Nomogram to Predict Successful Placement in Surgical Subspecialty Fellowships Using Applicant Characteristics

Tyler M. Muffly, MD, Matthew D. Barber, MD,* Matthew T. Karafa, PhD,† Michael W. Kattan, PhD,† Abigail Shniter, BS,‡ and J. Eric Jelovsek, MD**

*Center for Urogynecology and Pelvic Reconstructive Surgery, Obstetrics, Gynecology and Women's Health Institute, Cleveland Clinic, Cleveland, Ohio; †Department of Quantitative Health Sciencies, Cleveland Clinic, Cleveland, Ohio; ‡Department of Obstetrics and Gynecology, University of Missouri Kansas City, Kansas City, Missouri

PURPOSE: The purpose of the study was to develop a model that predicts an individual applicant's probability of successful placement into a surgical subspecialty fellowship program.

METHODS: Candidates who applied to surgical fellowships during a 3-year period were identified in a set of databases that included the electronic application materials.

RESULTS: Of the 1281 applicants who were available for analysis, 951 applicants (74%) successfully placed into a colon and rectal surgery, thoracic surgery, vascular surgery, or pediatric surgery fellowship. The optimal final prediction model, which was based on a logistic regression, included 14 variables. This model, with a c statistic of 0.74, allowed for the determination of a useful estimate of the probability of placement for an individual candidate.

CONCLUSIONS: Of the factors that are available at the time of fellowship application, 14 were used to predict accurately the proportion of applicants who will successfully gain a fellowship position. (J Surg 69:364-370. © 2012 Association of Program Directors in Surgery. Published by Elsevier Inc. All rights reserved.)

KEY WORDS: fellowship, personnel selection, nomogram

COMPETENCIES: Interpersonal and Communication Skills, Practice Based Learning and Improvement

INTRODUCTION

From 1993 to 2010 the number of applicants participating in the National Residency Matching Program Specialty Matching Service almost doubled.[1] In 2009, approximately 910 residents applied to a United States surgical subspecialty fellowship program and approximately 86% of applicants matched into 1 of these programs. Studies point toward an impending shortage of subspecialty physicians, as well as primary care physicians, where demand will outpace physician supply through at least 2025.[2,3] Because this increase in subspecialty physicians will likely continue due to population growth, aging, and other factors, it is important to consider tools that may assist trainees in selecting subspecialty areas of interest.

The career decision of medical students and residents to subspecialize involves a host of demographic factors, curriculum factors, debt level, and institutional factors.[4] Because of the finite population of residents who pursue subspecialty training, proportionally few studies examine factors that cause residents to consider applying to fellowship. Many of the factors that go into choosing a subspecialty fellowship are similar to factors that medical students use when making their decision about residency with a focus on lifestyle.[5] However, applicants to fellowship programs must consider a number of additional challenges during the application process when the decision to subspecialize is made: cost of travel for interviewing, cost of income lost due to continued postgraduate training, time lost from ongoing residency rotations, and possible need to use vacation time for interviews. These logistical and financial barriers to entering a fellowship can create doubts about submitting an application; especially, if the probability of successfully matching is not known. An individualized risk quantification would be helpful for applicant decision-making and counseling residents who are considering choosing a subspecialty fellowship. Moreover, such quantification may be useful for fellowship program directors who desire an objective method of enumerating an applicant's likelihood of matching before interviewing.

The study's objective was to create a prediction model based on past applicants' characteristics and assess its accuracy in pre-

Correspondence: Inquiries to Tyler M. Muffly, MD, Center for Urogynecology and Pelvic Reconstructive Surgery, Department of Obstetrics and Gynecology, Desk A-81, Obstetrics, Gynecology, and Women's Health Institute, Cleveland Clinic, 9500 Euclid Avenue, Cleveland, OH 44195; fax: 216-636-2620; e-mail: mufflyt@ccf.org

  1931-7204/$30.00 doi:10.1016/j.jsurg.2011.11.006

dicting their successful placement within a surgical subspecialty fellowship training program.

METHODS

Residency application data were available from 2007 to 2009 from the American Association of Medical Colleges (AAMC) using its Electronic Residency Application Service (ERAS) database. Residency and fellowship participation data were also available from the AAMC using its Graduate Medical Education (GME) Track database. GME Track is a residency and fellowship database and tracking system that assists GME administrators and program directors in the collection of GME data.

Demographic data, educational history, and successful placement were abstracted for those applicants who applied to a surgical fellowship approved by the Accreditation Council for Graduate Medical Education (ACGME). Fellowship subspecialties that did not use the Electronic Residency Application Service and were not approved by the Accreditation Council for Graduate Medical Education were not considered because they are not surveyed by GME Track. Using these constraints four surgical fellowships were selected: colon and rectal surgery, pediatric surgery, vascular surgery, and thoracic surgery. Applicants with incomplete data were excluded from the sample, and data were deidentified before the authors receiving the dataset from the AAMC. Approval for the study was obtained from the Cleveland Clinic Institutional Review Board and the AAMC.

A statistical model was created to take into account the factors which predict the numerical probability of successful placement for a given set of applicant characteristics. Demographic and educational data were categorized as predictors while successful placement (dichotomous yes/no) was categorized as the outcome. Predictors were selected based on a search of the published literature and plausibility. United States Medical Licensing (USMLE) scores and National Residency Matching Program rank lists were not provided by the AAMC because of confidentiality concerns thus were not included in the model. A priori we identified 14 predictors of interest and determined how they would be coded in the model because making these decisions afterward can have deleterious effects on the predictive ability of the model. We were unable to determine whether volunteer and research experiences were mutually exclusive; thus they were combined as 1 predictor variable and reported as frequency. Other continuously coded variables included age, number of research publications and the prior year's number of applicants. Because the number of applicants for an upcoming match year is not known until the match is completed, we accounted for how competitive a specialty may be from year to year using the prior year's number of applicants reported by National Residency Matching Program (NRMP) as a variable in the model to represent the number of applicants competing for positions. Categorically coded variables included gender, subspecialty, citizenship, Alpha Omega Alpha honor society status, whether the applicant was named in a malpractice suit, board certification, chief resident status, medical school type, military service, and other service obligations. Other service obligations included participation with the National Health Service, Corps and other such programs.

We created a multivariable logistic regression model using the methods proposed by Harrell et al.[6] Linearity assumptions of continuous variables were relaxed using 3-knot restricted cubic splines to allow a monotonic increase or decrease in the value of the factor. Internal validation of the predictive accuracy of the obtained model consisted of 2 components: discrimination and calibration. First, we quantified discrimination with the c-statistic, or the area under the receiver operating characteristic (ROC) curve, which ranges from 0 to 1. In our study, a c-statistic of 0.5 would indicate that the model has a prediction that is no different from chance. Cross validation was done by a split sample technique. The data were split into groups of 1 tenth of the number of observations. One group was removed, and the model was built on the reduced sample set, which is considered fixed. This fixed model derived from the reduced data were used to predict the group of applicants that was left out. Repeating this process, by leaving out each group once, provided predictions for all applicants in the original cohort and hence a model performance index (c statistic). To protect

TABLE 1. Descriptive Characteristics of the Study Population

Variable	Value
Female	286 (22)
Age (years)	
Younger than 30	32 (2.4)
30–32	476 (37)
33–34	360 (28)
35–37	259 (20)
38–39	52 (4)
40 and older	102 (9)
Number of applicants to specialty	
Vascular surgery	428 (34)
Colon and rectal surgery	349 (27)
Thoracic surgery	311 (24)
Pediatric general surgery	193 (15)
US citizenship	1,071 (84)
Alpha Omega Alpha member	188 (15)
Volunteer and research experiences	3.6 ± 3.8
Named in a malpractice suit	105 (8.2)
Board certified	106 (8.2)
Chief resident	980 (77)
Number of publications	10.9 ± 13.8
Medical school type	
US public allopathic	476 (37)
US private allopathic	405 (32)
International	317 (25)
Osteopathic	44 (3.4)
Canadian	39 (3.0)
Military service commitment	31 (2.4)
Other service commitments	17 (1.3)
Match status	
Successfully matched	951 (74)
Unsuccessfully matched	330 (26)

Values are n (%) or mean ± SD.

against the influence of the random splits, cross validation was repeated a large number of times (in our case 1000 resamples) and the average of the 1000 indexes was the bias-corrected index. Second, calibration was assessed. This was done by grouping applicants into quintiles with respect to their model-predicted probabilities and then comparing the mean probability of successful placement for each quintile with the observed quintile-specific estimate of matching success. Again, bootstrap resamples were used to correct for optimism. All analyses and nomogram construction were performed using R software, Version 2.12.0 (http://www.r-project.org).

RESULTS

During the three-year matching study period, 1281 applicants met inclusion criteria. The demographic and educational factors characteristics of the study population are presented in Table 1. Successful placement occurred in 951 of these applicants (74%). In our study, the intention was to predict a binary outcome of a fellowship placement model (success/failure) using the above-mentioned variables. Fourteen factors: gender, age, surgical subspecialty, citizenship status, Alpha Omega Alpha status, number of research and volunteer experiences, being named in at least 1 malpractice suit, board certification, chief resident status, number of publications, medical school type attended, the total number of applications submitted in the prior year, military service, and other service obligations, were found to contribute to an optimal final prediction model.

The graphic nomogram derived from the regression model is presented in Figure 1 or by visiting the web site located at (http://rcc.simpal.com/RCEval.cgi?RCID=WCBAPN). This nomogram is used by locating each applicant characteristic and finding the number of points, on the uppermost point scale, to which that characteristic corresponds (Fig. 2). When the numbers of points generated from all of an applicant's characteristics are added together, that sum is found on the "Total Points" scale, and the predicted probability of successful placement is the probability in the lowermost scale that is obtained by drawing a vertical line from the "Total Points" to the "Predicted Value" scale. The results of the multivariable logistic regression, presented with odds ratios and corresponding 95% confidence intervals, are presented in Table 2. The corresponding receiver operating characteristics curve is presented in Figure 3. Lastly, examples of the nomogram's predictive capability are illustrated by the calculation of fellowship success for 2 hypothetical applicants. Table 3 presents the characteristics of these applicants,

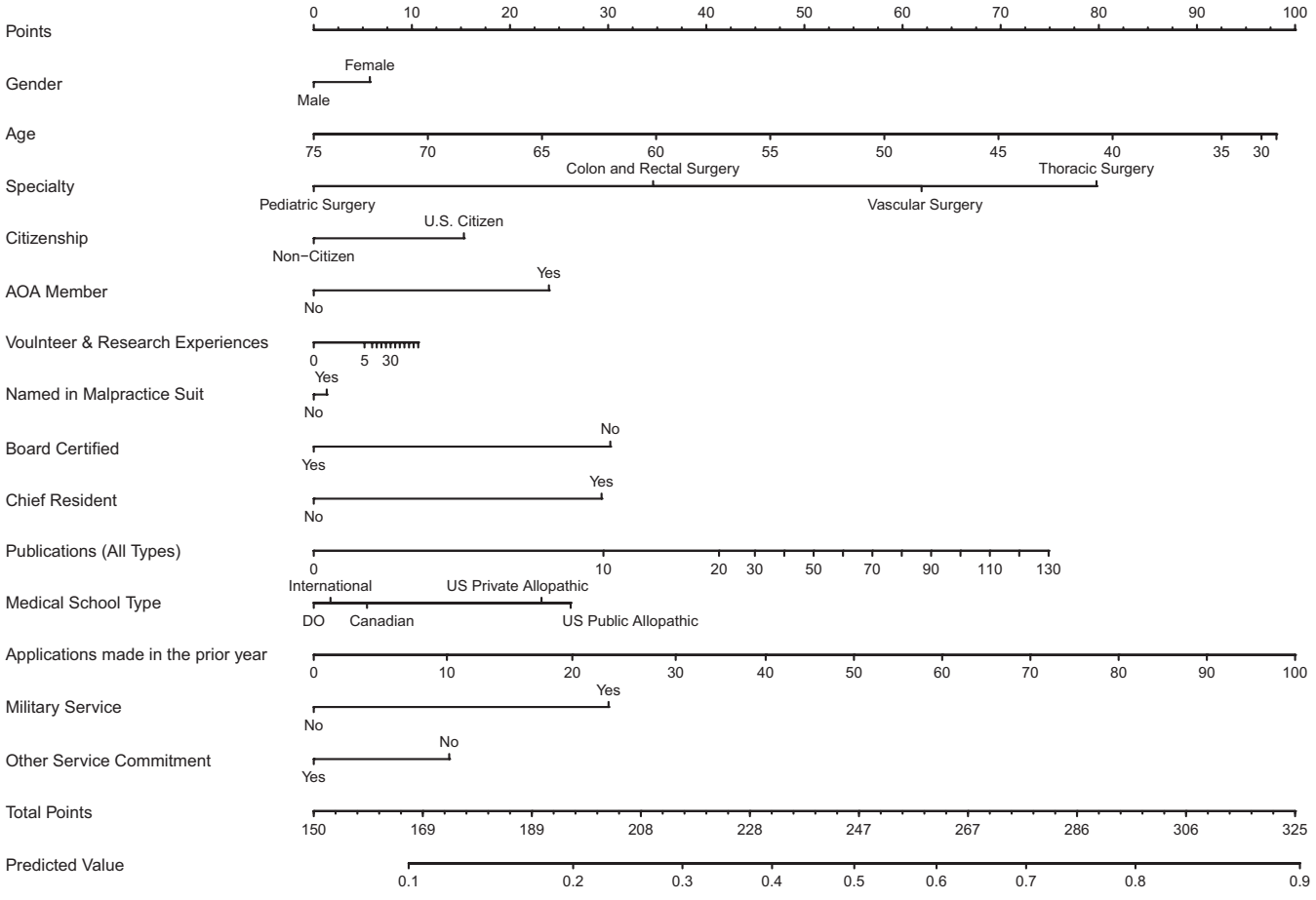

FIGURE 1. Predictive graphic nomogram for probability of successful placement into a surgical subspecialty fellowship.

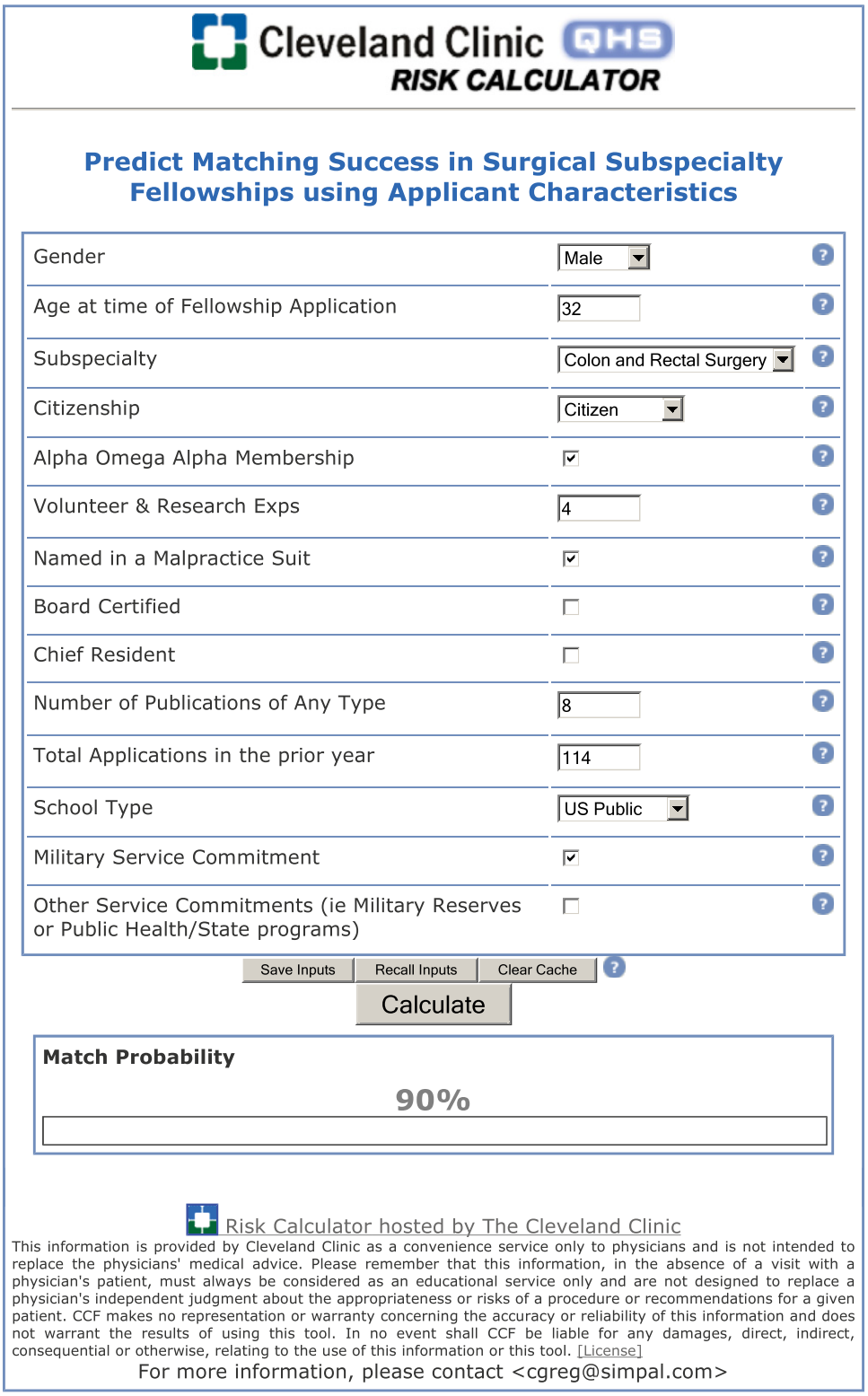

FIGURE 2. Screenshot of nomogram web site.

the total number of points generated by these characteristics from the nomogram, and the final predicted chance of successful placement for each individual applicant. With a c-statistic of 0.74 (95% confidence interval, 0.71-0.77), the nomogram explains 48% of the variation in those who successfully gain placement with a chosen fellowship subspecialty program.

The cross-validation procedure showed that the performance of the nomogram on the test set was similar to that originally

TABLE 2. Factors Associated With Successfully Matching into a Surgical Subspecialty Fellowship by Multivariable Logistic Regression

Variable	Odds Ratio	95% Confidence Interval	P Value
Female	1.149	0.80–1.639	0.45
Age	0.96	0.81–1.14	0.66
Subspecialty			
Colon and Rectal Surgery	Referent	—	—
Pediatric Surgery	0.03	0.0–3.37	0.14
Thoracic Surgery	2.13	0.95–4.75	0.07
Vascular Surgery	2.92	1.38–6.15	0.01
Non-US Citizen	0.66	0.43–1.02	0.06
AOA Member	1.68	1.03–2.75	0.04
Volunteer and Research Experiences	0.945	0.80–1.12	0.51
Named in a Malpractice Suit	1.1	0.66–1.83	0.73
Board Certified	0.43	0.26–0.70	0.001
Chief Resident	2.25	1.63–3.11	<0.001
Number of publications	0.879	0.81–0.96	0.004
Medical School Type			
US Private	Referent	—	—
US Public	1.06	0.74–1.53	0.75
International	0.65	0.41–1.03	0.07
Canadian	0.48	0.2–1.16	0.1
Osteopathic	0.52	0.25–1.07	0.07
Total Application Prior Year	1.05	0.96–1.14	0.3
Military Service	2.28	0.73–7.06	0.16
Other Service Commitment	0.68	0.21–2.25	0.53

AOA, Alpha Omega Alpha Medical Honor Society.

TABLE 3. Predicted Rates of Subspecialty Fellowship Matching for 2 Hypothetical Applicants

Characteristic	Applicant 1	Applicant 2
Sex	Male	Female
Age (years)	32	35
Applying to Subspecialty	Colon and Rectal Surgery	Colon and Rectal Surgery
Citizenship	US Citizen	Non-US Citizen
AOA Membership	AOA Member	Non-AOA Member
Volunteer and Research Experiences	4	8
Named in a Malpractice Suit	Yes	No
Board Certified	No	Yes
Chief Resident	No	Yes
Number of Publications (All Types)	8	8
Medical School Type	US Public	International
Number of Applications in the Prior Year	114	114
Military Service Commitment	Yes	No
Other Service Commitments	No	No
Total Points	240	194
Predicted Probability of Matching (%)	90%	48%

AOA, Alpha Omega Alpha Medical Honor Society.

determined from the training set. Figure 4 compares the predicted rates of successful placement with the empiric probabilities of success for applicants in the test set.

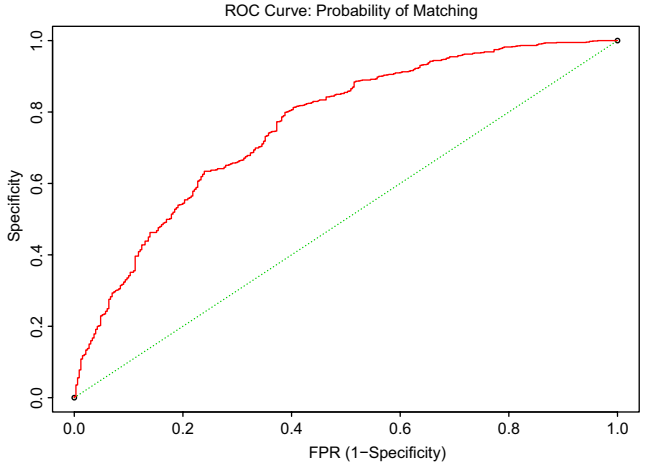

FIGURE 3. Receiver operating characteristics curve regarding probability of fellowship placement.

DISCUSSION

We have created a nomogram that allows an applicant or their mentor in a surgical subspecialty to predict the probability of

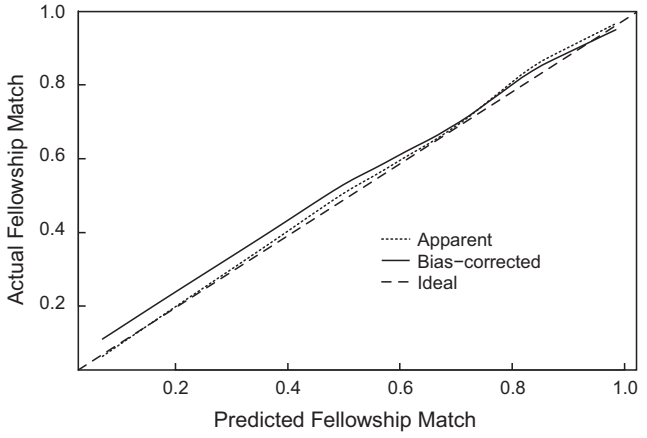

FIGURE 4. Nomogram calibration. The reference line, which an ideal nomogram would produce, is represented by the dashed diagonal line; the actual bias-corrected performance of the current nomogram is represented by the solid line.

successfully gaining placement into that subspecialty using basic application variables. A literature search (MEDLINE; January 1966-February 2011; English language; search terms: "personnel selection," "Internship and Residency," "nomogram," and "School Admission Criteria") shows there is no literature on tools to accurately predict the chance of a successful fellowship placement.

The uses of the surgical subspecialty nomogram are multiple. The nomogram is useful for comparatively visualizing the strength of associations between an applicant's demographic and educational variables and fellowship placement success. The nomogram illustrates that the competition is vastly different between the 4 specialties of colon and rectal surgery, thoracic surgery, vascular surgery, and pediatric surgery. Among specialty fellowships available in the United States, pediatric surgery fellowships are 1 of only 10 specialties that filled over 97% of all available positions.[7] Comparatively, thoracic surgery filled 72% of positions offered in the match, while among fellowships overall 89% filled. With regard to medical school type, the nomogram shows the full spectrum of matching in placement. Osteopathic medical school graduates portend a less favorable placement success into a surgical subspecialty compared with an allopathic public medical school graduate (63.1% vs 84.4%, respectively).[8] Applicants and fellowship directors could also use the nomogram to improve the efficiency of applying and screening applications. Applicants for these surgical subspecialties could make better informed decisions about applying to fellowship or determine which variables may have the greatest effect on success of placement. For example, residency applicants can improve their chances of success by increasing the number of publications, becoming chief resident, and volunteering more and the nomogram is able to provide information on how much influence each of these factors has. Fellowship directors could utilize the nomogram as a prescreening tool to objectively review critical characteristics for applicants to their program. This could make the interviewing and application process more objective and efficient. More competitive fellowship directors may set the target score higher for interviewing applicants who are predicted to gain placement at a higher rate. Conversely, a less competitive fellowship program that did not match in the last year may set their target score lower and thus have access to interview and rank more applicants.

We predicted successful placement without, including several common more subjective pieces of information in the nomogram. First, we did not include letters of recommendation. Fellowship applicants generally select their supporters as recommendation writers. The letters provide a subjective assessment of the applicant and we are unable to quantify them. Second, we did not include interviewing because this event often occurs after a prescreening of the application takes place and because it is mostly a subjective process. While the human intuition of an interview is no doubt critical in the selection process for extremely good or bad applicants, numerous biases interfere with the ability to mentally calculate probabilities which diminishes the ability to quantitate the likelihood of matching success for the majority of applicants.[9] The fellowship successful placement nomogram is designed to be used as an objective method to supplement the subjective nature of letters of recommendation and the interview.

The nomogram has limitations. The model becomes unstable at the extremes; thus the likelihood of gaining placement can only be predicted at $>$ 10% or $<$ 90% level. Additionally, we were unable to use USMLE or American Board of Surgery In-Training Examination scores due to confidentiality concerns that precluded our use of the data. The original purpose of the National Board of Medical Examiners was to provide high-quality examinations acceptable for licensing purposes by state agencies, not for ranking residency or fellowship candidates. Although USMLE Step 1 and 2 scores do correlate with how well individuals perform on standardized board examination multiple-choice questions, they do not correlate with reliable measures of clinical skill acquisition that would matter in an operating room or clinical environment.[10] As a substitute for USMLE scores, we used other predictors that may be affected by USMLE performance, such as Alpha Omega Alpha honor society membership. Additionally, we were unable to use an applicant's or director's rank order list. This ranking information would be helpful to determine the effect these characteristics have on matching success based on where an applicant ranks a program and vice versa. Finally, we were only able to use the limited number of surgical fellowships available to us in the AAMC dataset.

Overall, we feel that this nomogram for predicting surgical subspecialty fellowship placement success could be used to augment decision making by applicants and program directors. Larger samples of applicants, additional medical and surgical subspecialties and novel predictors are likely to improve nomogram accuracy. Although a nomogram, such as this will never replace the complex decisions required by both applicants and program directors, it may serve to provide objective data to those who choose to be informed in the high-stakes decisions of a career choice.

ACKNOWLEDGMENTS

The authors thank Lauren Johnson and Walter Fitz-William at the American Association of Medical Colleges, Washington, DC, for data collection. Funding was provided by the American College of Obstetricians and Gynecologists and the University of Missouri Kansas City. The study was given a review and approved by the Cleveland Clinic Institutional Review Board.

REFERENCES

1. National Residency Matching Program, Specialties Matching Service. Figure 2. Available at: http://www.nrmp.org/data/resultsanddatasms2010.pdf.

2. Center for Workforce Studies. The complexities of physician supply and demand: projections through 2025. Available at:

https://services.aamc.org/publications/showfile.cfm?file=version122.pdf&prd_id=244&prv_id=299&pdf_id=122.

3. Cooper RA. There's a shortage of specialists. Is anyone listening? *Acad Med*. 2002;77:761-766.

4. Robert Graham Center. Specialty and geographic distribution of the physician workforce: what influences medical student and resident choices? Available at: http://www.graham-center.org/online/etc/medialib/graham/documents/publications/mongraphs-books/2009/rgcmo-specialty-geographic.Par.0001.File.tmp/Specialty-geography-compressed.pdf.

5. Siddighi S, Barker M, Pancholy A, et al. Attitudes and perceptions regarding subspecialty training in female pelvic medicine and reconstructive surgery. *Int Urogynecol J Pelvic Floor Dysfunct*. 2008;19:1523-1526.

6. Harrell FE, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Statist Med*. 1984;3:143-152.

7. National Residency Matching Program, Specialties Matching Service. Table 1. Available at: http://www.nrmp.org/data/resultsanddatasms2010.pdf.

8. National Residency Matching Program, Specialties Matching Service. Table 4. Available at: http://www.nrmp.org/data/resultsanddatasms2010.pdf.

9. Hogarth RM. *Judgement and Choice: The Psychology of Decision*. New York, NY: Wiley; 1987.

10. McGaghie WC, Cohen ER, Wayne DB. Are United States medical licensing Exam step 1 and 2 scores valid measures for postgraduate medical residency selection decisions? *Acad Med*. 2011;86:48-52.