Table 1. Primary and Secondary Outcomes in Patients Undergoing Awake and Anesthetized Calibration

	Calibrated Awake	Calibrated Anesthetized	*P* Value
Primary outcome			
Total duration, normalized (min)	51 ± 14	51 ± 14	0.624
Secondary outcomes			
Total duration, non-normalized (min)	46 ± 12	47 ± 13	0.406
Duration of deep block (min)	19 ± 5	19 ± 5	0.573
Duration to train-of-four ratio 25% (min)	32 ± 7	32 ± 7	0.550
Duration to train-of-four ratio 50% (min)	36 ± 9	37 ± 9	0.125
Duration to train-of-four ratio 75% (min)	41 ± 12	42 ± 11	0.174
Onset time(s)	140 ± 51	139 ± 59	0.740
Baseline train-of-four ratio after calibration (%)	112 ± 6	111 ± 7	0.593
Stimulation current after calibration (mA)	45 ± 13	44 ± 13	0.751

The values are presented as mean ± SD. Total duration, normalized is the time in minutes from rocuronium injection until recovery to a normalized train-of-four ratio of 0.9. Total duration, non-normalized is the time in minutes from rocuronium injection until recovery to a non-normalized train-of-four ratio of 0.9. Duration of deep block is the time in minutes from rocuronium injection to reappearance of the first response to post-tetanic count stimulation. Duration to train-of-four ratio 25% is the time in minutes from rocuronium injection until 25% recovery of non-normalized train-of-four ratio. Duration to train-of-four ratio 50% is the time in minutes from rocuronium injection until 50% recovery of non-normalized train-of-four ratio. Duration to train-of-four ratio 75% is the time in minutes from rocuronium injection until 75% recovery of non-normalized train-of-four ratio. Onset time is the time in seconds from rocuronium injection to 95% depression of the first twitch of the train-of-four.

2. Fuchs-Buder T, Claudius C, Skovgaard LT, Eriksson LI, Mirakhur RK, Viby-Mogensen J; 8th International Neuromuscular Meeting: Good clinical research practice in pharmacodynamic studies of neuromuscular blocking agents II: The Stockholm revision. Acta Anaesthesiol Scand 2007; 51:789–808
3. Claudius C: Calibration of the neuromuscular monitor: Is it necessary? Can J Anaesth 2016; 63:354–5
4. Schreiber JU, Mucha E, Fuchs-Buder T: Acceleromyography to assess neuromuscular recovery: Is calibration before measurement mandatory? Acta Anaesthesiol Scand 2011; 55:328–31
5. Claudius C, Skovgaard LT, Viby-Mogensen J: Is the performance of acceleromyography improved with preload and normalization? A comparison with mechanomyography. Anesthesiology 2009; 110:1261–70
6. Bland JM, Altman DG: Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1:307–10
7. Nemes R, Nagy G, Murphy GS, Logvinov II, Fülesdi B, Renew JR: Awake volunteer pain scores during neuromuscular monitoring. Anesth Analg 2020; 130:941–8
8. Connelly NR, Silverman DG, O'Connor TZ, Brull SJ: Subjective responses to train-of-four and double burst stimulation in awake patients. Anesth Analg 1990; 70:650–3
9. Saitoh Y, Nakazawa K, Toyooka H, Amaha K: Optimal stimulating current for train-of-four stimulation in conscious subjects. Can J Anaesth 1995; 42:992–5
10. Baillard C, Bourdiau S, Le Toumelin P, Ait Kaci F, Riou B, Cupa M, Samama CM: Assessing residual neuromuscular blockade using acceleromyography can be deceptive in postoperative awake patients. Anesth Analg 2004; 98:854–7

(Accepted for publication August 30, 2022. Published online first on October 26, 2022.)

Personalized Surgical Transfusion Risk Prediction: Comment

To the Editor:

We have read the study titled "Personalized Surgical Transfusion Risk Prediction Using Machine Learning to Guide Preoperative Type and Screen Orders" by Lou *et al.*[1] and the accompanying editorial titled "Moving from 'Surgeries' to Patients: Progress and Pitfalls While Using Machine Learning to Personalize Transfusion Prediction" by Mathis *et al.* [2] The authors include 4 million surgical cases during a 3-yr period from the American College of Surgeons National Surgical Quality Improvement Program database. The authors used the American College of Surgeons National Surgical Quality Improvement Program database to develop a machine learning model that incorporates patient- and surgery-specific variables to predict transfusion risk and the associated need for preoperative type and screen. The authors hypothesize that their machine learning algorithm would outperform the traditional approach of relying primarily on historical surgery-specific transfusion rates and thus optimize resource allocation by decreasing blood bank waste. The machine learning algorithm recommends fewer preoperative type and screen orders.

The study presents in exceptional detail the methodologic approach to developing highly accurate algorithms to predict transfusion risk. Several authors have shown that race is an independent predictor of postoperative transfusion across surgical disciplines, associated with

either higher or lower rates of transfusion.[3,4] However, the work by Lou *et al.* did not include any mention of race or ethnicity. Experience from previous algorithms used to model resource allocation in health care demonstrate that omitting this information may lead to perpetuating bias that unfortunately exists within the United States healthcare system.[5,6] Although the intent from healthcare providers is to provide the best possible care to their patients, determinants of health are closely linked to race and ethnicity and availability of resources in the United States. Therefore, artificial intelligence models aiming to personalize medicine can present pitfalls for those already with low resource availability, unwittingly withholding care in marginalized communities.[5] On the other hand, it has been shown that including race or ethnicity in machine learning models may perpetuate bias, and therefore including race and ethnicity in artificial intelligence remains intensely debated.[7,8]

Our primary question is the following: Why did the authors choose not to include race and ethnicity in their table 1 or in their prediction model? Was the absence of any demographic data in the Lou *et al.* article an intended or inadvertent omission? Given the potential impact of a patient's race or ethnicity on clinician decision-making, and the ongoing controversy about the use of these variables in clinical prediction models,[8] we believe that an explanation for the absence of this data would be helpful. Inclusion of ethnic and racial minorities in research is important, and transparency is key in the design of prediction models to improve societal health.

Competing Interests

Dr. Cannesson has funding from Masimo (Irvine, California), Edwards Lifesciences (Irvine, California), and the National Institutes of Health (Bethesda, Maryland) for unrelated work and is a shareholder of Sironis (Newport Beach, California) and Perceptive Medical (Newport Beach, California). The other authors declare no competing interests.

Brittany N. Burton, M.D., M.H.S., M.A.S., Cecilia Canales, M.D., M.P.H., Jennifer Lucero, M.D., M.A., Maxime Cannesson, M.D., Ph.D. University of California-Los Angeles Health, Los Angeles, California (B.N.B.). bburton@mednet.ucla.edu

DOI: 10.1097/ALN.0000000000004397

References

1. Lou SS, Liu H, Lu C, Wildes TS, Hall BL, Kannampallil T: Personalized surgical transfusion risk prediction using machine learning to guide preoperative type and screen orders. Anesthesiology 2022; 137:55–66
2. Mathis MR, Singh K, Kheterpal S: Moving from "surgeries" to patients: Progress and pitfalls while using machine learning to personalize transfusion prediction. Anesthesiology 2022; 137:9–12
3. Menendez ME, Ring D: Minorities are less likely to receive autologous blood transfusion for major elective orthopaedic surgery. Clin Orthop Relat Res 2014; 472:3559–66
4. Francis JJ, Tinmouth A, Stanworth SJ, Grimshaw JM, Johnston M, Hyde C, Stockton C, Brehaut JC, Fergusson D, Eccles MP: Using theories of behaviour to understand transfusion prescribing in three clinical contexts in two countries: Development work for an implementation trial. Implement Sci 2009; 4:70
5. Obermeyer Z, Powers B, Vogeli C, Mullainathan S: Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019; 366:447–53
6. Canales C, Lee C, Cannesson M: Science without conscience is but the ruin of the soul: The ethics of big data and artificial intelligence in perioperative medicine. Anesth Analg 2020; 130:1234–43
7. Pfob A, Sidey-Gibbons C: Systematic bias in medical algorithms: To include or not include discriminatory demographic information? JCO Clin Cancer Inform 2022; 6:e2100146
8. Paulus JK, Kent DM: Race and ethnicity: A part of the equation for personalized clinical decision making? Circ Cardiovasc Qual Outcomes 2017; 10:e003823

(Accepted for publication September 27, 2022. Published online first on November 8, 2022.)

Personalized Surgical Transfusion Risk Prediction: Comment

To the Editor:

We have read with great interest the recent article by Lou *et al.*,[1] in which they used the American College of Surgeons National Surgical Quality Improvement Program participant use data file to expertly develop a transfusion prediction model with the goal of guiding type and screen ordering.

Lou *et al.* devised a clever method to broadly capture institution-specific transfusion information by redefining

the procedure-specific transfusion risk on local institutional data in their external validation experiments. The choice to use this process to improve model performance speaks to the importance of institution-specific data and to the assumption that inclusion of granular institutional data results in superior prediction. One example of an institution-specific variable that may confer additional predictive power is surgeon identifier, as there is evidence of inter-surgeon variability in transfusion requirements.[2–4] Also, it is unclear if anesthesiologist identifier is predictive of transfusion, which should be explored in greater detail. Widely externally valid approaches to modeling perioperative problems sacrifice data granularity that may be critical for practical implementation.

Research Support

Dr. Freundlich receives ongoing support from the National Institutes of Health (Bethesda, Maryland; No. K23HL148640 and No. UL1TR002243). Dr. Zapf receives ongoing support from the National Institutes of Health (No. T32GM108554).

Competing Interests

Dr. Freundlich has indicated the following financial relationships: Oak Hill Clinical Informatics (Oak Hill, Tennessee); 3M (Saint Paul, Minnesota). The other authors declare no competing interests.

Matthew A. C. Zapf, M.D.,
Robert E. Freundlich, M.D., M.S., M.S.C.I., F.A.S.A., F.C.C.M.,
Jonathan P. Wanderer, M.D., M.Phil., F.A.S.A., F.A.M.I.A.
Vanderbilt University Medical Center, Nashville, Tennessee
(M.A.C.Z.). matthew.a.zapf@vumc.org

DOI: 10.1097/ALN.0000000000004398

References

1. Lou SS, Liu H, Lu C, Wildes TS, Hall BL, Kannampallil T: Personalized surgical transfusion risk prediction using machine learning to guide preoperative type and screen orders. Anesthesiology 2022; 137: 55–66
2. Ejaz A, Spolverato G, Kim Y, Frank SM, Pawlik TM: Identifying variations in blood use based on hemoglobin transfusion trigger and target among hepatopancreaticobiliary surgeons. J Am Coll Surg 2014; 219:217–28
3. Aquina CT, Blumberg N, Probst CP, Becerra AZ, Hensley BJ, Noyes K, Monson JR, Fleming FJ: Large variation in blood transfusion use after colorectal resection: A call to action. Dis Colon Rectum 2016; 59:411–8
4. Aquina CT, Blumberg N, Probst CP, Becerra AZ, Hensley BJ, Iannuzzi JC, Gonzalez MG, Deeb AP, Noyes K, Monson JR, Fleming FJ: Significant variation in blood transfusion practice persists following upper GI cancer resection. J Gastrointest Surg 2015; 19:1927–37

(Accepted September 27, 2022. Published online first on November 8, 2022.)

Personalized Surgical Transfusion Risk Prediction: Reply

In Reply:

We thank Burton *et al.*[1] and Zapf *et al.*[2] for their thoughtful comments on our research on personalized surgical transfusion risk prediction.[3] Both letters raise important considerations regarding variable selection for predictive models in health care, which are worth discussing in further detail.

As Burton *et al.* note, race and ethnicity were not included as input variables in our machine learning model for surgical transfusion risk; we would like to clarify that this was intentional for several reasons, which we explain here. First, the inclusion of race in predictive models has been well-described to contribute to inequity.[4] One major limitation of machine learning is that a model can only learn from its training examples—in other words, real-world clinician behaviors. If such behaviors or the societal factors contextualizing that behavior are biased, the model will also be biased. The citation provided by Burton *et al.* is a perfect example of this[5]: in this study, researchers evaluated a model trained to predict healthcare utilization after hospital discharge, with the intention to allocate additional resources to patients predicted to have high utilization. Unfortunately, black patients had low utilization because they lacked access to care, which the model learned and perpetuated. Inclusion of race as an input variable in model development encourages machine learning models to explicitly encode such latent biases, and consequently the recommendations of such models will propagate systemic inequities in care.

Second, although race is a frequently collected variable in many datasets, it serves as a proxy for often unmeasured variables such as socioeconomic status, access to care, illness severity (due to poor access to care and delayed presentation), and other social determinants of health.[6] Thus, although the inclusion of race as a variable may improve

model discrimination, it potentially does so for the wrong reasons. Given two individuals, identical except for their skin color, it seems unjust for one to have a "better" prediction based on the population averages of their racial group, which may be due to unmeasured variables not applicable to the specific individual.

Third, to the best of our knowledge, there is little evidence that race itself contributes to risk for allogeneic blood transfusion after adjustment for disease burden, socioeconomic status, and other clinical variables that are known to contribute (*e.g.*, hematocrit). We thank Burton *et al.* for bringing attention to the potential pitfalls of racial adjustment and the critical importance of fairness in predictive modeling. As machine learning is increasingly used for clinical decision support, model developers must be vigilant for potential sources of bias, which can be introduced at every step of model development and implementation.[7,8] As a research community, we share a responsibility to ensure that the decision support tools we create do not exacerbate, and ideally help to reduce, the health disparities that are currently present in modern medicine.[9]

Zapf *et al.* raise important points about the benefits and limitations of model development using large registry datasets *versus* institution-specific datasets. We agree that inclusion of surgeon and anesthesiologist identifiers may further improve predictive performance. Variation in transfusion risk can occur due to differences in surgeon technique or case complexity, and it would be appropriate to adjust for these; however, they can also occur due to differences in preference for discretionary transfusion, which may be less appropriate to adjust for. By training our models on a large national database, we captured the average transfusion behavior of U.S. physicians, which we believe, on average, to be appropriate. Further customizing model predictions based on individual behavior patterns risks encoding undesirable physician practice patterns into the model; nonetheless, we acknowledge that such adjustment might be necessary for widespread adoption. Our transfer learning approach (*i.e.*, hospital-specific procedure-specific transfusion rate) could easily accommodate the addition of a surgeon- or anesthesiologist-specific adjustment, and it would be interesting to investigate such modifications in future work.

Research Support

Dr. Lou was supported by the National Institutes of Health (Bethesda, Maryland; 5T32GM108539-07) and the International Anesthesia Research Society (San Francisco, California).

Competing Interests

Dr. Kannampallil has consulting relationships with Pfizer Inc. (New York, New York) and Elsevier (Amsterdam, The Netherlands) that are unrelated to this work. Dr. Hall is the Consulting Director of the National Surgical Quality Improvement Program, American College of Surgeons (Chicago, Illinois). The other authors declare no competing interests.

Sunny S. Lou, M.D., Ph.D., Troy S. Wildes, M.D., Bruce L. Hall, M.D., Ph.D., M.B.A., Michael S. Avidan, M.B.B.Ch., Thomas Kannampallil, Ph.D. Washington University School of Medicine, St. Louis, Missouri (S.S.L.). slou@wustl.edu

DOI: 10.1097/ALN.0000000000004399

References

1. Burton BN, Canales C, Lucero J, Cannesson M: Personalized surgical transfusion risk prediction: Comment. Anesthesiology 2023; 138:116–7
2. Zapf MAC, Freundlich RE, Wanderer JP: Personalized surgical transfusion risk prediction: Comment. Anesthesiology 2023; 138:117–8
3. Lou SS, Liu H, Lu C, Wildes TS, Hall BL, Kannampallil T: Personalized surgical transfusion risk prediction using machine learning to guide preoperative type and screen orders. Anesthesiology 2022; 137:55–66
4. Vyas DA, Eisenstein LG, Jones DS: Hidden in plain sight - Reconsidering the use of race correction in clinical algorithms. N Engl J Med 2020; 383: 874–82
5. Obermeyer Z, Powers B, Vogeli C, Mullainathan S: Dissecting racial bias in an algorithm used to manage the health of populations. Science 2019; 366:447–53
6. Churchwell K, Elkind MSV, Benjamin RM, Carson AP, Chang EK, Lawrence W, Mills A, Odom TM, Rodriguez CJ, Rodriguez F, Sanchez E, Sharrief AZ, Sims M, Williams O; American Heart Association: Call to action: Structural racism as a fundamental driver of health disparities: A presidential advisory from the American Heart Association. Circulation 2020; 142:e454–68
7. Wiens J, Saria S, Sendak M, Ghassemi M, Liu VX, Doshi-Velez F, Jung K, Heller K, Kale D, Saeed M, Ossorio PN, Thadaney-Israni S, Goldenberg A: Do no harm: A roadmap for responsible machine learning for health care. Nat Med 2019; 25:1337–40
8. Chen IY, Pierson E, Rose S, Joshi S, Ferryman K, Ghassemi M: Ethical machine learning in healthcare. Annu Rev Biomed Data Sci 2021; 4:123–44
9. The Editors: Striving for diversity in research studies. N Engl J Med 2021; 385:1429–30.

(Accepted for publication September 27, 2022. Published online first on November 8, 2022.)