| Question | Drysdale, 2017 | Harris, 2022 | Hopman, 2021 | Kong, 2021 | Moreno-Ortega, 2019 | Pei, 2020 | Schultz, 2018 | Sun, 2020 | Tian, 2020 | van Waarde, 2015 | Wu, 2022 | Zhutovsky, 2019 | Zhutovsky, 2021 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1.1 Were appropriate data sources used, e.g., cohort, RCT, or nested case–control study data? | PN | Y | Y | PN | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| 1.2 Were all inclusions and exclusions of participants appropriate? | Y | Y | N | N | N | N | Y | Y | N | Y | Y | Y | Y |
| Final rating domain 1 (risk of bias) | low | low | high | high | high | high | low | low | high | low | low | low | low |
| 2.1 Were predictors defined and assessed in a similar way for all participants? | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| 2.2 Were predictor assessments made without knowledge of outcome data? | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| 2.3 Are all predictors available at the time the model is intended to be used? | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| Final rating domain 2 (risk of bias) | low | low | low | low | low | low | low | low | low | low | low | low | low |
| 3.1 Was the outcome determined appropriately? | PY | Y | Y | NI | NI | NI | NI | NI | NI | NI | NI | PY | Y |
| 3.2 Was a prespecified or standard outcome definition used? | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| 3.3 Were predictors excluded from the outcome definition? | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| 3.4 Was the outcome defined and determined in a similar way for all participants? | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| 3.5 Was the outcome determined without knowledge of predictor information? | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| 3.6 Was the time interval between predictor assessment and outcome determination appropriate? | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| Final rating domain 3 (risk of bias) | low | low | low | low | low | low | low | low | low | low | low | low | low |
| 4.1 Were there a reasonable number of participants with the outcome? | N | N | N | N | N | N | N | N | N | N | N | N | N |
| 4.2 Were continuous and categorical predictors handled appropriately? | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| 4.3 Were all enrolled participants included in the analysis? | NI | N | N | N | Y | N | Y | Y | N | Y | N | N | N |
| 4.4 Were participants with missing data handled appropriately? | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI |
| 4.5 Was selection of predictors based on univariable analysis avoided? | N | N | N | Y | N | Y | Y | N | Y | Y | N | N | Y |
| 4.6 Were complexities in the data (e.g., censoring, competing risks, sampling of control participants) accounted for appropriately? | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| 4.7 Were relevant model performance measures evaluated appropriately? | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI |
| 4.8 Were model overfitting and optimism in model performance accounted for? | Y | Y | N | Y | N | Y | Y | Y | NI | N | Y | Y | Y |
| 4.9 Do predictors and their assigned weights in the final model correspond to the results from the reported multivariable analysis? (development studies only) | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI | NI |
| Final rating domain 4 (risk of bias) | high | high | high | high | high | high | high | high | high | high | high | high | high |
| Final rating total (risk of bias) | high | high | high | high | high | high | high | high | high | high | high | high | high |