**Appendix A** Types of prediction models included in the revision according to TRIPOD Statement and their validations

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Authors of the model**  **(year)** | **Types of prediction models by TRIPOD statement1** | **Validation sample size** | **AUC2** | **Calibration3** |
| M. Jacob et al.  (2020) (26) | 1a |  |  |  |
| Chen W et al.  (2020) (23) | 1a |  |  |  |
| Wu Z et al.  (2020) (22) | 2a | - | - | - |
| Chen et al.  019) (21) | 3 | (without CEA level) :216 | Chen et al.: 0.847 (C-index) | 0.928 |
| (with CEA level) :216 | Chen et al.: 0.848 (C-index) | 0.866 |
| Wang et al.  (2018) (29) | 1a | - | - | - |
| She et al.  (2017) (24) | 2a | - | - | - |
| Yang et al.  (2017) (28) | 3 | 344 | Yang et al.: 0.784 | - |
| Swensen et al.:0.649 |
| Gould et al.: 0.599 |
| Van Gómez López et al. (2015) (16) | 1a | - | - | - |
| Zheng et al.  (2015) (20) | 2a | - | - | - |
| Zhang et al.  (2015) (25) | 3 | 120 | Zhang et al.: 0.910 | - |
| Zhang et al. (without CYFRA 21-1): 0.812 |
| Swensen et al.: 0.752 |
| Gould et al.: 0.730 |
| Li et al.: 0.833 |
| Li et al.  (2012) (19) | 3 | 145 | Li et al.: 0.874 | - |
| Swensen et al.:0.784 |
| Gould et al.: 0.754 |

(Continued)

**Appendix A** (Continued)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Authors of the model**  **(year)** | **Types of prediction models by TRIPOD statement1** | **Validation sample size** | **AUC2** | **Calibration3** |
| Dong et al.  (2013) (15) | 2a | - | - | - |
| Yonemori et al.  (2007) (18) | 3 | 148 | Yonemori et al.:0.840 | - |
| Gould et al.  (2007) (27) | 1b | - | - | - |
| Swensen et al.  (1997) (17) | 2a | - | - | - |

**Abbreviations**: CEA, carcinoembryonic antigen; CYFRA 21-1, Cytokeratin 19-fragment marker; TRIPOD, Transparent Reporting of a multivariable Prediction model for Individual Prognosis Or Diagnosis; AUC, the Area Under the Curve.

**Notes**: 1Types of prediction models by TRIPOD statement: **Type 1** corresponds to the construction of a predictive model without performing internal / external validation techniques (in **type 1a** the performance of the model is evaluated in the same group in which the model is developed, and therefore, it is an apparent performance ; in **type 1b**, resampling (for example bootstrapping or cross-validation techniques) is used to evaluate the performance and optimism of the model). **Type 2** corresponds to the construction of predictive models and performance an internal validation. In this type of model, the sample of participants is divided into two groups: one to develop the model and the other to validate it; if this division of the participants is carried out randomly, we are dealing with a **type 2a** study and if it is carried out in a non-randomized way (for example, by location or time), it is a **type 2b** study. According to TRIPOD, type 2b can also be considered as a development model halfway between internal validation and external validation. **Type 3** corresponds to the construction of a predictive model and the performance of an external validation in a separate population (that is, the model is developed in a group of participants and validated in a separate group other than the one used to create the model). 2AUC = **Discrimination** is the ability of the model to assign, on pairs of randomly selected subjects, one from the group with malignancy and the other without it, the correct result (greater probability in the group with malignancy). For binary results, the area under the ROC curve (AUC), or C statistic, is the most frequently used discrimination measure. 3 **Calibration** is a measure that expresses the agreement between the observed results and the model predictions. The most common calibration measurements are the calibration slope and the Hosmer-Lemeshow test.

**Appendix B** External validation of the included models by different authors from those who created the models

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Authors**  **(year)** | **Types of prediction models by TRIPOD statement1** | **Study population** | **Prevalence of malignancy** | **Prevalence of**  **active or former smokers** | **Validation sample size** | **Prediction model** | **AUC2** | **Calibration3** |
| Li et al.  (2020) (40) | 4 | Chinese patients with SPN who had surgery and had a clear pathological diagnosis. | 85.6% | NR | 496 | Swensen et al. | 0.62 | - |
| Gould et al. | 0.62 |
| Li et al. | 0.63 |
| Cui et al.  (2019) (39) | 4 | Chinese patients with newly detected SPN on CT scans that had been confirmed based on histopathological result or remained stable for at least 2 years. | 72.4% | NR | 277 | Swensen et al. | 0.77 | **-** |
| Gould et al. | 0.66 |
| Yang B et al.  (2018) (38) | 4 | Patients with biopsy-proven pulmonary lung nodules at a single centre in Korea. | 77.2% | 39% | 242 | Swensen et al. | 0.61 | - |
| Gould et al. | 0.60 |
| Hammer M et al.  (2017) (30) | 4 | USA patients with large nodules at high risk for lung cancer (non-screening cohort). | 69% | 84% | 86 | Swensen et al. | 0.62 | - |
| Gould et al. | 0.59 |
| Li et al. | 0.53 |
| Soardi G et al.  (2016) (44) | 4 | Italian patients with SPN with a definitive diagnosis by | 54.5% | 52% | 200 | Swensen et al. | 0.60 | - |

(Continued)

**Appendix B** (*Continued*)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Authors**  **(year)** | **Types of prediction models by TRIPOD statement11** | **Study population** | **Prevalence of malignancy** | **Prevalence of**  **active or former smokers** | **Validation sample size** | **Prediction model** | **AUC2** | **Calibration3** |
|  |  | biopsy or stability/volume reduction. |  |  |  |  |  |  |
| Talwar A et al.  (2016) (31) | 4 | Non-screening USA population with PNs (incidental finding on a chest CT and patients with either a known or prior cancer within the last 5 years scanned either as a staging or follow-up scan or scanned for another reason). | 46% | 65% | 702 | Swensen et al. | 0.58 | This model underestimated the probability of malignancy in 1st 2nd and 3th quintiles. It overestimated the probability of malignancy in 4th and 5th quintiles. |
| Gould et al. | 0.62 | This model underestimated the probability of malignancy in 1st 2nd and 3th quintiles. It overestimated the probability of malignancy in 4th and 5th quintiles. |
| Perandini S et al.  (2015) (46) | 4 | Patients with a newly discovered solid SPN with a definitive diagnosis by biopsy or by means of serial volume assessment found in electronic medical records. | 54.7% | NR | 285 | Swensen et al. | 0.77 | - |
| Li et al. | 0.88 |
| Al-Almeri A et al.  (2015) (41) | 4 | UK patients with PN detected in routine clinical practice. | 40.6% | 76.2% | 244 | Swensen et al. | 0.89 | - |
| Gould et al. | 0.73 |

(Continued)

**Appendix B** (*Continued*)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Authors**  **(year)** | **Types of prediction models by TRIPOD statement11** | **Study population** | **Prevalence of malignancy** | **Prevalence of**  **active or former smokers** | **Validation sample size** | **Prediction model** | **AUC2** | **Calibration3** |
| Tanner et al.  (2015) (32) | 4 | USA patients with a PN with a definitive diagnosis (identified by querying databases). | 25% | 73% | 377 | Swensen et al. | 0.77 | - |
| Gould et al. | 0.74 |
| Perandini et al.  (2014) (45) | 4 | Patients with SPN found in CT images from the hospital medical records. | 54.8% | 11.8% | 288 | Swensen et al. | 0.76 | - |
| Zhang X et al.  (2014) (37) | 4 | Patients with SPN who underwent surgical resection, mainly from the south of China. | 81.2% | 31,8% | 154 | Swensen et al. | 0.75 | This model underestimated the probability of malignancy. |
| Gould et al. | 0.72 | This model underestimated the probability of malignancy. |
| Li et al. | 0.80 | This model underestimated the probability of malignancy in all quintiles except the 4th quintile. |
| Shinohara S et al.  (2014) (36) | 4 | Japanese patients with SPN who underwent surgical resection. | 84.2% | 61.8% | 241 | Swensen et al. | 0.67 | - |
| Xiao F et al.  (2013) (35) | 4 | Patients with SPN confirmed by CT who underwent surgical procedure in China-Japan. | 72.8% | 37.4% | 107 | Swensen et al. | 0.78 | - |
| Gould et al. | 0.68 |
| Li et al. | 0.81 |
| Melo CB et al.  (2012) (42) | 4 | Brazilian patients submitted to resection of SPN in Brazil. | NR | NR | 110 | Swensen et al. | 0.79 | - |
| Gould et al. | 0.69 |

(Continued)

**Appendix B** (*Continued*)

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Authors**  **(year)** | **Types of prediction models by TRIPOD statement1** | **Study population** | **Prevalence of malignancy** | **Prevalence of**  **active or former smokers** | **Validation sample size** | **Prediction model** | **AUC2** | **Calibration3** |
| Isbell et al.  (2011) (33) | 4 | USA patients who underwent resection of a PN. | 73% | 74% | 189 | Swensen et al. | 0.79 | This model underestimated the probability of malignancy in 4th and 5th quintiles, the other quintiles were well calibrated. |
| Schultz E et al.  (2008) (34) | 4 | USA patients with SPN discovered incidentally in chest CT, who underwent surgery. | 45% | 89% | 151 | Swensen et al. | 0.80 | This model underestimated the probability of malignancy in all quintiles except the 1st.. |
| Gould et al. | 0.73 | This model overestimated the probability of malignancy. |
| Herder et al. (2005) (43) | 4 | Dutch patients with an indeterminate SPN, which had been detected during normal clinical work ,who had been referred for FDG-PET-CT. | 57.5% | 74.5% | 106 | Swensen et al. | 0.79 | This model tended to underestimate the probability of malignancy, particularly at lower probabilities. |

**Abbreviations**: SPN, Solitary Pulmonary Nodule; PN, Pulmonary Nodule; UK, United Kingdom; USA, United States of America; CT, Computed Tomography; FDG-PET, F-fluorodeoxyglucose-Positron Emission Tomography/Computed Tomography; NR, Not Reported; TRIPOD, Transparent Reporting of a multivariable Prediction model for Individual Prognosis Or Diagnosis ; AUC, the Area Under the Curve.

**Notes**: 1Types of prediction models by TRIPOD statement:**Type 4** corresponds to an external validation, in a separate population, of a previously constructed model (a new model is not created, only an external validation of an existing model is performed). 2AUC = **Discrimination** is the ability of the model to assign, on pairs of randomly selected subjects, one from the group with malignancy and the other without it, the correct result (greater probability in the group with malignancy). For binary results, the area under the ROC curve (AUC), or C statistic, is the most frequently used discrimination measure. 3 **Calibration** is a measure that expresses the agreement between the observed results and the model predictions. The most common calibration measurements are the calibration slope and the Hosmer-Lemeshow test.

**Appendix C** Mathematical equations of the included models

**M. JACOB ET AL. MODEL**

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = −8.61 + (1.56 × **Age Category**) + (1.87 × **Gender**) + (1.26 × **Smoking Status**) + (1.28 × **Current ExtraPulmonary Cancer**) + (1.27 × **Air Bronchogram**) + (0.22 × **Nodule Size**).

where e is the base of natural logarithms, while x is the regression coefﬁcient in the logistic regression; The rest is not reported.

In this study, the assigned weight of one predictor in the ﬁnal model do not correspond to the result from the reported multivariable analysis: the weight assigned to “Current ExtraPulmonary Cancer” variable is 1.28, to obtain this value the corresponding OR would be 3.59 (on the other hand, in the multivariate analysis table the OR = 8.94), therefore the assigned weight should be **2.19**.

**CHEN W ET AL. MODEL**

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = -2.957 - (0.004 x **lung nodule density**) + (1.096 x **vascular penetration sign**) + (1.198 x **nodule type**) - (1.811 x **incisure**).

where e is the base of natural logarithms, while x is the regression coefﬁcient in the logistic regression; lung nodule density indicates the lung nodule density in Hounsfield Units; vascular penetration sign (1 = if vascular penetration sign is present, 0 = otherwise), nodule type (1 = mixed Ground Glass Nodule (GGN), solid or pure GGN= 0); incisure (1 = if incisure surrounding nodules are present, otherwise = 0).

In this study, the assigned weights of the predictors in the ﬁnal model do not correspond to the results from the reported multivariable analysis: the weight assigned to “Nodule density” variable is 0.004, to obtain this value the corresponding OR would be 1.004, (on the other hand, in the multivariate analysis table the OR = 0.995), therefore the assigned weight should be **-0.005**; the weight assigned to “vascular penetration sign” variable is 1.096, to obtain this value the corresponding OR would be 2.99, (on the other hand, in the multivariate analysis table the OR = 3.49), therefore the assigned weight should be **1.25**; the weight assigned to “Nodule Type” variable is 1.198, to obtain this value the corresponding OR would be 2.99, (on the other hand, in the multivariate analysis table the OR = 4.27), therefore the assigned weight should be **1.451**; the weight assigned to “incisure” variable is 1.811, to obtain this value the corresponding OR would be 6.11, (on the other hand, in the multivariate analysis table the OR = 0.179), therefore the assigned weight should be **-1.720**.

**WU Z ET AL. MODEL**

The **Wu Z et al. model** is defined by the equations:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = − 2.8107 + (1.2454 × **smoking history**) + (− 1.4055 × **edge**) + (0.077 1 × **age**) + (− 1.0 728 × **gender**) + (− 0 .6228 × **clear border**) + (− 1.3319  ×  **calcification**) + (0.6 890  ×  **drinking** **history**).

where e is the base of natural logarithms, while x is the regression coefﬁcient in the logistic regression; age indicates the patient’s age in years; gender (1 = if the patient is male, 0 = otherwise); smoking history (1 = current or former smoker, otherwise = 0); calcification (1 = if calciﬁcation is present in the SPN, otherwise = 0); drinking history (1= if the patient has history of drinking, 0 = otherwise); edge (1= if smooth edge is present in the SPN, otherwise = 0); clear border (1= if a clear border is present in the SPN, otherwise = 0).

**CHEN ET AL. MODEL**

The **Chen et al. model** is defined by the equations:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = − 4.439 + (0.034 × **age**) + (1.347 × **marginal spiculation**) + (2.210 × **signiﬁcant enhancement**) + (1.211× **pleural indentation**).

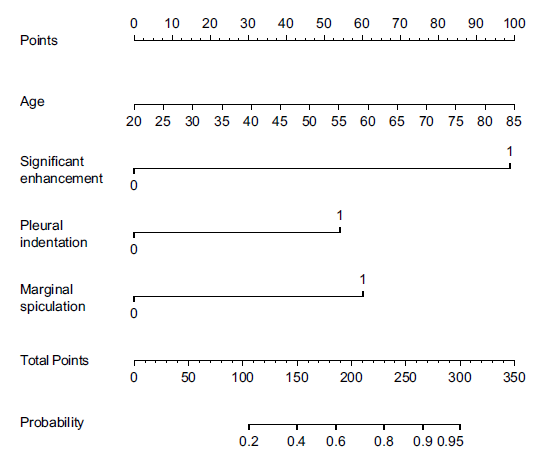
where e is the base of the natural logarithms, while x is the regression coefﬁcient in the logistic regression; age indicates the patient’s age in years; marginal spiculation (1 = if spiculated appearance is present in the SPN, otherwise = 0); signiﬁcant enhancement (1= if significant enhancement is present, otherwise = 0); pleural indentation (1 = if pleural indentation is present, otherwise = 0).

They described an alternative model adding CEA level:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = − 4.688 + (0.036× **age**) + (1.396 × **marginal spiculation**) + (2.174 × **signiﬁcant enhancement**) + (1.235 × **pleural indentation**) + (1.572 × **CEA**).

where e is the base of the natural logarithms, while x is the regression coefﬁcient in the logistic regression; age indicates the patient’s age in years; marginal spiculation (1 = if spiculated appearance is present in the SPN, otherwise = 0); signiﬁcant enhancement (1= if significant enhancement is present, otherwise = 0); pleural indentation (1 = if pleural indentation is present, otherwise = 0); CEA = serum CEA level (ng/mL).

This model is also expressed as a normogram as follows:



To obtain the nomogram-predicted probability, locate the patient values at each axis, and draw a vertical line to the “Points” axis to determine the number of points attributed to each variable value, determine total number of points for all variables, and locate the sum on the “Total Points” line to assess the individual probability of lung cancer.

**WANG ET AL. MODEL**

* Pre-test probability of a malignant SPN = ex/(1+ex)
* 𝑥 = − 7.363 + (0.079 × **age**) + (1.90 × **lobulation**) + (1.024 × **vascular convergence**) + (1.530 × **pleural retraction**) + (0.359 × **SUVmax**),

where 𝑒 is the base of the natural logarithms; while x is the regression coefﬁcient in the logistic regression; age indicates the patient’s age in years; lobulation (1 = if lobulation is present in the SPN, otherwise = 0); vascular convergence (1 = if vascular convergence is present in the SPN, otherwise = 0); pleural retraction (1 = if pleural retraction is present in the SPN, otherwise = 0); SUVmax is the maximum uptake value on the PET.

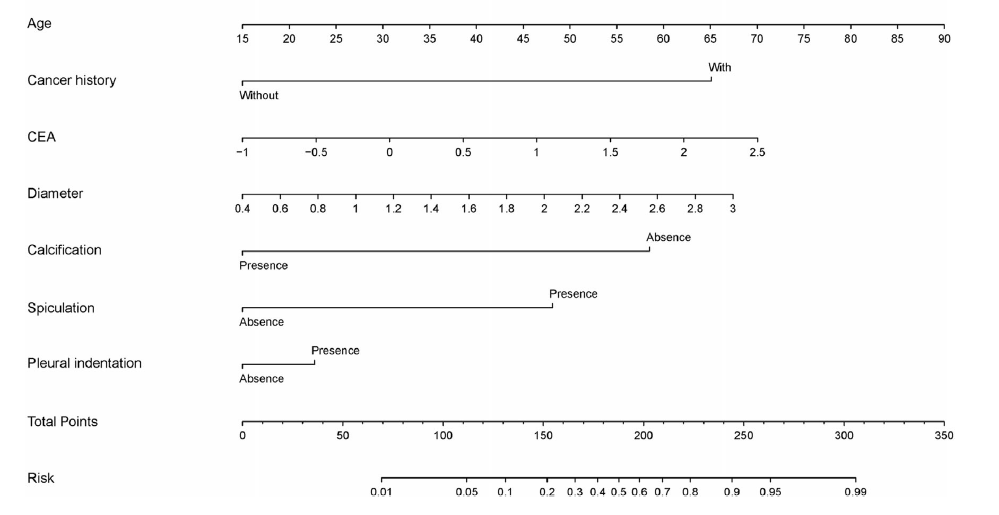
**SHE ET AL. MODEL**

The **She et al. model** is defined by the equations:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = - 5.4175 + (0.8149 × **Log10 CEA**) + (1.0447 × **diameter**) + (2.5978 × **cancer history**) + (0.0518 × **age**) + (1.7166 × **spiculation**) + (0.3986 × **pleural indentation**) - (2.2549 × **calcification**).

where e is the base of the natural logarithms, while x is the regression coefﬁcient in the logistic regression; Log10CEA = the log base 10 transformations of serum CEA value (ng/mL); the diameter of the SPN is measured in millimetres (mm); cancer history (1= if cancer history is present, otherwise = 0); age indicates the patient’s age in years; spiculation (1 = if spiculated appearance is present in the SPN, otherwise = 0); pleural indentation (1 = if pleural indentation is present, otherwise = 0); calcification (1 = if calciﬁcation is present in the SPN, otherwise = 0).

In this study, the assigned weights of three predictors in the ﬁnal model do not correspond to the results from the reported multivariable analysis: the weight assigned to “cancer history” variable is 2.5978, to obtain this value the corresponding OR would be 13.43, (on the other hand, in the multivariate analysis table the OR = 12.82), therefore the assigned weight should be **2.55**; the weight assigned to the “CEA” variable in the model is 0.8149, to obtain this value the corresponding OR would be 2,251 (on the other hand, in the multivariate analysis table the OR = 1.09), therefore, the weight should be **0.086**; the weight assigned to the diameter variable in the model is 1.0447, to obtain this value the corresponding OR would be 2.84 (on the other hand, in the multivariate analysis table the OR = 1.11) therefore, the weight should be **0.095**.

This model is also expressed as a normogram as follows:

To obtain the nomogram-predicted probability, locate the patient values at each axis, and draw a vertical line to the “Points” axis to determine the number of points attributed to each variable value, determine total number of points for all variables, and locate the sum on the “Total Points” line to assess the individual probability of lung cancer, CEA: Log10 (the serum carcinoembryonic antigen value).

**YANG ET AL. MODEL**

The **Yang et al. model** is defined by the equations:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = − 6.173 + (1.207 × **gender**) + (0.580 ×**age**) + (0.520 × **pack-years**) – (0.226 × **previous extrathoracic disease**) – (0.685× **previous chronic lung disease except cancer**) + (2.739 × **malignancy history**) + (0.933 × **diameter**) + (0.702 × **lobulation**) + (0.466 × **spiculation**) + (21.060 × **lobulation and spiculation**) – (1.428 × **irregular edges**) – (2.062 × **calcification**).

where e is the base of the natural logarithms, while x is the regression coefﬁcient in the logistic regression; gender (1 = if the patient is male, 2 = if the patient is female); age indicates the patient’s age in years; pack-years indicates the pack-years of smoking; Previous extrathoracic disease indicates the previous medical history (1 = if it is present, 0 = if it is not present); Previous chronic lung disease except cancer (1= if a previous chronic lung disease except cancer is present, if not = 0); malignancy history ( 1 = if malignancy history is present, otherwise = 0); the diameter of the SPN is measured in millimetres (mm); lobulation (1 = if lobulation is present in the SPN, otherwise = 0); spiculation (1 = if spiculated appearance is present in the SPN, otherwise = 0); lobulation and spiculation ( 1= if both are present in the SPN, otherwise = 0); irregular edges (1= if irregular edges are present in the SPN, otherwise = 0); calcification (1 = if calciﬁcation is present in the SPN, otherwise = 0).

We cannot confirm if the predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis because Yang et al. did not report the univariate and multivariate tables in their study.

**VAN GÓMEZ LÓPEZ ET AL. MODEL**

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = − 3.767+ (0.489× **SUVmax**) + (0.052× **age**)

where e is the base of the natural logarithms, while x is the regression coefﬁcient in the logistic regression; age indicates the patient’s age in years; SUVmax is the maximum uptake value on the PET.

**ZHENG ET AL. MODEL**

The **Zheng et al. model** is defined by two equations:

1. The clinical prediction model for malignancy in SPNs with less than 50% ground glass opacity (GGO) is defined by the equations:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = - 7.442 + (0.051 × **age**) + (0.711 × **presence of symptoms**) + (0.066 × **total protein** **concentration**) + (0.032 × **diameter**) + (1.071 × **lobulation**) **-** (1.220 × **calciﬁed nodes**).

where e is the base of the natural logarithms, while x is the regression coefﬁcient in the logistic regression; age indicates the patient’s age in years; presence of symptoms (1= if the patient has related symptoms\*, 0 = if the patient has not related symptoms); total protein concentration = serum total protein concentration (g/L); the diameter of the SPN is measured in millimetres (mm); lobulation (1 = if lobulation is present in the SPN, otherwise = 0); calciﬁed nodes (1 = if calciﬁcation is present in the SPN, otherwise = 0).

\*these symptoms are described in Zheng et al. article as: “lung cancer-related symptoms, such as cough, shortness of breath, hemoptysis, chest pain, fever, and so forth”.

1. The clinical prediction model for malignancy in SPNs with 50% or greater ground glass opacity (GGO) is defined by the equations:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = - 6.192 **+** (0.924 × **sex**) + (0.042 × **FEV1%)** + (0.131 × **diameter**) **-** (2.424 × **calciﬁed** **nodes**).

where e is the base of the natural logarithms, while x is the regression coefﬁcient in the logistic regression; sex (1 = if the patient is male, 2= if the patient is female); FEV1 = Forced expiratory volume 1 (%); the diameter of the SPN is measured in millimetres (mm); calciﬁed nodes (1 = if calciﬁcation is present in the SPN, otherwise = 0).

In both models, an error in the symbol that precedes two variables has been corrected (marked in bold), in coherence with the Odds ratios reported in the multivariate analysis.

**ZHANG ET AL. MODEL**

The **Zhang et al. model** is defined by the equations:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = −14.417 + (0.111 × **age**) + (1.009 × **smoking history**) + (2.597 × **nodule diameter**) + (1.056 × **spiculation**) + (−1.258 × **clear border**) + (1.184 × **CYFRA 21-1**).

where e is the base of the natural logarithms, while x is the regression coefﬁcient in the logistic regression; age indicates the patient’s age in years; history of smoking (1 = current or former smoker, otherwise = 0) nodule diameter refers to the maximum nodule diameter measured by chest radiography or CT prior to surgery (in cm); spiculation (1 = if spiculated appearance is present in the SPN, otherwise = 0); clear border (1 = if clear border is present in the SPN, otherwise = 0); CYFRA 21-1 = serum CYFRA 21-1 level (ng/mL).

**DONG ET AL. MODEL**

The **Dong et al. model** is defined by the equations:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = − 4.294 + (0.035 × **age**) + (0.221 × **CEA**) + (0.200 × **CYFRA 21-1**) + (1.029 × **smoking**) + (0.974 × **family history of cancer**) + (0.633 × **diameter**) + (−1.631 × **clear border**) + (−1.923 × **satellite lesions**) + (2.673 × **lobulation**) + (−3.295 × **calciﬁcation**) + (2.027 × **spiculation**).

where e is the base of the natural logarithms, while x is the regression coefﬁcient in the logistic regression; age indicates the patient’s age in years; the diameter of the SPN is measured in centimetres (cm); CEA = serum CEA level (ng/mL); CYFRA 21-1 = serum CYFRA 21-1 level (ng/mL); Smoking indicates smoking history (1= if smoking history is present, otherwise = 0); family cancer history (1 = if family cancer history is present, otherwise = 0); clear border (1 = if clear border is present in the SPN, otherwise = 0); satellite lesions (1 = if satellite lesions are present in the SPN, otherwise = 0); lobulation (1 = if lobulation is present in the SPN, otherwise = 0); calciﬁcation (1 = if calciﬁcation is present in the SPN, otherwise = 0); spiculation (1 = if spiculated appearance is present in the SPN, otherwise = 0).

**LI ET AL. MODEL**

The **Li et al. model** is defined by the equations:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = - 4.496 + (0.07 × **age**) + (0.676 × **diameter**) + (0.736 × **spiculation**) + (1.267 × **family history of** **cancer**) - (1.615 × **calcification**) – (1.408 × **border**).

where e is the natural logarithm, while x is the regression coefﬁcient in the logistic regression; age indicates the patient’s age in years; diameter indicates the largest nodule measurement (in cm) reported on initial chest radiograph or CT scan; spiculation (1 = if spiculated appearance is present, otherwise = 0); family cancer history (1 = if family history of cancer exist, otherwise = 0); calciﬁcation (1 = if calcification is present, otherwise = 0); border (1= if clear border is present, otherwise = 0).

**YONEMORI ET AL. MODEL**

The **Yonemori et al. model** is defined by the equations:

* Pre-test probability of a benign SPN = ex/(1+ex)
* x = 3.7009 + (3.0705 × **calciﬁcation**) + (-1.3243 × **CT bronchus sign**) + (- 5.3399 × **spiculation**) + (-1.16 × **√CEA**) + (-1.4987 × **CRP**).

where e is the base of the natural logarithm, while x is the regression coefﬁcient in the logistic regression; calciﬁcation (1 = if calciﬁcation is present in the SPN, otherwise = 0); CT bronchus sign (1 = if CT bronchus sign is present, otherwise = 0); spiculation (1 = if spiculated appearance is present, otherwise = 0); CEA = serum CEA level (ng/mL); CRP = serum CRP level (mg/L).

They constructed an alternative model in case the biological parameters were not available:

* Pre-test probability of a benign SPN = ex/(1+ex)
* x = 1.084 + (2.7851 × **calciﬁcation**) + (-1.1795 × **CT bronchus sign**) + (-5.4481 × **spiculation**).

where e is the base of the natural logarithm, while x is the regression coefﬁcient in the logistic regression; calciﬁcation (1 = if calciﬁcation is present in the SPN, otherwise = 0); CT bronchus sign (1 = if CT bronchus sign is present, otherwise = 0); spiculation (1 = if spiculated appearance is present, otherwise = 0).

In this case, the study of Yonemori et al. has created a model which predicts the probability of benignancy instead of malignancy of the SPN. To obtain the probability of malignancy you have to calculate the complementary probability (1- probability of benignancy).

We cannot confirm if the predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis because Yonemori et al. did not report the univariate and multivariate tables in their study.

**GOULD ET AL. MODEL**

The **Gould et al. model (also called VA model)** is defined by the equations:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* X = - 8.404+ (2.061× **smoke**) + (0.779 × **age10**) + (0.112 × **diameter**) - (0.567 × **yearsquit10**).

where e is the base of the natural logarithm, while x is the regression coefﬁcient in the logistic regression; smoke indicates smoking history (1 = current or former smoker, 0 = never smoker); age10 indicates age in years at the time of nodule identification, divided by 10; diameter indicates the largest nodule measurement (in mm) reported on initial chest radiograph or CT scan; yearsquit10 indicates the number of years since quitting smoking, divided by 10 (0 indicates not applicable).

VA model calculator is available on: <https://magarray.com/calculator-va/>

**SWENSEN ET AL. MODEL**

The **Swensen et al. model (also called The Mayo Clinic model)** is defined by the equations:

* Pre-test probability of a malignant SPN = ex/(1+ex)
* x = - 6.8272 + (0.0391× **age**) + (0.7917 × **smoke**) +(1.3388 × **cancer**) + (0.1274 × **diameter**) + (1.0407 × **spiculation**) + (0.7838 × **upper lobe**).

where e is the base of the natural logarithm, while x is the regression coefﬁcient in the logistic regression; age indicates the patient’s age in years; smoke indicates smoking history (1 = current or former smoker, 0 = never smoker); cancer indicates history of an extrathoracic cancer >5 years before nodule identification (1 = yes, otherwise = 0); diameter indicates the largest nodule measurement (in mm) reported on initial chest radiograph or CT scan; spiculation (1 = if spiculated appearance is present, otherwise = 0); and upper lobe is location of the nodule within the upper lobe (1 = yes, 0 = no).

Mayo Clinic model calculator is available on: <https://reference.medscape.com/calculator/solitary-pulmonary-nodule-risk> .

**Appendix D** Items included in each domain of PROBAST quality

PROBAST **M. JACOB ET AL. (2020).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **No.**

**1.2** Were all inclusions and exclusions of participants appropriate? **No.**

**Applicability: No.**

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes.**

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **No**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes.**

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability**: **No.**

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **No** (10.6 EPV).

**4.2** Were continuous and categorical predictors handled appropriately? **No**.

**4.3** Were all enrolled participants included in the analysis? **Yes**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only)? **Yes**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **No**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **No**.

PROBAST **CHEN W ET AL. (2020).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **No**.

**1.2** Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability: No.**

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **No**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability**: **No.**

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **Yes** (40 EPV).

**4.2** Were continuous and categorical predictors handled appropriately? **Yes**.

**4.3** Were all enrolled participants included in the analysis? **Yes**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only)? **Unclear**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **No**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **No**.

PROBAST **WU Z ET AL. (2020).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **Unclear**.

**1.2** Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability: No.**

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **No**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability**: **No**.

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **Unclear**.

**4.2** Were continuous and categorical predictors handled appropriately? **No**.

**4.3** Were all enrolled participants included in the analysis? **No**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only)? **Unclear**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **Yes**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **Unclear**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **Yes**.

PROBAST **CHEN ET AL. (2019).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **No**.

**1.2** Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability**: **No**.

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **No**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability**: **No**.

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **No** (88 participants with the outcome).

**4.2** Were continuous and categorical predictors handled appropriately? **No**.

**4.3** Were all enrolled participants included in the analysis? **Yes**.

**4.4** Were participants with missing data handled appropriately? **Yes**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only) **Yes**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **Yes**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **Yes**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **Yes**.

PROBAST **WANG ET AL. (2018).**

**1. PARTICIPANTS**

1.1 Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **Yes**.

1.2 Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability: No.**

**2. PREDICTORS**

2.1 Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

2.2 Were predictor assessments made without knowledge of outcome data? **Unclear**.

2.3 Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability: Unclear.**

**3. OUTCOME**

3.1 Was the outcome determined appropriately? **Yes**.

3.2 Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

3.3 Were predictors excluded from the outcome deﬁnition? **Yes**.

3.4 Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

3.5 Was the outcome determined without knowledge of predictor information? **Unclear**.

3.6 Was the time interval between predictor assessment and outcome determination appropriate? **No**.

**Applicability: No.**

**4. ANALYSIS**

4.1 Were there a reasonable number of participants with the outcome? **Yes** (23.8 EPV).

4.2 Were continuous and categorical predictors handled appropriately? **Yes**.

4.3 Were all enrolled participants included in the analysis? **No**.

4.4 Were participants with missing data handled appropriately? **Unclear**.

4.5 Was selection of predictors based on univariable analysis avoided? (Model development studies only) **Yes**.

4.6 Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

4.7 Were relevant model performance measures evaluated appropriately? **No**.

4.8 Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

4.9 Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **Unclear**.

PROBAST **SHE ET AL. (2017).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **No**.

**1.2** Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability**: **No**.

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **Unclear**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability**: **No**.

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **Yes** (86.42 EPV).

**4.2** Were continuous and categorical predictors handled appropriately? **Yes**.

**4.3** Were all enrolled participants included in the analysis? **Yes**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only) **No**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **Yes**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **Yes**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **No**.

PROBAST **YANG ET AL. (2017).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **Unclear**.

**1.2** Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability**: **No**.

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **Unclear**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability**: **No**.

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **Yes** (236 participants with the outcome).

**4.2** Were continuous and categorical predictors handled appropriately? **No**.

**4.3** Were all enrolled participants included in the analysis? **Yes**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only) **Unclear**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **Unclear.**

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **Unclear**.

PROBAST **VAN GÓMEZ LÓPEZ ET AL. (2015).**

**1. PARTICIPANTS**

1.1 Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **No**.

1.2 Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability: No.**

**2. PREDICTORS**

2.1 Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

2.2 Were predictor assessments made without knowledge of outcome data? **Unclear**.

2.3 Are all predictors available at the time the model is intended to be used **Yes**.

**Applicability: Unclear.**

**3. OUTCOME**

3.1 Was the outcome determined appropriately? **No**.

3.2 Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

3.3 Were predictors excluded from the outcome deﬁnition? **Yes**.

3.4 Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

3.5 Was the outcome determined without knowledge of predictor information? **Unclear**.

3.6 Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability: No.**

**4. ANALYSIS**

4.1 Were there a reasonable number of participants with the outcome? **Yes** (20 EPV).

4.2 Were continuous and categorical predictors handled appropriately? **Yes**.

4.3 Were all enrolled participants included in the analysis? **Yes**.

4.4 Were participants with missing data handled appropriately? **Unclear**.

4.5 Was selection of predictors based on univariable analysis avoided? (Model development studies only) **Unclear.**

4.6 Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

4.7 Were relevant model performance measures evaluated appropriately? **No**.

4.8 Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

4.9 Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **Yes**.

PROBAST **ZHENG ET AL. (2015).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **Unclear**.

**1.2** Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability**: **No**.

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **Unclear**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability**: **No**.

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **Unclear**.

**4.2** Were continuous and categorical predictors handled appropriately? **No**.

**4.3** Were all enrolled participants included in the analysis? **Unclear**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only) **Yes**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **Unclear**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **No.**

PROBAST **ZHANG ET AL. (2015).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **No**.

**1.2** Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability**: **No**.

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **Unclear**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability**: **No**.

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **No** (72 participants with the outcome).

**4.2** Were continuous and categorical predictors handled appropriately? **Yes**.

**4.3** Were all enrolled participants included in the analysis? **Yes**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only) **No**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **Unclear**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **Yes**.

PROBAST **DONG ET AL. (2013).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **No**.

**1.2** Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability**: **No**.

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **No**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability**: **No**.

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **Yes** (117.81 EPV).

**4.2** Were continuous and categorical predictors handled appropriately? **Yes**.

**4.3** Were all enrolled participants included in the analysis? **Unclear**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only) **Yes**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **Unclear**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **Yes**.

PROBAST **LI ET AL. (2012).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **No**.

**1.2** Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability**: **No**.

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **No**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability**: **No**.

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **No** (98 participants with the outcome).

**4.2** Were continuous and categorical predictors handled appropriately? **Yes**.

**4.3** Were all enrolled participants included in the analysis? **Yes**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only) **Unclear**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **Unclear**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **Yes**.

PROBAST **YONEMORI ET AL. (2007).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **No**.

**1.2** Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability**: **No**.

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Yes**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **No**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **Unclear**.

**Applicability**: **No**.

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **Yes** (131.72 EPV).

**4.2** Were continuous and categorical predictors handled appropriately? **Yes**.

**4.3** Were all enrolled participants included in the analysis? **No**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only) **No**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **Unclear**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **Unclear**.

PROBAST **GOULD ET AL. (2007).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **Yes**.

**1.2** Were all inclusions and exclusions of participants appropriate? **No**.

**Applicability: No.**

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **No**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability: Unclear.**

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **Unclear**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **No**.

**Applicability**: **No**.

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **Yes** (51 EPV).

**4.2** Were continuous and categorical predictors handled appropriately? **Unclear**.

**4.3** Were all enrolled participants included in the analysis? **Yes**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only) **Yes**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **Yes**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **Yes**.

PROBAST **SWENSEN ET AL. (1996).**

**1. PARTICIPANTS**

**1.1** Were appropriate data sources used, e.g., cohort, randomized controlled trial, or nested case–control study data? **Yes**.

**1.2** Were all inclusions and exclusions of participants appropriate? **Yes**.

**Applicability**: **Yes**.

**2. PREDICTORS**

**2.1** Were predictors deﬁned and assessed in a similar way for all participants? **Yes**.

**2.2** Were predictor assessments made without knowledge of outcome data? **Unclear**.

**2.3** Are all predictors available at the time the model is intended to be used? **Yes**.

**Applicability**: **Unclear**.

**3. OUTCOME**

**3.1** Was the outcome determined appropriately? **Yes**.

**3.2** Was a prespeciﬁed or standard outcome deﬁnition used? **Yes**.

**3.3** Were predictors excluded from the outcome deﬁnition? **Yes**.

**3.4** Was the outcome deﬁned and determined in a similar way for all participants? **Yes**.

**3.5** Was the outcome determined without knowledge of predictor information? **Unclear**.

**3.6** Was the time interval between predictor assessment and outcome determination appropriate? **No**.

**Applicability**: **No**.

**4. ANALYSIS**

**4.1** Were there a reasonable number of participants with the outcome? **Unclear**.

**4.2** Were continuous and categorical predictors handled appropriately? **Yes.**

**4.3** Were all enrolled participants included in the analysis? **Yes**.

**4.4** Were participants with missing data handled appropriately? **Unclear**.

**4.5** Was selection of predictors based on univariable analysis avoided? (Model development studies only) **Yes**.

**4.6** Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? **Yes**.

**4.7** Were relevant model performance measures evaluated appropriately? **Yes**.

**4.8** Were model overﬁtting and optimism in model performance accounted for? (Model development studies only) **No**.

**4.9** Do predictors and their assigned weights in the ﬁnal model correspond to the results from the reported multivariable analysis? (Model development studies only) **Yes**.

According to PROBAST, the answers to each item of the domains should be answered as follows: "Probably yes", "Probably no", "Yes", "No", "No information" or "Unclear". We, by consensus, have answered the questions avoiding the first two answers, given the difficult definition of them objectively, therefore, the categories have been "Yes", "No" and "Unclear".