Machine learning risk prediction of mortality for patients undergoing surgery with perioperative SARS-CoV-2: the COVIDSurg mortality score

COVIDSurg Collaborative Co-authors*†

*Correspondence to: Mr Aneel Bhangu, NIHR Global Health Research Unit on Global Surgery, Heritage Building, University of Birmingham, Mindelsohn Way, Birmingham B15 2TH, UK (e-mail: A.A.Bhangu@bham.ac.uk)

†Collaborating authors are listed in the supplement

Introduction

Since the beginning of the COVID-19 pandemic tens of millions of operations have been cancelled1 as a result of excessive postoperative pulmonary complications (51.2 per cent) and mortality rates (23.8 per cent) in patients with perioperative SARS-CoV-2 infection2. There is an urgent need to restart surgery safely in order to minimize the impact of untreated non-communicable disease.

As rates of SARS-CoV-2 infection in elective surgery patients range from 1–9 per cent3–8, vaccination is expected to take years to implement globally9 and preoperative screening is likely to lead to increasing numbers of SARS-CoV-2-positive patients, perioperative SARS-CoV-2 infection will remain a challenge for the foreseeable future.

In order to inform consent and shared decision making, a robust, globally applicable score is needed to predict individualized mortality risk for patients with perioperative SARS-CoV-2 infection. The authors aimed to develop and validate a machine learning-based risk score to predict postoperative mortality risk in patients with perioperative SARS-CoV-2 infection.

Methods

Cohort study design

An international, multicentre, prospective, cohort study (COVIDSurg Cohort Study) included consecutive patients who were diagnosed with SARS-CoV-2 in the 7 days before or the 30 days after surgery. Patients undergoing any type of surgery were included from 1 February 2020 to 31 July 2020. Full study methodology has been published previously2.

Machine learning methods were used to analyse the COVIDSurg Cohort Study dataset. Sixteen patient and operative variables were entered into the analysis and, based on time, data were split into a derivation set, where all analysis and modelling were performed, and a validation set, where the final model was assessed. Five features were selected following variable importance measurements from both linear and non-linear modelling. These features were then combined into 26 different predictor sets, which were fitted through three different algorithms (logistic regression, decision trees and random forest), generating a total of 78 different models. These models were tuned with a 10-fold cross-validation, fitted in the 75 per cent split of the derivation set and assessed in the remaining 25 per cent. In order to ascertain the model’s stability, this training and testing split was randomly repeated 100 times (bootstraps). Finally, to decide which model to select, performance was evaluated through the mean area under the receiver operating characteristic curve (AUROC) value. A full description of methods is available in Appendix S1.

Results

Patients

From 10,029 patients in the cohort study, data from 8492 patients entered the machine learning processes (Fig. 1). The most common surgical procedures were abdominal (40.6 per cent, 3446 of 8492 patients), orthopaedic (33.8 per cent, 2867 of 8492) and head and neck surgery (9.8 per cent, 835 of 8492). Emergency surgery accounted for 80.8 per cent (6862 of 8492 patients) and benign disease for 57.3 per cent (4868 of 8492) (Table S1).

Of the 8492 patients, data from 6777 patients who had surgery from February to May 2020 (726 hospitals, 65 countries) were entered into the derivation set and data from 1715 who had surgery from June to July 2020 (296 hospitals, 53 countries) were entered into the validation set. Despite changes in patient demographics and treatment options, such as the recommendations from the RECOVERY trial10 in June 2020, model performance remained robust between the derivation and validation sets. Mortality rates were 17.2 per cent overall (1462 of 8492 patients), 18.9 per cent (1279 of 6777) in the derivation set and 10.7 per cent (183 of 1715) in the validation set. Missing data handling is described in Appendix S2.
Feature selection

A full breakdown of feature and model selection are in Tables S3–S7. Random forest and elastic net were applied to the derivation set and the most important features were ranked (Fig. S1). Logit model performance of all different feature combination (runs) divided by each of the features is shown in Fig. 2a. The five features taken forward into model building were age, preoperative respiratory support, ASA grade, specialty and revised cardiac risk index (RCRI) score.
Model building and validation

Multivariable models were built using all possible combinations of the five selected features, resulting in 26 model runs (Table S8). The smallest, higher-performing and most robust model was a logistic regression model with four features (age, ASA grade, RCRI score and preoperative respiratory support). A full description is in Appendix S2.

The AUROC for the final score was 0.73 (95 per cent c.i. 0.71 to 0.74) in the derivation set and 0.80 (95 per cent c.i. 0.77 to 0.83) in the validation set (Fig. 2b). Calibration in the validation set is shown in Fig. S2 (Brier score 0.084, intercept 0.078, slope 1.08).

Mortality rates were studied for different probability risk cut-offs (Fig. S3), with a balance between distribution and clinical importance leading to cut-off values at approximately 2, 10, 30 and 45 per cent mortality. These showed high concordance between derivation and validation sets (Supplementary Table S7) with the final mortality percentages in the combined dataset being very similar.

Deployment

The COVIDSurg Mortality Score has been deployed online at http://covidsurgrisk.app/.

Discussion

Understanding risks related to SARS-CoV-2 and COVID-19 will remain important for the foreseeable future. There is an urgent need to restart elective surgery in order to reduce the backlog of cancelled operations11. Vaccination programmes are expected to take years to deliver, particularly in low- and middle-income countries9, and may not protect against new variants12. COVID-19-free surgical pathways are unlikely to be universally implemented, particularly in emergency settings8, leaving patients at...
risk of in-hospital SARS-CoV-2 transmission. Research is critical to mitigate impact. The COVIDSurg Morality Score is a tool that can help overcome these challenges and has clear clinical applications. It can inform consent and shared decision making by patients and surgeons. If individual patient mortality risk is too high, they may choose to postpone surgery. If surgery is unavoidable, risks should be clearly documented as part of informed consent, and postoperative critical care can be identified and justified. It can also identify patients who are at very low risk, where surgery can continue as planned.

The COVIDSurg Mortality Score shows high discrimination for predicting mortality in patients with perioperative SARS-CoV-2. The AUROCs in the derivation (0.73) and validation (0.80) sets are consistent with published non-surgical COVID risk scores. The COVIDSurg Mortality Score is globally applicable and consists of variables that are readily available across all resource settings. The score was derived and validated in a global data set (all types of surgery, 756 hospitals, 69 countries) and demonstrated validity across geographic regions and over time as community SARS-CoV-2 infection rates fluctuated.

The COVIDSurg cohort study was prospectively collected and conducted, but it had limitations. Variables were coded following standardized medical criteria but a further analysis of continuous variables as well as the inclusion of interactions could be incorporated in future studies. As there are no other international, multispecialty data sets for SARS-CoV-2 surgical patients, further external validation was not possible, but should be considered if such data sets emerge. High mortality percentages (approximately 47 per cent) at higher probability outputs (greater than 0.5) should be interpreted with caution, as numbers of patients with these feature combinations were low. The authors acknowledge that the model will warrant updates as more data become available.

Despite differences in mortality between the derivation and validation sets, model performance remained robust. There are likely to have been structural and system differences between hospitals that were not captured in the cohort study. These differences are further compounded by variation in community SARS-CoV-2 across time and across geographical locations. Nonetheless, the score performed well in the validation set, suggesting the score can be used reliably both during COVID waves and troughs.

**CovidSurg Collaborative Co-authors (PubMed Citable)**

**Writing group:** Laura Bravo* (UK), Dmitri Nepogodiev* (UK), James C Glasby* (UK), Elizabeth Li* (UK), Joana FF Simoes* (Portugal), Sivesh K Kamarajah (UK), Maria Picciochi (Portugal), Tom EF Abbott (UK), Adesoji O Ademuyiwa (Nigeria), Alexis P Arnaud (France), Arnav Agarwal (Canada), Amanpreet Brar (Canada), Mohammad Khosravi (Iran), Masoud Mozaafari (Iran), Ahmed Adil (Iraq), Helen M. Mohan (Ireland), Oded Zmori (Israel), Marco Fiore (Italy), Gaetano Gallo (Italy), Francesco Pata (Italy), Gianluca Pellino (Italy), Soheil Sato (Japan), Faris Ayasra (Jordan), Mohammad Chaaar (Jordan), Ildar R Fakhradiyev (Kazakhstan), Mohammad Jamal (Kuwait), Muhammed Elhadi (Libya), Aiste Guilla (Lithuania), April Rosiani (Malaysia), Laura Martinez (Mexico), Antonio Ramos De La Medina (Mexico), Oumaima Outani (Morocco), Pascal Jonker (Netherlands), Schelto Kruifje (Netherlands), Milou Noltes (Netherlands), Pieter Stein-kamp (Netherlands), Willemijn van der Plas (Netherlands), Adesoji Ademuyiwa (Nigeria), Babatunde Osinaike (Nigeria), Justin Seyi-olajide (Nigeria), Emmanuel Williams (Nigeria), Sofija Pejkova (North Macedonia), Knut Magne Augestad (Norway), Kjetil Soreide (Norway), Zainab Al Balushi (Oman), Ahmad Qureshi (Pakistan), Raza Sayyed (Pakistan), Mustafa Abu Mohsen (Palestine), Sadi Abukhalaf (Palestine), Moises Cukier (Panama), Hugo Gomez (Paraguay), Sebastian Shu (Peru), Ximena Vasquez (Peru), Marie Dionne Parreno-Sacadan (Philippines), Piotr Major (Poland), Jose Azevedo (Portugal), Miguel Cunha (Portugal), Irene Santos (Portugal), Ahmad Zarour (Qatar), Eduard-Alexandru Bonci (Romania), Ionut Negoi (Romania), Sergey Efetov (Russia), Andrey Litvin (Russia), Faustin Ntirenganya (Rwanda), Ehab Al-Ameer (Saudi Arabia), Dejan Radenkovic (Serbia), Frederick Koh Hong Xiang (Singapore), Chew Min Hoe (Singapore), James Ng Chi Yong (Singapore), Rachel Moore (South Africa), Ncamsile Nhlabathi (South Africa), Ruth Blanco Colino (Spain), Ana Minaya Bravo (Spain), Ana Minaya-Bravo (Spain), Umesh Jayarajah (Sri Lanka), Dakshitha Wickramasinghe (Sri Lanka), Mohammed Elmuftija (Sudan), William Jebril (Sweden), Martin Rutegard (Sweden), Malin Sund (Sweden), Arda Isik (Turkey), Sezai Leventoglu (Turkey), Tom EF Abbott (UK), Ruth Benson (UK), Ed Caruna (UK), Sohini Chakrabortee (UK), James M. Keatley, Samuel Lawday, Elizabeth Li, Harvinder Mann, Ella J Marson, Kenneth A Mclean, Maria Picciochi, Elliott H. Taylor, Abhinav Tiwari, Joana F.F. Simoes, Isobel M. Trout, Mary L. Venn, Richard J.W. Wilkin, Aneel Bhangu*, Dmitri Nepogodiev*

*Co-chairs

**Dissemination Committee:** Irida Dajti (Albania), Arben Gjata (Albania), Luis Boccalatte (Argentina), Maria Marta Modolo (Argentina), Daniel Cox (Australia), Peter Pockney (Australia), Phillip Townend (Australia), Felix Aigner (Austria), Irmgard Kronberger (Austria), Kamral Hossain (Bangladesh), Gabriele vanRamshorst (Belgium), Ismail Lawani (Benin), Gustavo Ataide (Brazil), Glauco Baiocchi (Brazil), Igor Buarque (Brazil), Muhammad Gohar (Bulgaria), Mihail Slavchev (Bulgaria), Arnav Agarwal (Canada), Amranpreet Brar (Canada), Janet Martin (Canada), Maria Marta Modolo (Chile), Maricarmen Olivos (Chile), Jose Calvache (Colombia), Carlos Jose Perez Rivera (Colombia), Ana Danic Hadzibegovic (Croatia), Tomislav Kopjar (Croatia), Jakov Mihanovic (Croatia), Jaroslaw Klat (Czech Republic), René Novyseidlak (Czech Republic), Peter Christensen (Denmark), Ala El-Hussuna (Denmark), Sylvia Batista (Dominican Republic), Eddy Lincango (Ecuador), Sameh H Emile (Egypt), Mengistu Gebreyohanes Mengesha (Ethiopia), Samuel Hailu (Ethiopia), Hailu Tamiru (Ethiopia), Joonas Kaupulla (Finland), Johanna Laukkari (Finland), Alexis Arnaud (France), Markus Albertsmeiers (Germany), Hans Lederhuber (Germany), Markus Loffler (Germany), Stephen Tabiri (Ghana), Symeon Metallidis (Greece), Georgios Tsoulfas (Greece), Maria Aguilera Lorena (Guatemala), Gustavo Grecinos (Guatemala), Tamas Mersich (Hungary), Daniel Wettstein (Hungary), Dhruv Ghosh (India), Gabriele Kembuan (Indonesia), Peiman Brouk (Iran), Mohammad Khosravi (Iran), Masoud Mozaafari (Iran), Ahmed Adil (Iraq), Helen M. Mohan (Ireland), Oded Zmori (Israel), Marco Fiore (Italy), Gaetano Gallo (Italy), Francesco Pata (Italy), Gianluca Pellino (Italy), Soheil Sato (Japan), Faris Ayasra (Jordan), Mohammad Chaaar (Jordan), Ildar R Fakhradiyev (Kazakhstan), Mohammad Jamal (Kuwait), Muhammed Elhadi (Libya), Aiste Guilla (Lithuania), April Rosiani (Malaysia), Laura Martinez (Mexico), Antonio Ramos De La Medina (Mexico), Oumaima Outani (Morocco), Pascal Jonker (Netherlands), Schelto Kruifje (Netherlands), Milou Noltes (Netherlands), Pieter Stein-kamp (Netherlands), Willemijn van der Plas (Netherlands), Adesoji Ademuyiwa (Nigeria), Babatunde Osinaike (Nigeria), Justin Seyi-olajide (Nigeria), Emmanuel Williams (Nigeria), Sofija Pejkova (North Macedonia), Knut Magne Augestad (Norway), Kjetil Soreide (Norway), Zainab Al Balushi (Oman), Ahmad Qureshi (Pakistan), Raza Sayyed (Pakistan), Mustafa Abu Mohsen (Palestine), Sadi Abukhalaf (Palestine), Moises Cukier (Panama), Hugo Gomez (Paraguay), Sebastian Shu (Peru), Ximena Vasquez (Peru), Marie Dionne Parreno-Sacadan (Philippines), Piotr Major (Poland), Jose Azevedo (Portugal), Miguel Cunha (Portugal), Irene Santos (Portugal), Ahmad Zarour (Qatar), Eduard-Alexandru Bonci (Romania), Ionut Negoi (Romania), Sergey Efetov (Russia), Andrey Litvin (Russia), Faustin Ntirenganya (Rwanda), Ehab Al-Ameer (Saudi Arabia), Dejan Radenkovic (Serbia), Frederick Koh Hong Xiang (Singapore), Chew Min Hoe (Singapore), James Ng Chi Yong (Singapore), Rachel Moore (South Africa), Ncamsile Nhlabathi (South Africa), Ruth Blanco Colino (Spain), Ana Minaya Bravo (Spain), Ana Minaya-Bravo (Spain), Umesh Jayarajah (Sri Lanka), Dakshitha Wickramasinghe (Sri Lanka), Mohammed Elmuftija (Sudan), William Jebril (Sweden), Martin Rutegard (Sweden), Malin Sund (Sweden), Arda Isik (Turkey), Sezai Leventoglu (Turkey), Tom EF Abbott (UK), Ruth Benson (UK), Ed Caruna (UK), Sohini Chakrabortee (UK),
University of Rome, Rome, Italy); A. Sagnotta*, L. Marino Cosentino, S. Mancini (Ospedale San Filippo Neri, Rome, Italy); G. Lisi*, D. Spoletini (Sant’Eugenio Hospital, Rome, Italy); V. Bellato, M. Campanelli, G. Sica*, L. Siragusa, (Policlinico Tor Vergata Hospital, Rome, Italy); L. Bonavina*, E. Asti, D. Bernardi, A. Lovece (IRCSS Policlinico San Donato, University of Milan, Milan, Italy); T. Perra, A. Porcu*, A. Fancellu, C. F. Feo, A. M. Scano (Cliniche San Pietro, A.O.U. Sassari, Sassari, Italy); F. Tuminello, R. Galleano*, A. Franceschi, A. Langone (San Paolo, Italy); F. Fleres*, A. Spolini, P. Bordoni, M. Franzini, G. Clarizia, A. Grechi, A. Longhini (Ospedale Di Sondrio (Asst Valtellina E Alto Lario), Sondrio, Italy); E. Guaitoli, G. Manca (Pernosti Hospital Brindisi, Italy); U. Grossi, S. Novello, G. Zanus*, M. Romano, S. Rossi (Ca’ Foncello Treviso—DISCOG, Università di Padova, Padua, Italy); G. Pirollozzi*, A. Recordare (Dell’Angelo Hospital, ULS53 Serenissima, Venezia, Italy); S. Faiella, G. Turri, S. Rattizzato, T. Campagnaro*, A. Guglielmi, C. Pedrazzani, A. Ruzzenente, E. Polletto, S. Conci, L. Casetti, M. Fontana, R. Salvia*, G. Malieo, A. Esposito, L. Landoni, M. De Pastena, C. Bassi, M. Tuveri, S. Noble, M. Marchegiani, L. Bortolasi* (Azienda Ospedaliera Universitaria Integrata Di Verona, Verona, Italy); F. Ferrara (San Carlo Borromeo Hospital, ASST Santi Paolo e Carlo, Italy); M. La Torre (Fabia Mater Hospital, Rome, Italy); E. Sambaguro*, M. Malavolta*, G. Moreto, H. Impellizzeri, M. Inama* (Ospedale Pederzoli, Peschiera del Garda, Italy); G. Barugola, F. Ascarì, G. Ruffo (IRCCS Ospedale Sacro Cuore Don Calabria, Negar di Valpolicella, Verona, Italy); S. Granieri*, C. Cotsoglou (ASST Vimercate, Vimercate, Italy); M. Berselli*, M. Desio, V. Marchionini, E. Cocozza (ASST Settiglioni, Varese, Italy); S. Di Saverio*, G. Jetto, D. Iovino, G. Carcano (University of Insubria, Ospedale di Circolo e Fondazione Macchi, Varese, Italy); F. Ayara*, A. Qasem, Y. Ayara (Al-Basheer Hospital, Amman, Jordan); M. Al-Marsi*, M. K. Abou Chaar, H. Al-Najjar, K. Ghandour, F. Alawneh, R. Abdel Jalil, S. Abdel Al, M. Elayyan, R. Ghanem, I. Lataifeh, O. Alsalairhe (King Hussein Cancer Centre, Amman, Jordan); F. J. Abu Za’nomeen, T. Fahmawee, A. Ibrahim, K. Obeidat* (King Abdullah University Hospital, Ar-Ramtha, Jordan); K. J. Lee (Keimyung University School of Medicine, Daegu, Korea); S. J. Shin, H. Chung (Keimyung University Hospital, Daegu, Korea); I. Albader, J. Alabbad, M. A. R. Bachiri, L. C. Groen (Noordwest Ziekenhuisgroep, Alkmaar, Netherlands); T. Verhagen*, F. M. ter Brugge (Ziekenhuisgroep Twente, Almelo, Netherlands); J. C. G. Scheijmans, M. A. Boermesteer*, R. Hompes*, E. M. Meima-van Praag, S. Sharabany (Amsterdam UMC, Amsterdam, Netherlands); A. B. J. Borgstein, S. G. Giibertz*, M. I. van Berge Henegouwen* (Amsterdam UMC VUM, Amsterdam, Netherlands); S. Gans*, P. Van Duijvendijk, T. HerkLOTS, T. De Hoop, R. M. De Graaff, D. Sloothak, M. Bolster van Enneman, J. Baaij, J. H. G. Klinkenberg (Gelre ziekenhuis, Apeldoorn, Netherlands); R. Van Eekeren, E. J. P. Spillenlaar Blijen (Rijnstate, Arnhem, Netherlands); S. W. Van der Burg, N. J. Harlaar*, F. H. W. Jonker (Rode Kruis Ziekenhuis, Beverwijk, Netherlands); M. Vermaas, K. R. Voigt (Jisselland Ziekenhuis, Capelle aan den IJssel, Netherlands); D. Nellensteijn*, E. A. B. Bensi (Curacao Medical Center, Willemstad, Curaçao, Netherlands); L. Posma-Bouman* (Slingeland Ziekenhuis, Doetinchem, Netherlands); M. Van Sambeek*, M. Holscher (Catharina Ziekenhuis, Eindhoven, Netherlands); W. T. Van den Broek (St. Anna Ziekenhuis, Geldrop, Netherlands); S. H. Kluijf*, J. P. P. M. De Vries, P. J. Steinkamp, K. C. Jonker, W. Y. Van der Plas, W. Bierman, Y. Janssen (University Medical Centre Groningen, Netherlands); J. Franken, S. Oosterling* (Spaarne Gasthuis, Zoetermeer, Netherlands); G. Boerma*, D. Schweitzer, M. H. F. Keulen, S. Ketting (Zuyderland MC, Sittard/Heerlen, Netherlands); J. A. Wegdam, T. S. De Vries Rellingh, E. Schipper, P. H. E. Teween (Eikeleriek Ziekenhuis, Helmond, Netherlands); E. R. Hendriks, A. A. W. Van Geloven (Tergooi Hospital, Hilversum, Netherlands); M. Emous*, R. Poelstra, M. Teunissen (Medisch Centrum Leeuwarden, Leeuwarden, Netherlands); S. L. Gerritsen, D. Boerma, D. Stenstra (Antonius Ziekenhuis, Nieuwegein, Netherlands); P. R. De Reuver*, F. Thunnissen, B. A. M. Vermeulen, A. Groen (Dradouw University Medical Center Medisch Centrum, Nijmegen, Netherlands); T. M. Van Ginhoven*, C. L. Vietor, M. J. W. van der Oest (Erasmus Medisch Centrum, Rotterdam, Netherlands); P. W. H. E. Vriens*, T. Houwen, J. Heisterkamp (Elisabeth Tweesteden Ziekenhuis, Tilburg, Netherlands); A. S. Van Petersen, W. Van der Meij, C. T. Stevens (Bovenhoven, Uden, Netherlands); A. Pronk, W. J. Bakker (Diakonessenhuis, Utrecht, Netherlands); M. C. Richir*, M. R. Vriens, M. D. Filipe (University Medical Center Utrecht, Netherlands);
USA); A. Abbas, A. Soliman, B. Kim, C. Jones, E. Dauer, E. Renza-Stingone, E. Hernandez, E. Gokcen, E. Kropf, H. Sufrin, H. Hirsch, H. Ross, J. Engel, J. Seward, J. Diaz, J. Poggio, K. Sanserino, L. Rae, M. Philp, M. Metro, P. McNelis, R. Petrov, S. Rehman, T. Pazions (Temple University Hospital, Philadelphia, USA); B. Till, R. Lamm, A. J. Rios-Diaz, F. Palazzo* (Thomas Jefferson University Hospital, Philadelphia, USA); M. Rosengart*, K. Nicholson (University of Pittsburgh Medical Center, Pittsburgh, USA); M. M. Carrick*, K. Rodkey (Medical City Plano, Plano, USA); A. Suri, R. Callcut* (UC Davis Medical Center, Sacramento, USA); S. Nicholson*, N. Talathoti (UT Health San Antonio (UTHSA), University Hospital (UHS), San Antonio, USA); D. Klaristenfeld* (Kaiser Permanente San Diego Medical Center, USA); Biffl W*, Marsh C, Schaffer K (Scripps Memorial Hospital, San Diego, USA); A. E. Berndtson*, S. Averbach, T. Curry (University of California, San Diego, USA); R. Kwan-Feinberg*, E. Consorti, R. Gonzalez, R. Grolman, T. Liu, O. Merzlikin (Alta Bates Summit Medical Center (Sutter Health), Oakland, USA); M. K. Abel, D. Ozgediz, M. Boeck, L. Z. Kornblith*, B. Nunez-Garcia, B. Robinson, P. Park (University of California San Francisco (UCSF), San Francisco, USA); A. F. Utria, S. E. Rice-Townsend*, P. Javid, J. Hauptman, K. Kieran (Seattle Children’s Hospital, Seattle, USA); D. Nehra*, A. Walters, J. Cuschieri, G. H. Davidson (Harborview Medical Center, Seattle, USA); J. Nunez*, R. Cosker (University of Utah Healthcare, Salt Lake City, USA); S. Eckhouse* (Washington University School of Medicine, St Louis, USA); A. Choudhry, W. Marx* (SUNY Upstate University Hospital, Syracuse, USA); T. Jamil*, S. Seeger, S. Al-Embiedeen (ProMedica Toledo Hospital, Toledo, USA); M. Quintana*, H. Jackson (The George Washington University Hospital, Washington, D.C., USA); S. D. Wexner, I. Kent (Cleveland Clinic Florida, Weston, USA); F. N. Martins* (University of Massachusetts, UMass Memorial Hospital, Worcester, USA); M. Alshehri*, H. Al-Naggar*, M. Alsayadi, 

**PROBAST: Prediction model Risk Of Bias A5essment Tool**

**Step 1: Specify your systematic review question(s) – not applicable.**

**Step 2: Classify the type of prediction model evaluation – development and validation.**

**Step 3: Assess risk of bias and applicability**

<table>
<thead>
<tr>
<th>Risk of Bias</th>
<th>Participants</th>
<th>Predictors</th>
<th>Outcome</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Were appropriate data sources used, e.g., cohort, RCT, or nested case–control study data?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were all inclusions and exclusions of participants appropriate?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were predictors defined and assessed in a similar way for all participants?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were predictor assessments made without knowledge of outcome data?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are all predictors available at the time the model is intended to be used?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was the outcome determined appropriately?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was a prespecified or standard outcome definition used?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were predictors excluded from the outcome definition?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was the outcome defined and determined in a similar way for all participants?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was the outcome determined without knowledge of predictor information?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was the time interval between predictor assessment and outcome determination appropriate?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were there a reasonable number of participants with the outcome?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were continuous and categorical predictors handled appropriately?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were all enrolled participants included in the analysis?</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were participants with missing data handled appropriately?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was selection of predictors based on univariable analysis avoided?</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were relevant model performance measures evaluated appropriately?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Were model overfitting, underfitting, and optimism in model performance accounted for?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do predictors and their assigned weights in the final model correspond to the results from the reported multivariable analysis?</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

ROB = risk of bias. + indicates low ROB/low concern regarding applicability. – indicates high ROB/high concern regarding applicability, and ? indicates unclear ROB/unclear concern regarding applicability.

**Step 4: Overall judgment**

<table>
<thead>
<tr>
<th>ROB</th>
<th>Applicability</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>NA</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Funding
This report was funded by a National Institute for Health Research (NIHR) Global Health Research Unit Grant (NIHR 16.136.79), Association of Coloproctology of Great Britain and Ireland, Bowel & Cancer Research, Bowel Research UK, Association of Upper Gastrointestinal Surgeons, British Association of Surgical Oncology, British Gynaecological Cancer Society, European Society of Coloproctology, Medtronic, NIHR Academy, The Urology Foundation, Sarcoma UK, Vascular Society for Great Britain and Ireland, Yorkshire Cancer Research, and the MRC Health Data Research UK (HDRUK/CFC/01), an initiative funded by UK Research and Innovation, Department of Health and Social Care (England) and the devolved administrations, and leading medical research charities. L.B.M. and G.V.G. also acknowledge the Wellcome Trust 4-year studentship programme in mechanisms of inflammatory disease (MIDAS, 215182/Z/19/Z). The funders had no role in study design, data collection, analysis and interpretation, or writing of this report. The views expressed are those of the authors and not necessarily those of the National Health Service, the NIHR, or the UK Department of Health and Social Care.

Acknowledgements

Supplementary material
Supplementary material is available at BJS online.

References