were compared to non-program director home institution and non-home institution elective letters based on inclusion of components of the 5-point composite endpoint using Chi square testing.

**Results:** 715 letters were examined (398 non-program director home institution letters, 201 program director letters, and 116 non-home institution elective letters). High-quality letter components were: nature of relationship, duration of relationship, In Training Evaluation Report information, research involvement and comments on areas for improvement. Program director letters had a significantly higher proportion (10.4%) of all 5 high-quality components, compared to 0% in both non-program director home institution letters and elective letters ( $p < 0.001$ ). A significantly higher proportion of program director letters had 4-5 high-quality components (62.5%) compared to 2% of non-program director home institution letters and 0% of elective letters ( $p < 0.0001$ ).

**Conclusions:** Letters of recommendation from elective rotations are of the poorest quality and such rotations should not be pursued for the sole purpose of obtaining a letter. The low quality of elective letters leads to the recommendation that writers should decline to write them, programs should not require them and trainees should not request them. Program directors write the highest quality letters and should be a resource for faculty development. Clinical supervisors can use the 5-point composite endpoint as a guide when writing letters for applicants.

## Background

The Letter of Recommendation (LOR) is an integral part of applicant selection for internal medicine subspecialty training programs. LORs are considered one of the most important factors in ranking candidates to postgraduate subspecialty programs and may be helpful in predicting residency retention rates (1-5). However, the writing of LORs has also been described as a process that very few understand (6) and the role of the LOR has been deemed worthy of further examination.

LORs have been found to lack a meaningful comparison of the applicant to peers and often fail to include concrete examples demonstrating applicant performance (7). LORs will often include "lengthy reiterations of already available objective data" and render all applicants "excellent," making it extremely difficult to discriminate between them (8-9). In response to these noted shortcomings, program directors

and medical educators are often asked to comment on what could make LORs more helpful. Researchers have examined, in a limited fashion, the role of the letter writer (10), key elements and phrases in letters and correlations between LORs with objective data (11-12).

Our previous research which included a survey of Canadian internal medicine program directors has provided some insight into the challenges with letter writing and interpretation. Program directors (PDs) felt LORs lacked a common vocabulary and format and that referees often used variable rating scales or similar statements that did not differentiate between candidates. One example of this is the observation that a significant proportion of residents appeared to be in the top 5% of applicants (13). The Canadian internal medicine PDs also suggested that there may be certain reference letter features that are more useful than others, and this led us to explore LORs further (13).

We partnered with the Canadian Resident Matching Service (CaRMS) to evaluate LORs for internal medicine residents applying to subspecialty training programs. The study objectives were to: 1a) elicit the key components of letters which affect quality through a large audit of LORs to subspecialty internal medicine programs in Canada between 2011 and 2014 and b) subsequently define the components of a high-quality letter; and 2) compare PD letters to non-PD home institution and non-home institution elective letters (elective letters) based on the inclusion of high-quality components.

## Methods

Ethics approval was obtained from the Queen's University Health Sciences & Affiliated Teaching Hospitals Research Ethics Board and the Canadian Resident Matching Service (CaRMS) Board.

This study was designed in two phases.

### ***Phase 1: Creation of composite endpoint for a high-quality LOR***

The objective of phase one was to identify common components present in LORs and then by consensus, select the components that are necessary for a LOR to be deemed high-quality.

#### *Common components of LORs*

CaRMS provided 738 de-identified reference letters that had been submitted to subspecialty internal medicine programs (for example, gastroenterology, rheumatology, cardiology etc.) over a 3-year period (2011-2014). Letters had been randomly selected by staff at CaRMS and the distribution across specialties correlated well with the overall distribution for the match from 2011-2014. An audit tool was established based on a literature review of LORs in an effort to collect quantitative and qualitative data, including information and examples according to categories of CanMeds roles and details of letter writer (type of supervisor, relationship to trainee, duration of relationship).

The research assistant conducted content analysis (14-15) using open coding of the reference letters with the audit tool as a coding framework to identify common components. She independently reviewed 10%

of reference letters. The research team subsequently reviewed the same data set. Discrepancies among the team were resolved and the audit tool was refined for clarity. The resultant iteration of the audit tool provided a clear list of components for ongoing coding (*Appendix A*). The research assistant was not associated with any resident or faculty investigators to address reflexivity and mitigate bias (16).

### *Defining a high-quality letter: construction of the composite endpoint*

A composite endpoint consisting of five items was created to define a high-quality LOR. The composite endpoint was created by selection of the five most valuable components of LORs from the list of identified common components, based on: 1) the opinions of Canadian internal medicine program directors obtained through a survey (13) (Figure 1), 2) results of a detailed literature review, and 3) consensus from a panel of local medical educators, including subspecialty medicine program directors and the dean of postgraduate education.

### *Phase 2: Reference letter comparison based on referee type*

Letters from PDs (PD letters) were compared to (i) letters by non-PDs from the home school (non-PD home institution letters), and (ii) letters obtained on elective rotations (elective letters). Letters were compared with respect to the inclusion of individual components of the 5-point composite endpoint. The inclusion of In Training Evaluation Report (ITER) information was identified as one of the components of the composite endpoint. Upon reflection that elective letters were unlikely to include this data, a separate comparison was also performed, removing the ITER component and leaving a 4-point composite endpoint. Analysis was also performed to assess if a writer had met the criteria for a good to high-quality letter by including a majority of the components.

Chi square testing was performed to compare frequencies using SPSS Statistics, version 24 (IBM, Armonk, New York). A sample size of 105 per group was necessary to demonstrate a 20% difference in proportion of high-quality letters using three-step pairwise comparison in an ANOVA.

## **Results**

### *Phase 1: Creation of composite endpoint for a high-quality LOR*

The components necessary for a high-quality letter were determined to be:

1. nature of relationship between the referee and candidate, such as rotation supervisor, research supervisor, and/or program director
2. duration of this relationship-measured in weeks.
3. inclusion of In Training Evaluation Report (ITER) information
4. research involvement comments, often outlining role in specific projects
5. area for improvement or critique comments.

Given the possibility that only PDs would have ready access to ITER information, we also performed the analysis excluding this component, using a 4-point composite endpoint.

**Phase 2: Reference letter comparison based on referee type**

738 letters were reviewed, with 715 included in the final analysis. Twenty-three were excluded due to duplicates or redaction beyond ability to extract meaningful information. The distribution of reference letters was: 398 non-PD home institution letters (55.7%), 201 PD letters (28.1%), and 116 elective letters (16.2%). Letters were written for 206 applicants, with a range of 1-5 letters per applicant reviewed. Applicants were applying for 9 specific sub-specialties, with good representation of the breath of sub-specialty internal medicine. There were 113 letters for which the specialty being applied for was not evident. There were no letters of application for allergy, infectious disease, or endocrinology. Letter writers were affiliated with 33 different institutions within Canada and the United States.

With respect to the primary outcome (inclusion of all 5 high-quality components of the 5-point composite endpoint), PD letters had a significantly higher proportion of letters that met this target (10.4%), compared to 0% in both non-PD home institution letters and elective letters ( $p < 0.001$ ) (Table 1a). Using the “non ITER” 4-point composite endpoint, PDs again had a significantly higher proportion with all four components (12.4%), compared to non-PD home institution letters (1.3%) and elective letters (0%) (Table 1b).

Table 1a. Comparison of composite score based on the 5-point composite endpoint (ITER included) across letter writers

		Letter writer			P value
		Non-program director home institution letters, n=398	Program director letters, n=201	elective letters, n=116	
Composite score	0	9	0	6	<0.0001
	1	67	5	50	
	2	172	17	47	
	3	142	54	13	
	4	8	104	0	
	5	0	21	0	

Findings are significant for subgroup analysis of home versus elective as well ( $p < 0.0001$ )

Table 1b. Comparison of composite score based on the 4-point composite endpoint (ITER excluded) across letter writers

		Letter writer			P value
		Non-program director home institution letters, n=398	Program director letters, n=201	elective letters, n=116	
Composite score	0	9	0	6	<0.0001
	1	68	5	51	
	2	174	49	47	
	3	142	122	12	
	4	5	25	0	

Findings are significant for subgroup analysis of home versus elective as well ( $p < 0.0001$ ) Table 2 shows the prevalence of each component of the 5-point composite endpoint across letter writers.

When evaluated for having a majority of the components of the 5-point composite endpoint, a significantly higher proportion of PD letters had 4-5 components (62.5%) compared to 2% of non-PD home institution letters, and 0% of elective letters ( $p < 0.0001$ ) (Table 1a). Of PD letters, 94% contained 3 or more components, compared to 37.7% of non-PD home institution letters, and 11.2% of elective letters ( $p < 0.0001$ ) (Table 1a). Likewise, for the 4-point composite endpoint, 73.1% of PD letters had 3-4 components compared to 36.9% of non-PD home institution letters and 10.3% of elective letters ( $p < 0.0001$ ) (Table 1b).

Table 2. Prevalence of each component of the 5-point composite endpoint across letter writers

		Letter writer			P value
		Non-program director home institution (n=398)	Program director (n=201)	Elective supervisor (n=116)	
Any mention of	Nature of relationship	379	194	104	0.025
	Duration of relationship	323	181	82	<0.0001
	ITER comments	7	153	2	<0.0001
	Research involvement	220	139	61	0.002
	Area of improvement/critique	23	36	3	<0.0001

## Discussion

Letters of recommendation continue to be a key component of the modern fellowship application (6). However, this current study of over 700 LORs reveals a significant variability in letter quality. Variability in letter quality appeared dependent on the relationship between the writer and applicant. Not surprisingly, PDs' letter writing practice is consistent with their recommendations. This is demonstrated by the fact that PD letters had the highest proportion of high-quality letters (10.4% included all 5 components of the 5-point composite endpoint and 94% had 3 or more components), while letters from elective supervisors were of much lower quality, with 0% having all 5 components of the 5-point composite endpoint, and only 11.2% having 3 or more components.

Although there is existing literature examining LORs, the current study adds to this in several respects. Most studies to date have looked at trainees in the United States medical training system (10, 11, 17), while the current study examined letters from across Canada, with a different healthcare and education system, thus adding the experience of another country and enriching the data pool. This study also highlights the relationship between letter quality and *letter writer*, leading to opportunity for possible faculty development with a goal of improving letter quality. Lastly, it brings forth the issue of elective rotations, and their role as auditions for position, and whether letters of recommendation should be a part of key outcomes of this effort, particularly in light of the poor quality of the letters that come from the experience.

### ***The letter writer***

Program directors practice what they preach. They understand the value of specific components of LORs and include these components with the highest frequency. This indicates that PDs are an untapped resource for providing faculty development to all clinical supervisors who are potential letter writers. In stark contrast to PD letters, LORs from elective rotation supervisors were of the lowest quality. One possible reason for this finding is that a short amount of time spent with an unknown trainee does not provide sufficient data to allow the supervisor to be comfortable in rating the trainee in key domains. As many residents will seek out elective rotations at other sites with the intention of not only introducing themselves, but also to obtain a LOR from this site, the overall poor quality of these letters calls into question at least one element of the perceived value of these rotations.

### ***Elective rotations***

While letters from elective supervisors represent approximately 15% of overall letters in our sample, their utility appears very limited; a meagre 11.2% met the criteria for a good to high-quality letter by including 3 or more components of the 5-point composite endpoint. Questions raised by this include 1) why would trainees get these letters, and 2) did this study miss some important aspect of the letter that would make letters from electives important? While our data suggest elective letters are not of good quality, and thus seemingly not useful overall, they may be a necessity for a particular institution, regardless of how non-specific the information within them, as suggested by recent papers from the surgery field (18-19). These could be considered letters from an 'audition elective'- a term used to describe a clinical elective taken by trainees to distinguish themselves from their peers in order to improve their chances of being selected by a residency program (20). Internal medicine residents will spend a great amount of time and expense completing sub-specialty audition electives to improve their chances of success in the subspecialty match, possibly via obtaining a strong reference letter from their elective supervisors. As we have shown the LORs from elective rotations to be of low quality, this calls into question the value of these rotations, given their associated costs. With the current travel restrictions related to the COVID-19 pandemic, completion of these away rotations may be even more difficult. This is consistent with other studies questioning the value of the 'away rotation' and that the value of such rotations should not depend solely on the obtaining of a LOR (21).

The current study does have certain limitations. While exposing key differences in letter quality according to the letter writer, further exploration of this data is somewhat limited. To ensure anonymity, the letters underwent redaction prior to release by CaRMS, with demographic data including age, gender, and candidate home institution name being removed. Moreover, we were unable to correlate these findings with subspecialty match or residency performance results, thus limiting us to examining the process rather than outcomes that may be of interest to a PD. We put a lot of weight on the opinions of internal medicine PDs, largely because of their experience in both writing and reading letters of recommendation. However, there are other members of selection committees, whose views and opinions of a quality letter may differ. In consideration of this, our composite endpoint was developed with input from the internal medicine PD survey, but also with input from a panel of academic internal medicine subspecialists who sit on various residency training selection committees. Lastly, the current study only looked at letters for application to the subspecialty internal medicine match and results may not be transferable to programs accepting medical students into their first year of residency training (PGY-1).

## **Recommendations**

1. LORs from elective rotations: We recommend that letters from elective rotations should not be mandatory for consideration of acceptance into a program, as well as consideration of a faculty wide approach/position that writers will not provide letters based solely on elective rotations. This recommendation would presumably lead to trainees asking for letters from other sources, likely faculty that know them better and thus can provide a higher quality letter.
2. Faculty development: Program directors are a potential resource for faculty development and can provide guidance to letter writers based on their experience and expertise in postgraduate medical education. Institutions should approach program directors for faculty development strategies that will lead to improved quality of LORs.
3. Letter writers: Clinical supervisors who are writing letters for applicants in the Medicine Subspecialty Match can use the components of the 5-point composite endpoint as guide for the inclusion of important components (relationship between the referee and candidate, duration of relationship, inclusion of In Training Evaluation Report (ITER) information, research involvement and comments on areas for improvement). Also, if a supervisor does not feel they can comment on one or more of these components, then it would be acceptable to decline a request for a LOR from an applicant.
4. Letter readers: Similarly, those reviewing applicants' portfolios can use the components of the 5-point composite endpoint to judge the quality of the letter.
5. Applicants: Applicants should be aware of what program directors value in LORs and should direct their requests to supervisors who are appropriate and capable of commenting on the key components.

## **Conclusions**

In conclusion, LORs continue to be variable in quality. There continues to be a quality gap in LORs for subspecialty medicine positions, in particular with respect to letters from elective rotations. Program directors write the highest quality LORs and should be a resource for faculty development. LORs from away rotations are of the poorest quality and such rotations should not be pursued for the sole purpose of obtaining a letter. Future studies are needed to assess outcome and to study clinical performance of trainees who have letters from elective rotations. Future research could involve surveying trainees and program directors for deeper insight into what drives LORs and opportunities for improving them.

## List Of Abbreviations

LOR - letter of recommendation

PD - program director

CaRMS - Canadian Resident Matching Service

ITER - In Training Evaluation Report

MSM – Medicine Subspecialty Match

## Declarations

### *Ethics approval and consent to participate*

Ethics approval was obtained from the Queen's University Health Sciences & Affiliated Teaching Hospitals Research Ethics Board and the Canadian Resident Matching Service (CaRMS) Board. Data from letters of recommendation were redacted so that there were no identifiers.

### *Consent for publication*

Not applicable.

### *Availability of data and materials*

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

### *Competing interests*

The authors declare that they have no competing interests.

### *Funding*

Funding sources include the Thomas M. and Louise A. Brown Research Studentship (Dr. Chopra) and a Southeastern Ontario Academic Medical Organization Clinician Scientist Award (Dr. Hookey).

## ***Authors' contributions***

DC, MJ, GS, CAS, LH were responsible for study design, data analysis, manuscript preparation and revision. CMS was responsible for data analysis, manuscript preparation and revision.

## ***Acknowledgements***

Logistical support was provided by CaRMS.

Prior Presentations: Deepti Chopra presented a portion of this work at the International Conference on Residency Education, October 2017.

## **References**

1. Kudlas MJ. Effects of radiography program admissions practices on student retention. *J Allied Health*. 2006;35(3):162-8.
2. Dobovsky SL, Gendel MH, Dubovsky AN, Levin R, Rosse J, House R. Can admissions interviews predict performance in residency? *Acad Psychiatry*. 2008;32(6):498-503.
3. Siu E, Reiter HI. Overview: what's worked and what hasn't as a guide towards predictive admissions tool development. *Adv Health Sci Educ Theory Pract*. 2009;14(5):759-75.
4. Khan K, Levstik M. Ranking in Canadian gastroenterology residency match: what do residents and program directors want? *Can J Gastroenterol*. 2010;24(6):369-72
5. Nallasamy S, Uhler T, Nallasamy N, Tapino PJ, Volpe NJ. Ophthalmology resident selection: current trends in selection criteria and improving the process. *Ophthalmology*. 2010;117(5):1041-7.
6. Nehler M. Letters of recommendation: how do they fit into the modern application? *J Grad Med Educ*. 2018;10(3):267-8.
7. Fortune J. The content and value of letters of recommendation in the resident candidate evaluation process. *J. Curr Surg*. 2002;59(1):79-83.\
8. McCarthy J, Goffin, R. Improving the validity of letters of recommendation: an investigation of three standardized reference forms. *Mil Psychol*. 2001;13(4):199-222.
9. Prager JD, Myer CM 3rd, Pensak ML. Improving the letter of recommendation. *Otolaryngol Head Neck Surg*. 2010;143(3):327-30.
10. Beskind DL, Hiller KM, Stolz U, Bradshaw H, Berkman M, Stoneking LR, et al. Does the experience of the writer affect the evaluative components on the standardized letter of recommendation in emergency medicine? *J Emerg Med*. 2014;46(4):544-50.
11. Saudek K, Saudek D, Treat R, Bartz P, Weigert R, Weisgerber M. Dear program director: deciphering letters of recommendation. *J Grad Med Educ*. 2018;10(3):261-6.
12. Rajesh A, Rivera M, Asaad M, Chandra A, Baloul MS, Backstrom CM, et al. What are we really looking for in a letter of recommendation? *J Surg Educ*. 2019;76(6):e118-24.

13. Chopra D, Sandhu G, Smith C, Hookey L. Reference letters for subspecialty medical residents: pearls or perils? Abstract presented at: Canadian Conference on Medical Education; 2013 Apr 20-23; Quebec City, QC.
14. Mayring P. Qualitative content analysis [28 paragraphs]. Forum Qualitative Sozialforschung/Forum: Qualitative Social Research [Internet]. 2000 June [cited 2020 June 24];1(2). Available from: <http://www.qualitative-research.net/index.php/fqs/article/view/1089/2385>
15. Cho JY, Lee E. Reducing confusion about grounded theory and qualitative content analysis: similarities and The Qualitative Report. 2014;19(32):1-20.
16. Berger R. Now I see it, now I don't: researcher's position and reflexivity in qualitative research. Qualitative research. 2015;15(2):219-34.
17. Garmel GM, Grover CA, Quinn A, Lotfipour S, Langdorf MI, Cheaito MA, et al. Letters of recommendation. J Emerg Med. 2019;57(3):405-10.
18. Krauss EM, Bezuhly M, Williams JG. Selecting the best and brightest: a comparison of residency match processes in the United States and Canada. Plast Surg. 2015;23(4):225-30.
19. Weissbart SJ, Stock JA, Wein AJ. Program directors' criteria for selection into urology residency. Urology. 2015;85(4):731-6.
20. Higgins E, Newman L, Halligan K, Miller M, Schwab S, Kosowicz L. Do audition electives impact match success? Med Educ Online [Internet]. 2016 June [cited 2020 June 28]; 21(1):31325. Available from: <https://doi.org/10.3402/meo.v21.31325> DOI: 10.3402/meo.v21.31325
21. Villwock JA, Hamill CS, Ryan JT, Nicholas BD. The role of the away rotation in otolaryngology residency. Otolaryngol Head Neck Surg. 2017;156(6):1104-7.
22. Alweis R, Collichio F, Milne CK, Dalal B, Williams CM, Sulistio MS, et al. Guidelines for a standardized fellowship letter of recommendation. Am J Med. 2017;130(5):606-11.
23. Roberts C, Khanna P, Rigby L, Bartle E, Llewellyn A, Gustavs J, et al. Utility of selection methods for specialist medical training: a BEME (best evidence medical education) systematic review: BEME guide no. 45. Med Teach. 2018;40(1):3-19.

## Figures

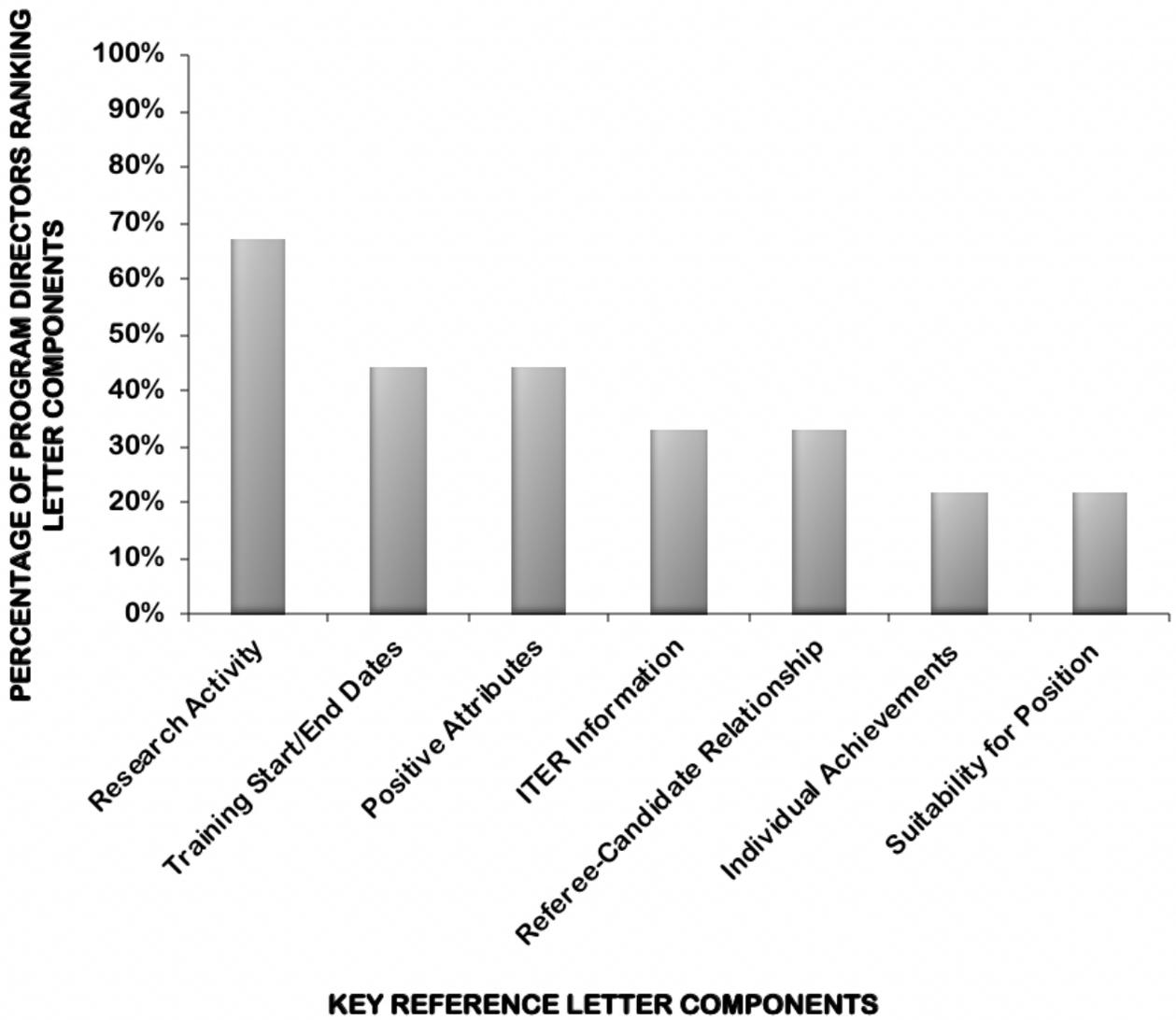


Figure 1

Program director opinions of essential letter of recommendation components

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Appendix.pdf](#)