|  |  |  |
| --- | --- | --- |
| Section/Topic | # | Checklist Item |
| **TITLE** |  |  |
| Nature of study | 1 | Identify the report as introducing a predictive model |
| **ABSTRACT** |  |  |
| Structured summary | 2 | Background |
|  | 3 | Objectives |
|  | 4 | Data sources |
|  | 5 | Performance metrics of the predictive model or models, in point estimates |
|  | 6 | Conclusion, including the practical value of the developed predictive model or models |
| **INTRODUCTION** |  |  |
| Rationale | 7 | Identify the clinical goal |
|  | 8 | Review the current practice of any existing models |
|  | 9 | Review the prediction accuracy of any existing models |
| Objectives | 10 | State the nature of study being predictive modeling |
|  | 11 | Define the target of prediction |
|  | 12 | Identify how the prediction problem may benefit the clinical goal |
| **METHODS** |  |  |
| Describe the setting | 13 | Identify the clinical setting for the target predictive model |
| Define the prediction problem | 14 | Define a measurement for the prediction goal |
|  | 15 | Defining quality metrics for prediction models |
|  | 16 | Define the success criteria for prediction (e.g. based on metrics in internal validation or external validation in the context of the clinical problem) |
| Prepare data for model building | 17 | Identify relevant data sources |
|  | 18 | States that relevant data sources were approved by ethics committee or Institutional Review Board |
|  | 19 | States the inclusion criteria for data |
|  | 20 | States the exclusion criteria for data |
|  | 21 | Describe the time span of data |
|  | 22 | Describe the sample or cohort size |
|  | 23 | Define the observational units on which the response variable is defined |
|  | 24 | Define the observational units on which the predictor variable(s) are defined |
|  | 25 | Define the predictor variables. |
| Data | 26 | Describe the feature selection |
|  | 27 | Using the normalization techniques |
|  | 28-30 | Explain the model validation (0/1) If they did train and test on internal data-->add 1  If they used external validation--> add 2 |
|  | 31 | If inhouse dataset was used |
|  | 32 | If they used a method to resolve imbalance |
|  | 33 | Determine a set of candidate modeling techniques (eg. logistic regression, random forest, or deep learning). |
|  | 34 | Define the performance metrics to select the best model |
|  | 35 | Specify the model selection strategy. |
| **RESULTS** |  |  |
| Report the final model and performance-- | 36 | Report the predictive performance of the final model in terms of the performance metrics specified in the methods section |
| **DISCUSSION** |  |  |
|  | 37 | Interpretation of the final model |
|  | 37 | Comparison with other models in the literature |
| Limitations of the model | 39 | Sufficient data available for a good fit of the model. |
|  | 40 | Assumed variances in data format: For example, input data format (e.g. inter-scanner variability, sample size, difference in sequences used) or output data format |
|  | 41 | Potential bias of the data used in modeling |
|  | 42 | Generalizability of the data |
|  | 43 | Exact number of patients in each grade |
| **TOTAL SCORE** |  |  |