Potential Impact of Pass/Fail Scores on USMLE Step 1: Predictors of Excellence in Obstetrics and Gynecology Residency Training

Sejal Tamakuwala1, Joshua Dean2, Katherine J. Kramer3, Adib Shafi4, Sarah Ottum5, Joshua George6, and Maurice-Andre Recanati4

1Emory University, Atlanta, GA, USA. 2Beaumont, Royal Oak, MI, USA. 3St. Vincent’s Catholic Medical Centers, New York, NY, USA. 4Wayne State University, Detroit, MI, USA. 5Detroit Medical Center, Detroit, MI, USA. 6University of New Mexico, Albuquerque, NM, USA.

ABSTRACT

AIM: The study aims to determine resident applicant metrics most predictive of academic and clinical performance as measured by the Council of Resident Education in Obstetrics and Gynecology (CREOG) examination scores and Accreditation Council for Graduate Medical Education (ACGME) clinical performance (Milestones) in the aftermath of United States Medical Licensing Examination Scores (USMLE) Step 1 becoming a pass/fail examination.

METHODS: In this retrospective study, electronic and paper documents for Wayne State University Obstetrics and Gynecology residents matted over a 5-year period ending July 2018 were collected. USMLE scores, clerkship grade, and wording on the letters of recommendation as well as Medical Student Performance Evaluation (MSPE) were extracted from the Electronic Residency Application Service (ERAS) and scored numerically. Semiannual Milestone evaluations and yearly CREOG scores were used as a marker of resident performance. Statistical analysis on residents (n = 75) was performed using R and SPSS and significance was set at P < .05.

RESULTS: Mean USMLE score correlated with CREOG performance and, of all 3 Steps, Step 1 had the tightest association. MSPE and class percentile also correlated with CREOGs. Clerkship grade and recommendation letters had no correlation with resident performance. Of all metrics provided by ERAS, none taken alone, were as useful as Step 1 scores at predicting performance in residency. Regression modeling demonstrated that the combination of Step 2 scores with MSPE wording restored the predictive ability lost by Step 1.

CONCLUSIONS: The change of USMLE Step 1 to pass/fail may alter resident selection strategies. Other objective markers are needed in order to evaluate an applicant’s future performance in residency.

KEYWORDS: USMLE, Step 1, in-service examination, CREOG, resident selection

Introduction

The success of an obstetrics and gynecology (Ob/Gyn) residency program relies upon recruiting and retaining candidates who will excel both academically and clinically. Most candidates apply to Ob/Gyn residencies through the Electronic Residency Application Service (ERAS), which delivers the applicant’s data to programs. Such data includes medical school transcripts, letters of recommendation, Medical Student Performance Evaluation (MSPE) and United States Medical Licensing Examination Scores (USMLE). These are used by programs to select candidates for an interview and to compile the National Resident Matching Program (NRMP) match list.

In March 2019, the Invitational Conference on USMLE Scoring (InCUS)1 recommended that, to reduce the adverse impact of the current overemphasis on USMLE performance in residency selection, Step 1 would change to pass/fail reporting in 2022.2 Instead of reporting the examinee’s score as it is currently, programs and test takers will no longer have access to this numerical value. While the USMLE was designed to assess a physician’s ability to apply the knowledge, concepts, and principles and demonstrate patient-centered skills that constitute the basis of safe and effective care and was intended for eligibility for licensure, many programs relied on this objective metric in resident ranking. While the National Resident Matching Program (NRMP)’s Program Director Survey, which sheds light on the factors used to select residents to interview and rank, demonstrates that Step 1 scores are of paramount importance, with 76% of programs reporting a target score.3 In Ob/Gyn, other factors, such as failed Steps, NRMP match violations, perceived commitment to the specialty, clerkship grades (particularly in Ob/Gyn), Dean’s letter, Alpha–Omega–Alpha Status, lack of gaps in training, leadership qualities and, after interviews, interactions with faculty/staff and interpersonal skills were also cited in the survey.
With the impending loss of USMLE Step 1 scores, and the limitations in audition rotations due to the paucity of available spots, we looked at other metrics available to program directors at the time of application to see how they compared with Step 1 scores in predicting clinical and academic performance. Ultimately the goal of this investigation was to establish an evidence-based algorithm for initial evaluation of residency applicants.

### Materials and Methods

Under Institutional Review Board (IRB) approval (042317MP2X), data for this retrospective study was collected from electronic and paper documents on residents who were matriculated in Wayne State University’s Obstetrics and Gynecology Residency program between July 2013 and July 2018. All data was fully anonymized and IRB waived requirements for informed consent.

#### Undergraduate medical education data

Medical students applied to our program through ERAS (AAMC, Washington DC). Data supplied by ERAS included: USMLE Step 1 and 2 scores (or the Comprehensive Osteopathic Medical Licensing Examination (COMLEX) scores), Step 3 scores (only if available), transcript containing basic science and clerkship grades, and medical school class ranking. Individual USMLE Step scores as well as the “average USMLE score,” defined as the average of Step 1, 2, and 3 scores was included in our analysis. Medical school class rank was normalized to percentile (ranking/total class size) and clerkship grade was translated into an honors/pass/fail scale. When the MSPE (also known as the dean’s letter) specified academic standing scores or displayed this data graphically, that score was converted into quintiles and incorporated into the dataset. When only textual descriptions were presented, the MSPE score was based on the strength of the wording in the letter as well as the final adjective using the validated 5-point scale used by the Dean’s office at the Wayne State School of Medicine (Figure 1). Letters of recommendation were scored on a 5-point scale corresponding to the subjective strength of the recommendation as determined by 2 blinded faculty members experienced in the residency selection process.

#### Graduate medical education data

Scores on the Council on Resident Education in Obstetrics and Gynecology (CREOG) multiple choice in-training examination, which objectively assesses a resident’s knowledge in the specialty on a yearly basis, were used as a marker of academic performance. Each yearly CREOG score, as available, as well as the “average CREOG” value, defined as the average score obtained over the 4 yearly exams, was used in our analysis. Clinical performance was evaluated using the Accreditation Council for Graduate Medical Education (ACGME) Milestones. This metric, graded 1 to 5, is comprised of fine-grained developmental levels that correspond to and expand on the core competencies in Ob/Gyn and summarize the consensus of a committee formed of multiple attendings, include assessments of technical and clinical skills, peer-review, end-of-rotation evaluations, patient feedback, and performance in simulations. We analyzed the 7 ACGME Milestone groups: obstetrics, gynecology, office, systems-based practice (SBP), practice-based learning (PBL), professionalism (PRO), and interpersonal (ICS) (Supplemental Figure 1), individually using matrix analysis. Clinical success was defined as more than 3 (out of 7) Milestone groups rated as excellent (4.0) or outstanding (5.0) in any single evaluation cycle.

#### Data analysis

Data was de-identified and tabulated in an Excel (Microsoft) spreadsheet. Statistical analysis was performed using R programming (version 3.51, The R Foundation). Significance for differences in mean values was inferred with the Mann–Whitney test. Enrichment analysis was performed using the Fisher’s exact test. Correlations between different measurements were calculated using Pearson’s correlation. Bonferroni correction was used when appropriate. For each case, statistical significance was considered to be a P-value < .05.

| MSPE/Dean’s letter score | Outstanding | 5 |
|--------------------------|-------------|---|
|                          | Excellent/High Regard | 4 |
|                          | Very Good    | 3 |
|                          | Good         | 2 |
|                          | Satisfactory | 1 |
| OBGYN clerkship grade    | Honors       | 2 |
|                          | Pass         | 1 |
|                          | Fail         | 0 |

*Figure 1. Scoring system. Scoring used for Medical Student Performance Evaluation (MSPE)/Dean’s letter and clerkship grade.*
Predictive modeling was performed using SPSS (version 26 from IBM). A stepwise routine was used to build models with only the predictors that are significant, generally with $P < .05$; however, $P < .10$ was used when trying to model Milestones. The routine takes a stepwise approach when entering the eligible predictor variables into the equation, which takes the form $\text{Dependent} = (\text{Coefficient} \times \text{Predictor 1}) + (\text{Coefficient} \times \text{Predictor 2}) + \text{Constant}$. Linear regression was used for CREOG and Milestones since all variables are continuous.

**Regression modeling**

Predictive modeling was performed using SPSS (version 26 from IBM). A stepwise routine was used to build models with only the predictors that are significant, generally with $P < .05$; however, $P < .10$ was used when trying to model Milestones. The routine takes a stepwise approach when entering the eligible predictor variables into the equation, which takes the form $\text{Dependent} = (\text{Coefficient} \times \text{Predictor 1}) + (\text{Coefficient} \times \text{Predictor 2}) + \text{Constant}$. Linear regression was used for CREOG and Milestones since all variables are continuous.

**Step 1 Cutoff Scores:** Using our dataset, each resident’s USMLE step 1 score was plotted against CREOG scores and graphed using Excel (Microsoft)A. Data was fitted using polynomial curve analysis and extrapolation of Step 1 score corresponding to CREOG values was performed using Matlab (Mathworks).

**Results**

**Descriptive statistics on cohort of residents**

The dataset, which includes residents who were enrolled from July 2013 to 2018, comprises $n = 75$ residents, 22 males (29.3%), and 53 females (70.7%; Figure 2).
Correlation between USMLE and markers of success as a resident

The mean USMLE score was 219.9 ± 1.65 SEM (range: 191.3–255.7) and was positively correlated with the mean CREOG score over the 4 years of residency, 195.4 ± 2.05 SEM (range: 154–248.8) (r = 0.67; P < .001). Of the 3 USMLE scores, Step 1 (mean 220.5 ± 1.94 SEM, range: 188–255) had the highest correlation with average CREOG scores (r = 0.64; P < .001). However, the average of the 3 USMLE scores has a higher statistical correlation with average CREOG scores than the USMLE Step 1 score alone (Figure 3). The MSPE and class percentile were correlated with CREOGs, but performance on the Ob/Gyn core clerkship and wording in the letters of recommendation were not correlated with CREOGs.

Milestone scores were associated with receiving a score of “honors” in the Ob/Gyn clerkship and in scores >200 on the USMLE Step 1; however, these did not achieve statistical significance. Letter of recommendation score, MSPE score, and average USMLE score were not correlated with average Milestone scores over the entire residency training period.

Predictors of a student’s success as a resident

Regression analysis mirrors these results. Mean USMLE score was the dominant predictor variable for CREOG scores and the only variable that significantly explains CREOG1, 2, 4 and mean CREOG (P-values all <.001 and R² = 0.4–0.6). While USMLE 1 and 2 (available at the time of application) are predictive of mean CREOGs (R² = 0.435), the effect of losing Step 1 and relying only on Step 2 scores significantly reduces the regression (R² = 0.316). However, adding back the MSPE letter’s wording to the regression restores predictive value (R² = 0.494). Similarly, while USMLE 1 and 2 together are predictive of CREOG 1 scores (R² = 0.480), relying on Step 2 alone reduces the predictive ability (R² = 0.331) while adding back class percentile to the regression restores the regression’s predictive ability (R² = 0.496) (Figure 4).

Similar stepwise regression analysis shows that the strength of recommendation letters is a mild predictor of Milestones’ performance. MSPE wording is a mild predictor of ICS, PRO, and mean Milestones (P = .036). Regression analysis showed that neither average USMLEs nor any of its components could be used to predict an applicant’s future Milestone scores.

Predictors of CREOG scores based on USMLE step 1

Regression using polynomial curve fitting analysis, comparing USMLE1 to average CREOGs, was fitted with an R² = 0.45. Setting a threshold USMLE1 score > 209 corresponded to average CREOGs of 199 (Figure 5).

Discussion

Correlation of USMLE performance with resident success

The InCUS recommended to “accelerate research on the correlation of USMLE performance to measures of residency performance.” Thus, in this study, we sought out to assess which medical student applicant metrics could predict success in residency. These predictors are important to program directors in order to identify the strongest candidates in a pool of applicants that could help contribute the most toward a culture of academic success, patient safety, and development of excellence within a training program. We used CREOGs as a marker of academic performance and ACGME Milestones as a measure of clinical performance. Our study suggests that USMLE scores exhibited the best correlation with CREOG scores and senior year Milestone scores. This association of markers of residency performance with USMLE scores was supported by other studies.10,11 In 1 analysis, USMLE scores > 200 were indicative of higher CREOG scores and 100% written board examination pass rate when analyzing 40 residents with available USMLE scores from 8 accredited obstetrics and gynecology programs.12 Furthermore, USMLE Step 1 scores correlated with performance on CREOGs for PGY1 to 3 years, USMLE Step 2 scores correlated with CREOG scores for all 4 PGY years,13 and was associated with successfully matching into fellowship.14 While USMLE Step 1 had the strongest correlation for success on CREOGs, when compared to Step 2 or 3 (and Steps 2 and 3 combined), the composite score of all 3 USMLEs was a better indicator. Unfortunately, at the time most medical students apply for residency, Step 3 data are unavailable.

Model projecting performance with USMLE step 1 as a pass/fail exam

With the controversial change of Step 1 to pass/fail in 2022, this useful metric will no longer be available. Thus, there is an impetus to analyze datasets and identify alternative metrics

| USMLE2 | USMLE3 | Avg. USMLE | Avg. CREOG |
|--------|--------|------------|------------|
| USMLE1 | 0.65***| 0.67***    | 0.89***    | 0.64***    |
| USMLE2 | 0.67***| 0.89***    | 0.87***    | 0.61***    |
| USMLE3 |        |            |            | 0.67***    |
| Avg. USMLE |        |            |            |            |

Figure 3. Correlation of USMLE and CREOG scores. CREOG scores are highly correlated with average USMLE scores. Of the 3 USMLEs, Step 1 scores were most highly correlated with CREOGs. The inter-correlation for each of the USMLE exams individually and for CREOGs is shown. *** signifies p ≤ .001. (n = 75).

Abbreviations: USMLE, United States Medical Licensing Examination Scores; CREOG, Council of Resident Education in Obstetrics and Gynecology.
Figure 4. CREOG regression analysis summary. Stepwise linear regressions were run to establish a relationship between CREOG performance and USMLEs. Step 1 was highly related with CREOGs and the loss of this metric impacts predicting applicant’s performance. Unshaded rows include Step 1, while shaded rows omit Step 1, and dark-shaded rows add additional metrics. $R^2$ shows that it is possible to retain predictive ability after the loss of Step 1 if other metrics, such as MSPE wording or class rank are included.

Abbreviations: CREOG, Council of Resident Education in Obstetrics and Gynecology; USMLE, United States Medical Licensing Examination Scores
which could be used in its place and which would offer similar reliability. Residency programs tend to consider applicant’s Ob/Gyn clerkship grade and strength of letter of recommendation; however, our study demonstrated that these elements were not predictive of performance. This finding was supported by other studies, which found that letters of recommendation tended to be biased, suppressed negative factors, and highly subjective, leading to their poor predictability of resident success. While some studies suggested that USMLE Step 2 scores correlated with CREOG scores for all 4 PGY years, and was associated with successfully matching into fellowship, our study showed minimal correlation for Step 2 scores alone. As the model suggests, the predictive ability lost by numeric score on Step 1 may be regained by using the wording on the MSPE or the class percentile at graduation from medical school.

USMLE step 1 cutoffs as a pass/fail exam

After Step 1 becomes pass/fail, can a certain passing threshold score be established, and be predictive of good performance on CREOGs? The simple redefinition of the “passing grade” to 209 (from the current value of 194) would predict the ability of an applicant to score about 199 on CREOGs (corresponding to the average grade of all test-takers taking the examination). While some program directors have advocated raising the passing score, our mathematical model demonstrates that an increase would be unrealistic as it would make licensure (the primary purpose of the test) more challenging to obtain.

Exploring alternative approaches

While our algorithm integrating Step 2 with other metrics can mitigate the impact of the changes, many program directors have advocated integrating a greater number of metrics for a more holistic review of applicants across medical schools. Such may include medical school awards (Gold Humanism, Alpha–Omega–Alpha), leadership roles, extracurricular activities, participation in research, content of the personal statement, perseverance, grades from core clerkships, data from audition electives, and knowledge of the applicant. Because programs need to screen a large number of candidates in a short amount of time, a holistic review of each applicant may not be feasible. Tools, such as the National Board of Medical Examiners (NBME) Clinical Science Examination (“shelf”) scores, technical skills assessments, professionalism and communication skills, emotional quotients, and fitness/aptitude testing for a particular specialty will need to be developed and validated. However, the lack of other good metrics may result in rethinking the residency selection process, create new assessment tools, widen the definition of success in medical school beyond academics, and create a selection process free of bias and structural racism. These changes may eventually help fulfill our obligation to diversity and inclusion in the healthcare field. Alternatively, if the paradigm is kept as it currently is with the exception of Step 1 quantitation, students may feel that Step 2 scores are the single most important exam during medical school and would create excessive stress. A bad performance 1 day and on a single exam should not be lethal to a young doctor’s career. Furthermore, as more emphasis is placed on MSPE and class rank, program directors may be biased in

Figure 5. Regression analysis. Setting passing score on USMLE Step 1 to 209 is predictive of performance on CREOGs. Regression analysis shows that Step 1 was highly related with CREOGs. A Step 1 score >209 was predictive of average CREOG scores of 199. Since the CREOG exam is scored with a mean of 200 and a standard deviation of 20 for all test-takers, setting the passing score for Step 1 to 209 would reassure program directors that the candidate will do well on the CREOG exam (n = 75).

Abbreviations: CREOG, Council of Resident Education in Obstetrics and Gynecology; USMLE, United States Medical Licensing Examination Scores
selecting residents from well-known schools. In turn, this may make it even harder for well-qualified applicants from other schools, and particularly foreign programs, to match at all.

Weakness of this study includes the small sample size, the use of subjective data such as the letters of recommendations and the Milestone data, and, the narrow study period. While Milestones are subjectively scored, and thus may have bias, in our program the clinical competency committee balances out the subjective impressions of the faculty attendings evaluating the resident, somewhat mitigating this weakness. Strengths of this study include the single program site data with 10 residents per year, eliminating interprogram variability found in other papers, and the predictive regression analysis. While the diversity in our program, including a component of osteopathic physicians, is a strength, our results, however, may not be applicable to other institutions.

Conclusion
As of this writing, there is no consensus among program directors as to an optimal approach to help guide applicant selection in the absence of numerical Step 1 data. While the literature discusses various approaches, this paper creates a model which suggests that, by adding to Step 2 scores MSPE data and class rank, the ability of predicting residency performance lost by Step 1 can be, at least partially, regained. Other approaches, such as changing the cutoff for Step 1, were mathematically explored and found to be impractical. New objective metrics will need to be created or validated to help programs sort out applicants as we embark on an era of a pass/fail USMLE.

Acknowledgments
This research was made possible by the work and dedication the obstetrics and gynecology residents at Wayne State University and by the faculty and staff at the Detroit Medical Center, in particular, the Sorin Draghichi laboratory, which helped provide statistical analysis. MAR also wishes to thank the NIH-Women’s Reproductive Health Research Career Development Award at Wayne State University for ongoing research and support.

Ethical Approval
This work was conducted under the approval of the Wayne State Institutional Review Board (042317MP2X) as well as the Detroit Medical Center’s contract research organization (CRO), (DMC14380). All data for this retrospective chart review of adult learners was completely anonymized and IRB waived informed consent.

Previous Presentation
This work was presented in abstract form for poster publication at the Council on Resident Education in Obstetrics and Gynecology (CEOG) & the Association of Professors of Gynecology and Obstetrics (APGO) in New Orleans, LA, February 2019.

Author Contributions
ST: conception, data extraction, wrote parts of manuscript. JD: data extraction, wrote parts of manuscript. KK: conception, design, drafting manuscript, created figures. AS: statistical analysis and interpretation. SO: statistical analysis, created figures. JG: manuscript revisions. SK: data extraction, interpretation of data, rewrote parts of manuscript. CC: revising the manuscript, important intellectual contributions. MR: conception of study, wrote part of manuscript, analyzed statistics and provided guidance and direction, responsible for the accuracy and integrity of the work presented here. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript.

ORCID iDs
Sarah Ottum https://orcid.org/0000-0002-8729-1095
Joshua George https://orcid.org/0000-0001-6188-6755
Conrad R. Chao https://orcid.org/0000-0003-1107-8703
Maurice-Andre Recanati https://orcid.org/0000-0003-0974-3962

Supplemental Material
Supplemental material for this article is available online.

REFERENCES
1. United States Medical Licensing Examination, Summary Report and Preliminary Recommendations from the Invitational Conference on USMLE Scoring 2019; https://www.usmle.org/pdfs/incus/incus_summary_report.pdf.
2. United States Medical Licensing Examination. Invitational Conference on USMLE Scoring: change to pass/fail score reporting for Step 1, 2020; https://www.usmle.org/incus/
3. The National Resident Matching Program. Results of the 2018 NRMP Program Director Survey. 2018; https://www.nrmp.org/wp-content/uploads/2018/07/NRMP-2018-Program-Director-Survey-for-WWW.pdf.
4. Wells M, Polimare M, Nesi LJ, Dunn JC. Factors considered when ranking military orthopedic surgery residency candidates. Mil Med. 2020;185:e1603-e1607.
5. Ward MA, Palazzi DL, Lorin MI, Agrawal A, Frankenthal H, Turner TL. Impact of the final adjective in the Medical Student Performance Evaluation on determination of applicant desirability. Med Educ Online. 2018;23:1542922.
6. Visscher FC. Validity and purpose of in-training examination in obstetrics and gynecology, 1979-1982. Obstet Gynecol. 1984;63:253-259.
7. Carraccio CL, Benson BJ, Nixon LJ, Derstine PL. From the educational bench to the clinical bedside: translating the dreyfus developmental model to the learning of clinical skills. Acad Med. 2008;83:761-767.
8. Watson RS, Borger AJ CT OH, et al. A multicenter prospective comparison of the accreditation council for graduate medical education milestones: clinical competency committee versus Resident self-assessment. J Surg Educ. 2017;74:e8-e14.
9. Lyle B, Borger AJ, Kallies KJ, Jarman BT. Do attending surgeons and residents See Eye To Eye? An evaluation of the accreditation council For graduate medical education milestones in general surgery residency. J Surg Educ. 2016;73:e54-e58.
10. Spellacy WN, Downes KL. United States medical licensing examination scores as a predictor of performance on the annual council of resident education in obstetrics and gynecology examinations. J Reprod Med. 2014;59:17-19.
11. Hamstra SJ CM, Jurich D, Yamazaki PL, et al. Exploring the association between USMLE scores and ACGME milestone ratings: a validity study using national data from emergency medicine. Acad Med. 2021;96(1):1324-1331.
12. Armstrong A, Alvero R, Nielsen P, et al. Do U.S. Medical licenses examination step 1 scores correlate with council on resident education in obstetrics and gynecology in-training examination scores and American board of obstetrics and gynecology written examination performance? Mil Med. 2007;172:640-643.
13. Bell JG, Kanellitis I, Shaffer L. Selection of obstetrics and gynecology residents on the basis of medical school performance. Am J Obstet Gynecol. 2002;186:1091-1094.
14. Franasiak J, Kaser D, Goodman L, et al. Characterization of reproductive endocrinology and infertility (REI) fellowship applicants: guiding our mentees toward success. Fertil Res Pract. 2017;3:18.

15. Weiner S. Should the USMLE be pass/fail? AAMC News 2019; https://www.aamc.org/news-insights/should-usmle-be-pass-fail.

16. Edmond M, Roberson M, Hasan N. The dishonest Dean’s letter: an analysis of 532 Dean’s letters from 99 U.S. Medical schools. Acad Med. 1999;74:1033-1035.

17. Belovich AN, Bahner I, Bonaminio G, et al. USMLE Step-1 is going to pass/fail, Now what Do We Do? Med Sci Educ. 2021;31(4):1-6.

18. Willett LL. The impact of a pass/fail step 1 - A residency program director’s view. N Engl J Med. 2020;382:2387-2389.

19. Gu A, Farrar J, Fassihi SC, et al. Effect of change in USMLE step 1 grading on orthopaedic surgery applicants: a survey of orthopaedic surgery residency program directors. J Am Acad Orthop Surg Glob Res Rev. 2021;5:e20 00216.

20. Mun F, Scott AR, Cui D, Chisty A, Hennessikus WL, Hennessikus EF. Internal medicine residency program director perceptions of USMLE step 1 pass/fail scoring: a cross-sectional survey. Medicine (Baltimore). 2021;100:e25284.

21. Makhoul AT, Pontell ME, Ganesh Kumar N, Drolet BC. Objective measures needed - program Directors’ perspectives on a pass/fail USMLE step 1. N Engl J Med. 2020;382:2389-2392.

22. Youmans QR, Essien UR, Qt Capers. A test of diversity - what USMLE pass/fail scoring means for medicine. N Engl J Med. 2020;382:2393-2395.