Positive SLNB

Contents

1 Model development

The primary goal is to create a model that predicts recurrence using the available variables except the variable with additional positive lymph nodes. The set of variables presented below was considered of importance for the following analyses:

- `age`
- `sex`
- `simpleloc` (simple tumor location)
- `clark` level
- `histology_simple`
- `breslow` thickness
- `AJCC_sub_SLNB_8th`
- `ulceration`
- `SN_tumor_burden`
- `no_removed_SNs`
- `no_removed_nonSNs` (not used for the first set of analyses)
- `no_pos_SNs`
- `no_pos_SNs_cat`

1.1 Models for recurrence

We performed a backwards selection of to come up with the final model that included `ulceration`, `age`, `breslow` and `SN_tumor_burden` (Table 1). Logarithmic transformations of the continuous covariates -i.e. `age`, `breslow` and `SN_tumor_burden`- adequately represented their effects.

We used the variables in the final model to predict 5-year recurrence (Table 2).

The final 5-year recurrence model had c-index of 0.68 (95 percent c.i. 0.65 to 0.7).

We assess calibration of the final model using a leave-one-center-out cross validation approach. The prediction model is built on 8 centers and calibration is evaluated on the 9th. That is performed recursively, each time

Table 1: Final model for recurrence

	Hazard ratio	Lower 0.95	Upper 0.95
age	1.28	1.12	1.45
SN_tumor_burden	1.59	1.39	1.81
breslow	1.41	1.23	1.61
ulceration - present:absent	1.41	1.16	1.73

Table 2: Final model for 5-year recurrence

	Hazard ratio	Lower 0.95	Upper 0.95
age	1.28	1.12	1.45
SN_tumor_burden	1.59	1.39	1.81
breslow	1.41	1.23	1.61
ulceration - present:absent	1.41	1.16	1.73

leaving a different center out for validation. We separate multiple imputation in the training set and the test set to avoid using information of missingness in the training centers to the test center.

In general, we see quite adequate performance across centers, where confidence intervals include the diagonal. However, in smaller centers such as the one in Groningen there is substantial underestimation of risk.

1.2 Distant metastasis

For the assessment of distant metastasis we considered a calibrated version of the 5-year recurrence model. The association between distant metastasis and was of the same size (calibration slope of 1.01, 95 percent c.i. 0.87 to 1.16). We compare the performance of the considered model to that of multivariable Cox regression model including all 9 covariates of interest (`sex`, `ulceration`, `no_removed_SNs`, `no_pos_SNs`, `simpleloc`, `histology_simple`, `breslow`, `SN_tumor_burden` and `age`). The full model had a c-index of 0.7 (95 percent c.i. 0.68 to 0.73) while the calibrated model had a c-index of 0.7 (95 percent c.i. 0.67 to 0.72).

In terms of calibation, we first assess a leave-one-center-out cross validation, where the baseline hazard is estimated from the training set of 8 centers based on the cox regression model for 5-year recurrence and the calibration slope for the linear predictor is derived from a cox model predicting risk of distant metastasis form the linear predictor of the previous model (Figure 2).

1.3 Overall mortality

For the assessment 5-year overall mortality we considered a calibrated version of the 5-year recurrence model. We compare the performance of the considered model to that of multivariable Cox regression model including all 9 covariates of interest. The association between recurrence and overall mortality was not different (calibration slope 1.04, 95 percent c.i. 0.88 to 1.20). The full model had a c-index of 0.7 (95 percent c.i. 0.68 to 0.73) while the calibrated model had a c-index of 0.7 (95 percent c.i. 0.67 to 0.73). We assess calibation as previously using leave-one-center-out cross validation (Figure 3).

1.4 Nomogram

We developed a 4-item score assigning points to each prognostic factor based on the magnitude of their association with recurrence. The nomogram to calculate the score is given in 4. The calibrated results for 5-year distant recurrence and 5-year mortality can be found in Figure 5.

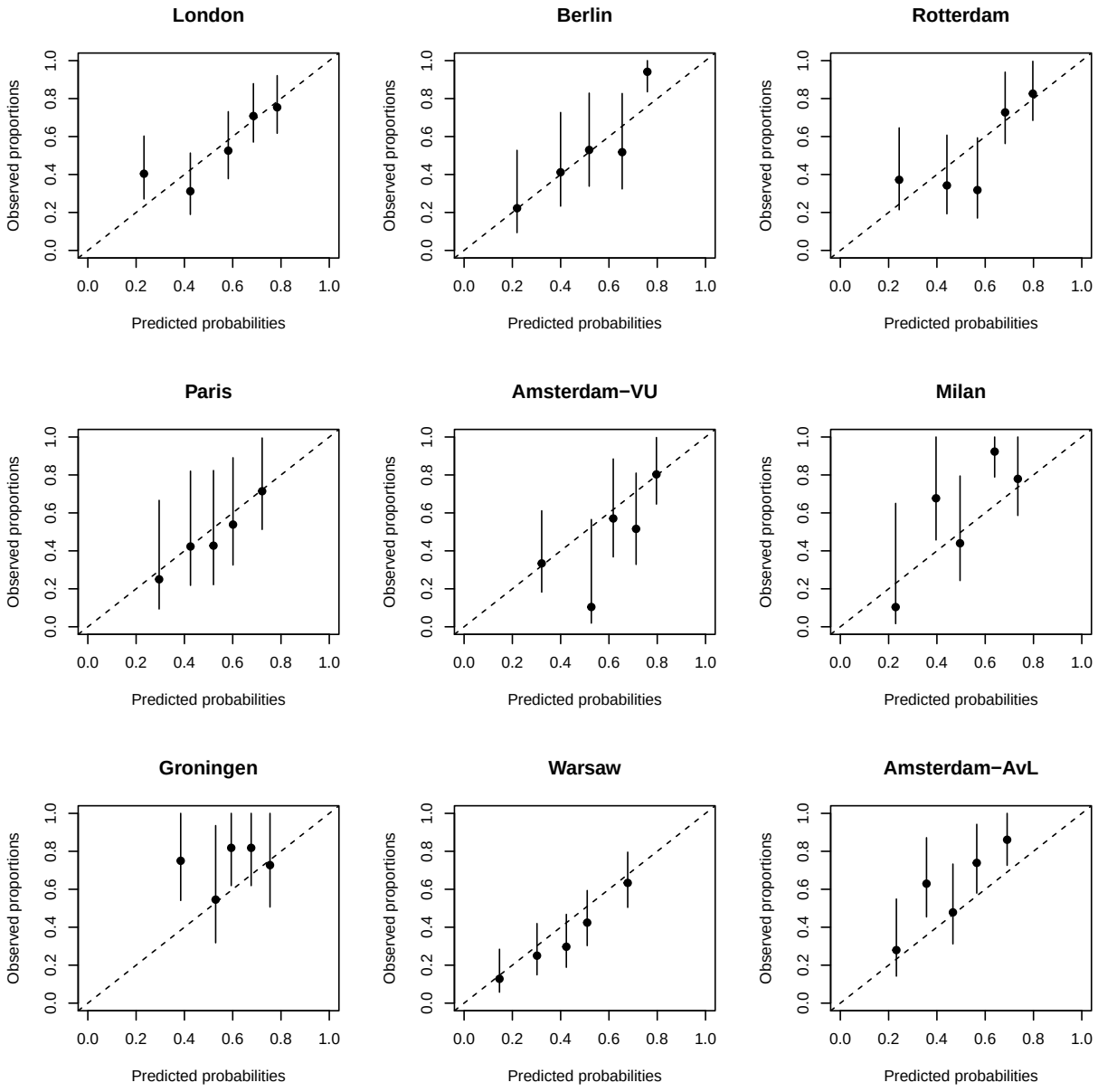

Figure 1: Leave one center out cross validation for the prediction of 5-year recurrence

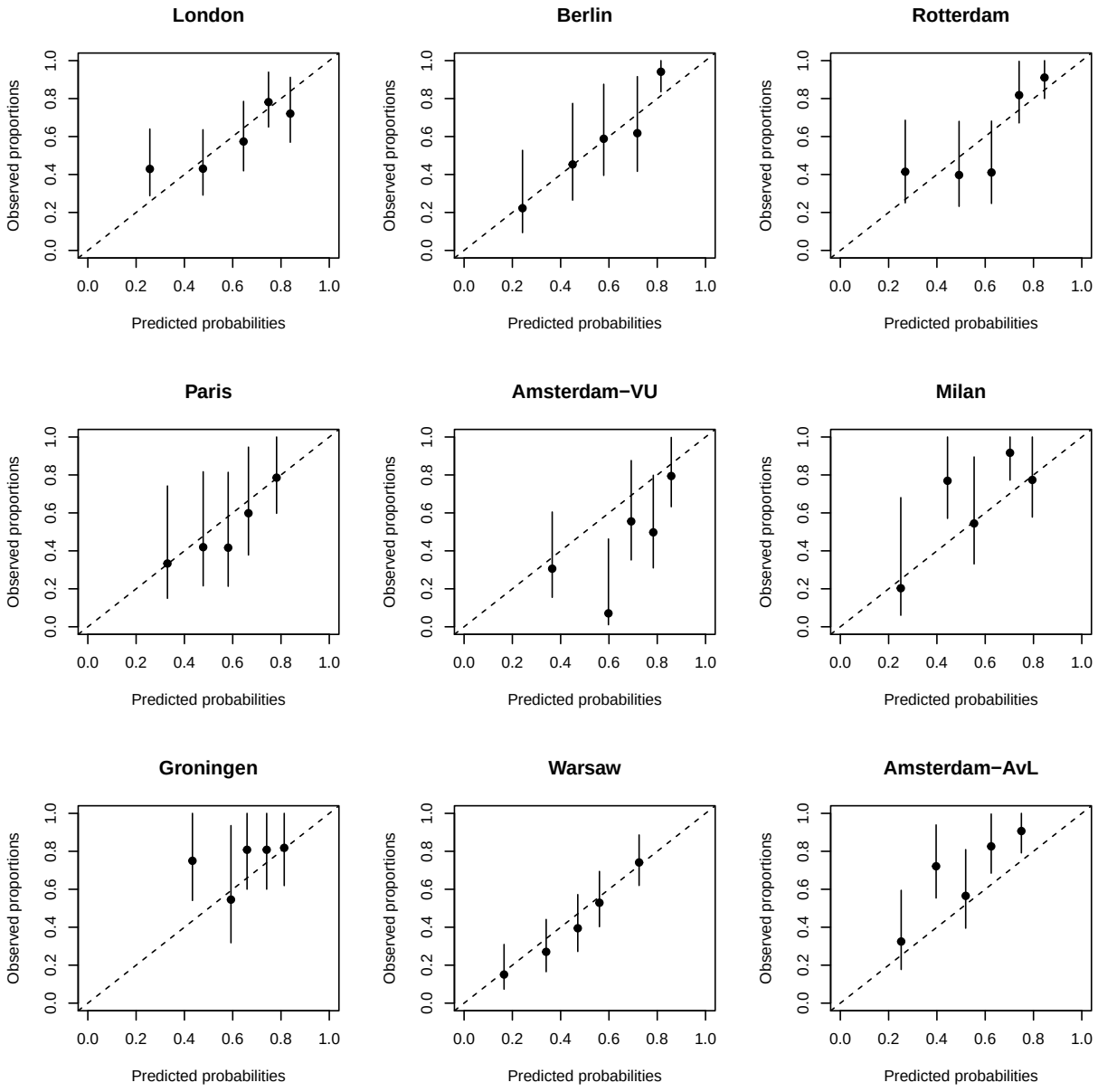

Figure 2: Leave one center out cross validation of the calibrated model for 5-year distant metastasis

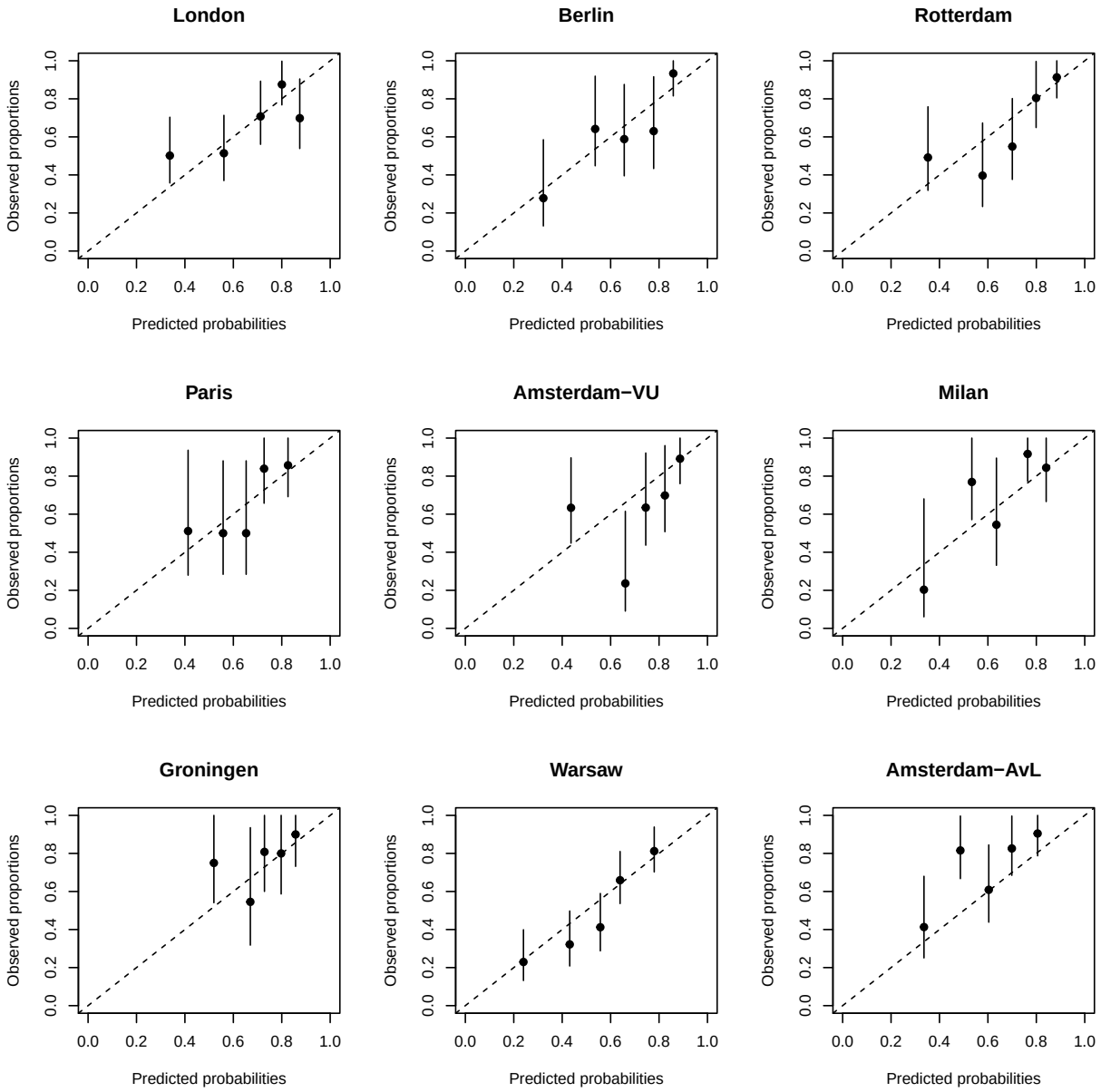

Figure 3: Leave one center out cross validation of the calibrated model for 5-year overall mortality

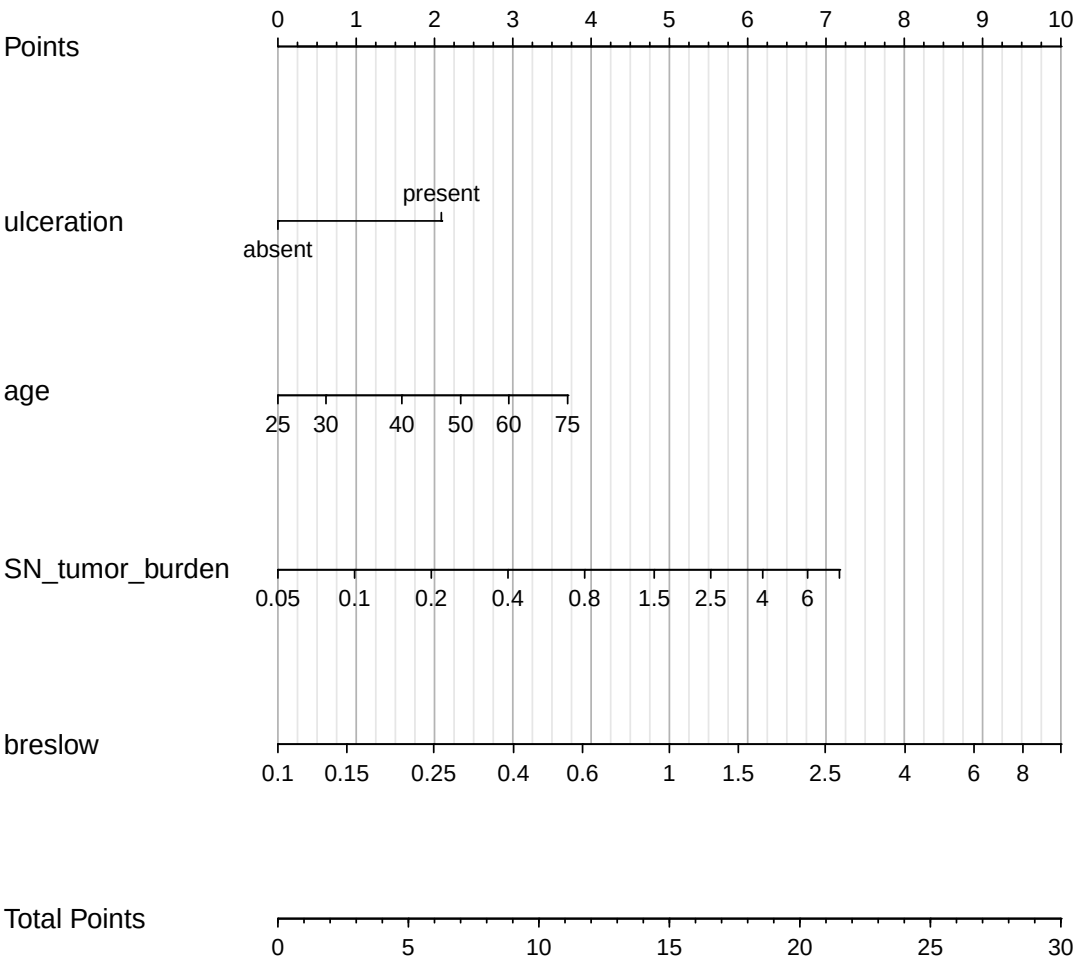

Figure 4: Nomogram for 5-year recurrence

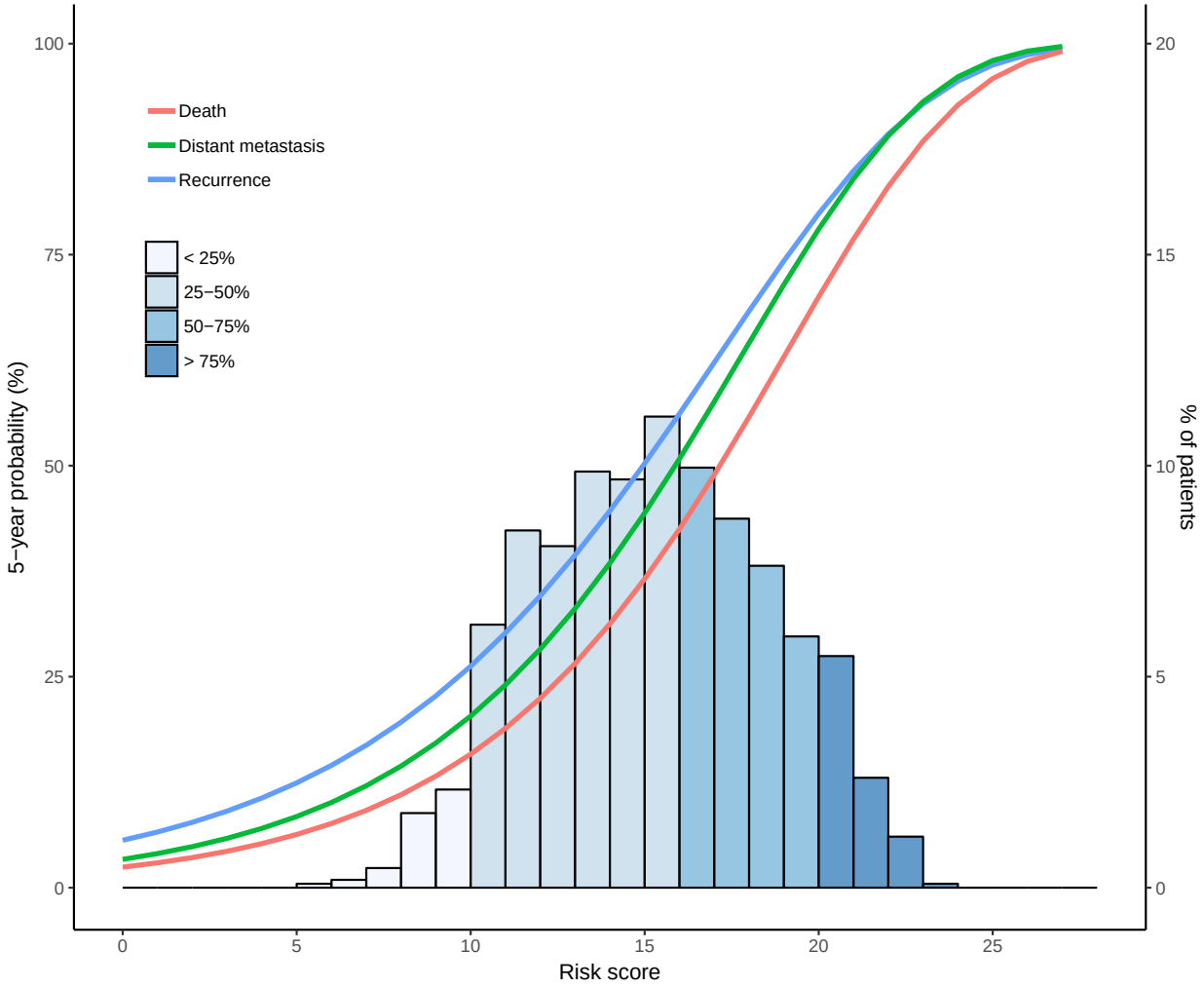

Figure 5: Absolute risks along with risk distribution

Table 3: Model for recurrence to be validated

	Hazard ratio	Lower 0.95	Upper 0.95
age	1.28	1.13	1.46
breslow	1.42	1.24	1.63
ulceration - present:absent	1.42	1.16	1.73
SN_tumor_burden_extended - single cells:0.5 - 1.0 mm	0.47	0.30	0.74
SN_tumor_burden_extended - <0.5 mm:0.5 - 1.0 mm	0.84	0.62	1.13
SN_tumor_burden_extended - >1.0 - 2.0 mm:0.5 - 1.0 mm	1.19	0.90	1.57
SN_tumor_burden_extended - >2.0 - 5.0 mm:0.5 - 1.0 mm	1.57	1.20	2.06
SN_tumor_burden_extended - >5.0 mm:0.5 - 1.0 mm	1.64	1.21	2.23

2 External validation

The variable `SN_tumor_burden` is not available in the valaidation set, but `SN_tumor_burden_extended` is. For that reason we first need to assess the performance of a model derived in the training set, where we substitute `SN_tumor_burden`. The new model has an apparent c-index of 0.68 (95 percent c.i. 0.65 to 0.7). The model to be validated in the validation set is given in Table 3

The altered model for recurrence gave very similar performance in the validation set, with a c-index of 0.7 (95 percent c.i. 0.67 to 0.74). From the calibration plot (Figure 6) we see that there may be slight under-estimation for higher risk patients.

For the case of distant metastasis, the c-index of the calibrated model in external validation was 0.72 (95 percent c.i. 0.68 to 0.75). The model performed very well in terms of calibration as well (Figure 7).

Similar conclusions can be drawn for the case of overall mortality. The c-index of the calibrated model in external validation was 0.74 (95 percent c.i. 0.71 to 0.78) while calibration to the test set was again very good plot except for the case of high risk patients where risks were slightly underestimated (Figure 8)

3 Tertiary objective

The tertiary goal is to create a prediction model for recurrence (and DMFS and OS) using the extra variable positive additional positive nodes after CLND. This model can only be based on those patients who actually underwent a CLND of course. This would tell us something about how much this variable would add to the discrimination of the prediction model.

More specifically, in the case of 5-year recurrence a model including `ulceration`, `age`, `SN_tumor_burden`, `breslow` AND `no_pos_nonSNs_cat` has a c-index of 0.69 with a 95 percent c.i. 0.67 to 0.72, which a slight increase from 0.68 (95 percent c.i. 0.65 to 0.7).

For 5-year distant metastasis the c-index is 0.72 with a 95 percent c.i. 0.69 to 0.74 compared to 0.7 (95 percent c.i. 0.67 to 0.72) of the origial model.

Finally, for 5-year overall mortality the c-index is 0.72 with a 95 percent c.i. 0.69 to 0.75 comapared to 0.70 (95 percent c.i. 0.67 to 0.72) of the origial model.

When we use the simpler EJC risk groups to assign patients to risk categories the performance drops significantly. In Table 4 we give an overview of the perfomrmance regarding discrimination for the different modeling approaches considered.

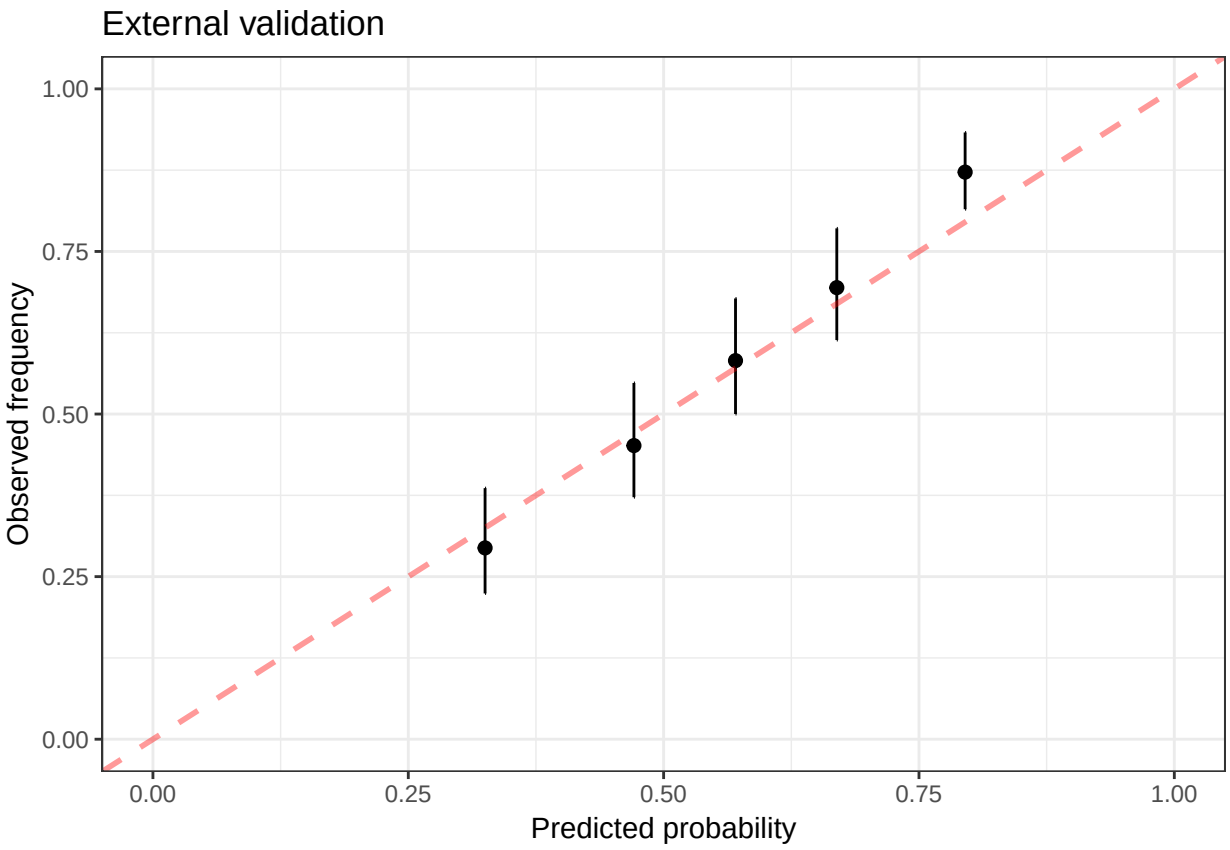

Figure 6: Calibration plot of 5-year recurrence model

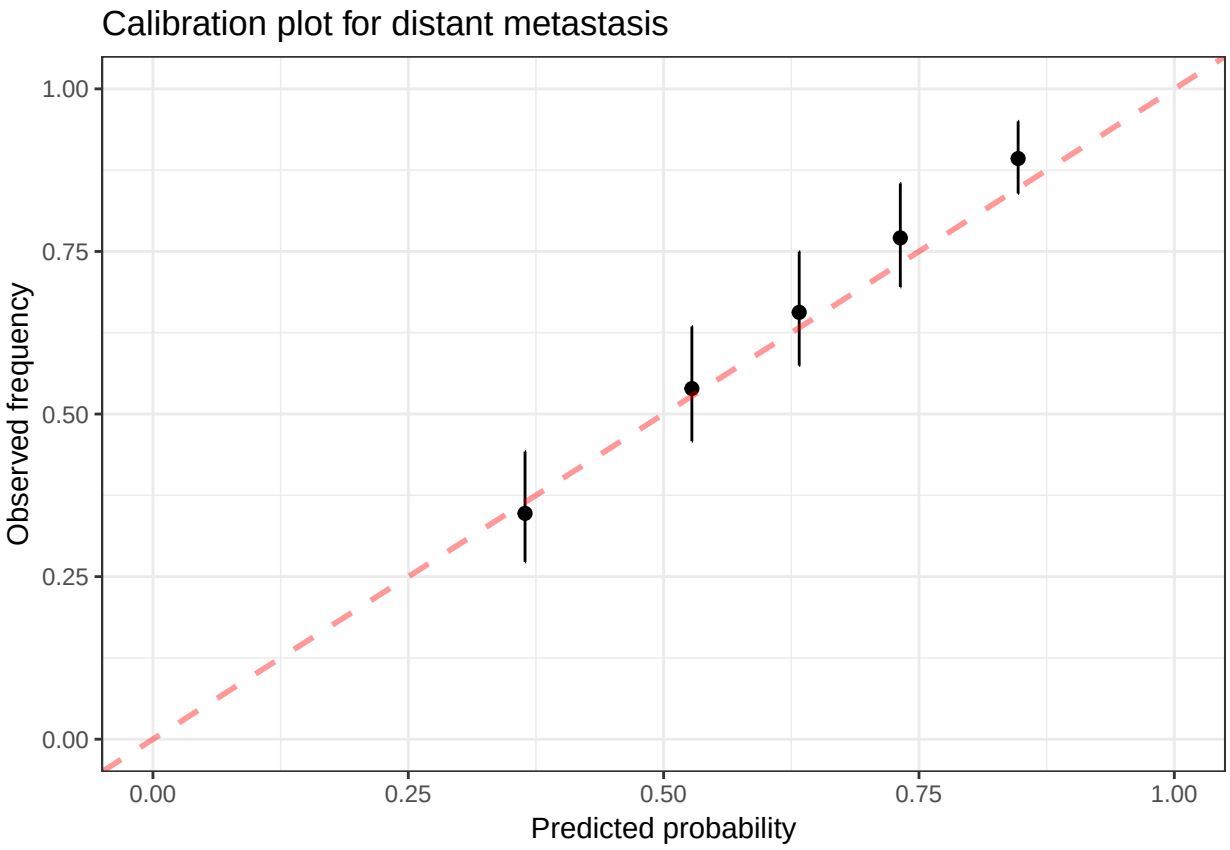

Figure 7: Calibration plot of 5-year distant metastasis model

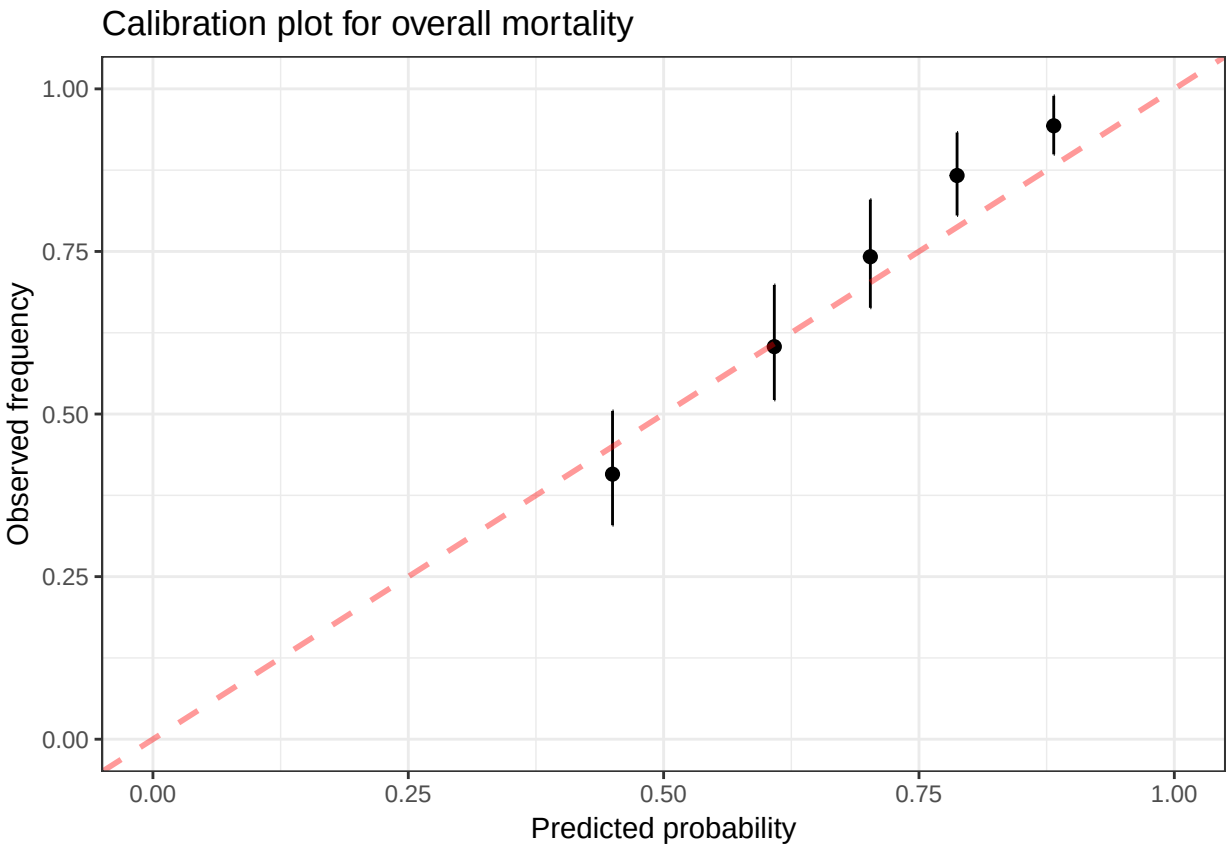

Figure 8: Calibration plot of 5-year overall mortality model

Table 4: C-indices for the different prediction models considered internal discrimination

Models	Recurrence	Distant_Metastasis	Overall_Mortality
EORTC prediction model	0.68 (0.65 - 0.70)	0.70 (0.67 - 0.72)	0.70 (0.67 - 0.73)
EORTC prediction model with no nonSN's	0.69 (0.67 - 0.72)	0.72 (0.69 - 0.74)	0.72 (0.69 - 0.75)
EORTC - simple EJC groups	0.61 (0.59 - 0.63)	0.61 (0.59 - 0.63)	0.60 (0.58 - 0.63)

Table 5: Final model for 5-year recurrence using the merged dataset

	Hazard ratio	Lower 0.95	Upper 0.95
age	1.41	1.26	1.57
SN_tumor_burden	1.33	0.99	1.80
breslow	1.54	1.35	1.77
ulceration - present:absent	1.44	1.20	1.74

4 Merging data sets

Due to the very good performance of the developed model in the test dataset, we combine the training with the test datasets to for the development of the final model.

4.1 Recurrence

The final model for the prediction of 5-year recurrence from the merged set can be found in Table 5. The c-index of that model is 0.68 (95 percent c.i. 0.65 to 0.71). Calibration is assessed using a leave-one-center-out cross validation, based on the merged dataset. In this way, information from the 9 development centers of the first objective can be used to impute missing values for `SN_tumor_burden` in the German (validation) dataset. Multiple imputation is again separated between the training and the test set as was done in the original analysis. When leaving out the German dataset, because we cannot use the data from the other centers to impute the missing values, we use the approach of External validation (the same holds for the assessment of calibration in the calibrated models for distant metastasis and overall mortality). The results of this analysis can be found in Figure 9 for the 5-year recurrence model, in Figure 11 for the 5-year distant recurrence model and in Figure 12 for the 5-year overall mortality model.

4.2 Nomogram

The updated nomogram based on the combined dataset can be found in Figure 10.

4.3 Distant metastasis

For the assessment of distant metastasis we considered a calibrated version of the 5-year recurrence model. The association between distant metastasis and was of the same size (calibration slope of 1.01, 95 percent c.i. 0.87 to 1.16). We compare the performance of the considered model to that of multivariable Cox regression model including all 9 covariates of interest (`sex`, `ulceration`, `no_removed_SNs`, `no_pos_SNs`, `simpleloc`, `histology_simple`, `breslow`, `SN_tumor_burden` and `age`). The full model had a c-index of 0.7 (95 percent c.i. 0.66 to 0.73) while the calibrated model had a c-index of 0.7 (95 percent c.i. 0.68 to 0.72).

Again we perform a leave-one-center-out cross validation to assess the calibration of our final model for distant metastasis. The approach is the same as the one considered for the 5-year recurrence model (Figure 11).

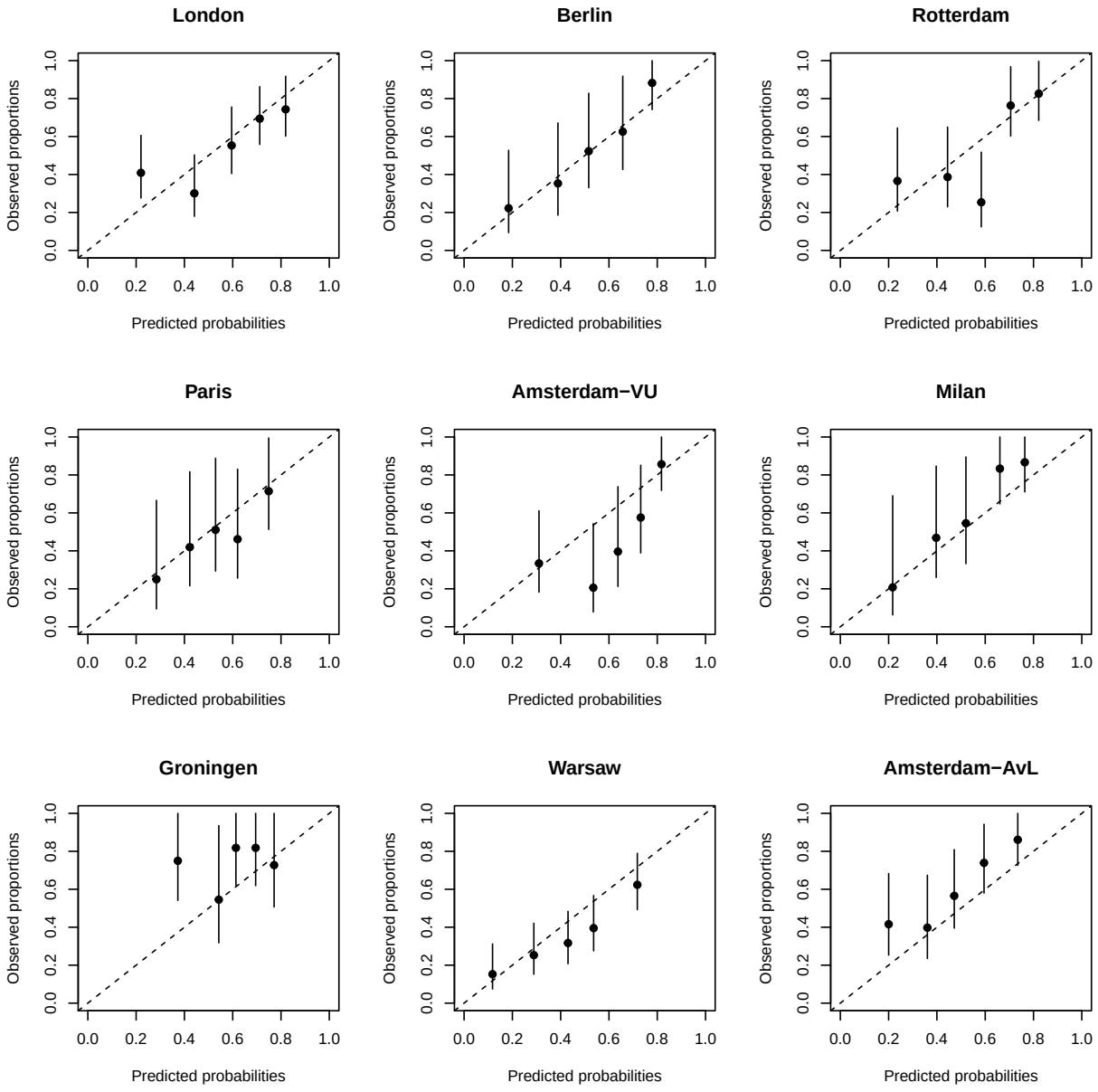

Figure 9: Leave-one-center-out cross validation for prediction of recurrence in the merged dataset

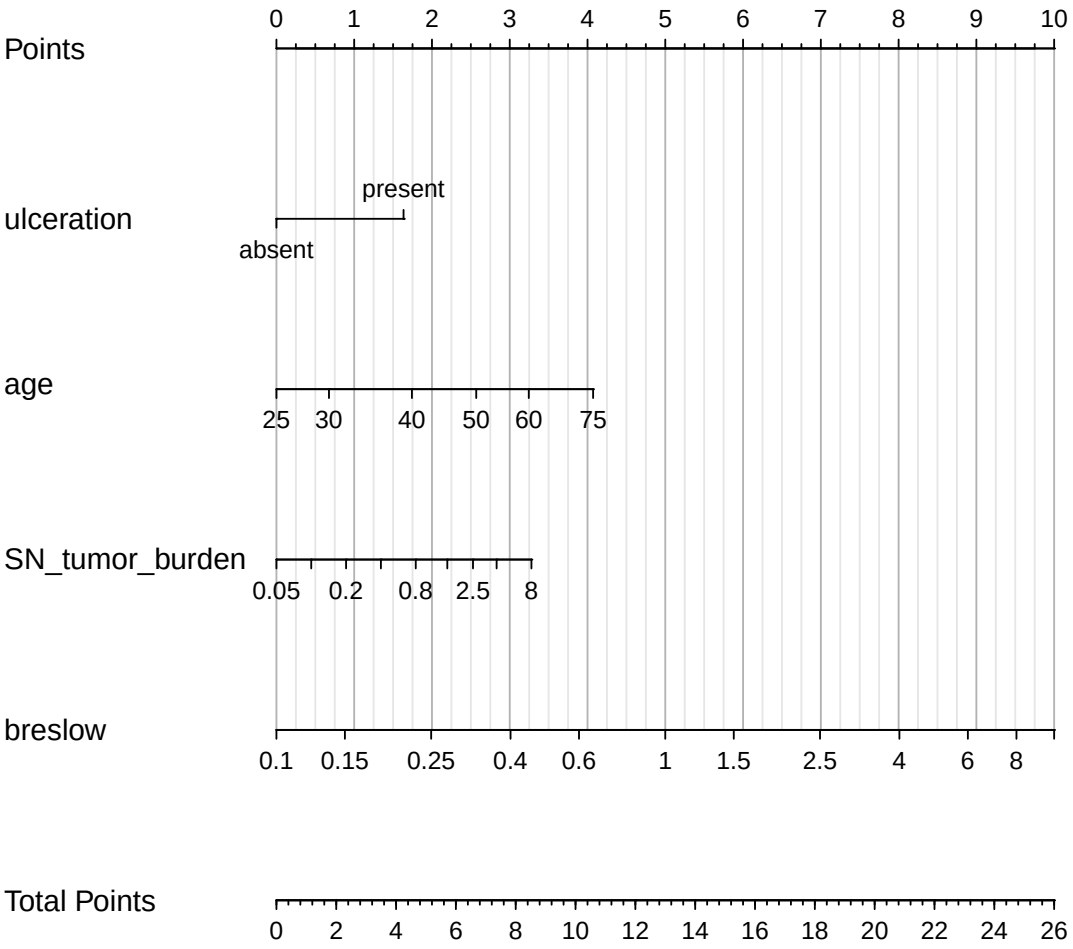

Figure 10: Nomogram for 5-year recurrence for the merged dataset

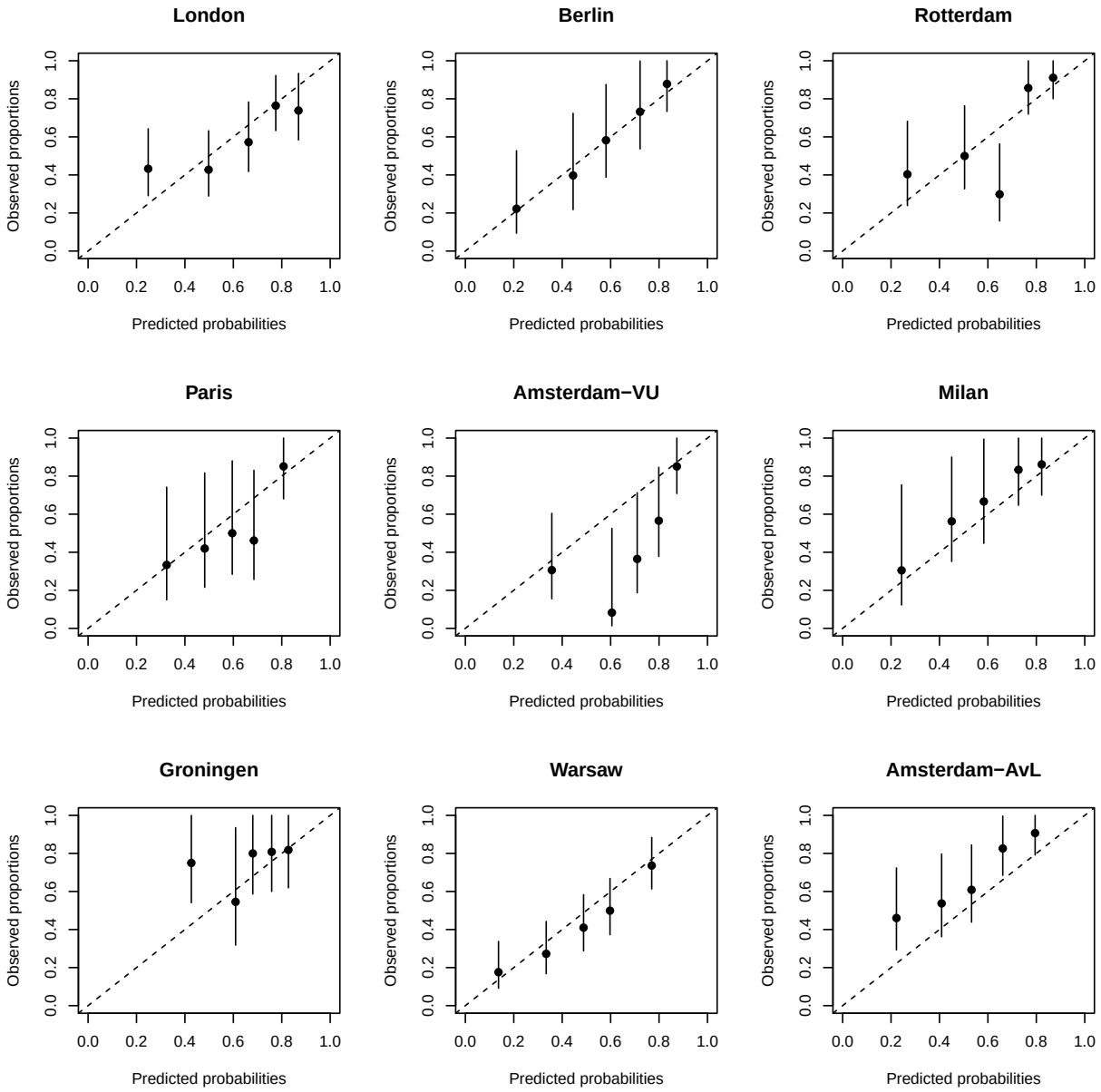

Figure 11: Leave-one-center-out cross validation for prediction of distant metastasis in the merged dataset

4.4 Overall mortality

For the assessment 5-year overall mortality in the merged dataset we considered a calibrated version of the 5-year recurrence model. We compare the performance of the considered model to that of multivariable Cox regression model including all 9 covariates of interest. The association between recurrence and overall mortality was not different (calibration slope 1.04, 95 percent c.i. 0.88 to 1.20). The full model had a c-index of 0.7 (95 percent c.i. 0.66 to 0.75) while the calibrated model had a c-index of 0.71 (95 percent c.i. 0.69 to 0.73). We assess calibation as previously using leave-one-center-out cross validation.

The results of leave-one-center-out cross validation can be found in Figure 12.

The combined results regarding the calibrated models for 5-year distant recurrence and 5-year overall mortality along with the predictions from the 5-year recurrence model can be found in Figure 13.

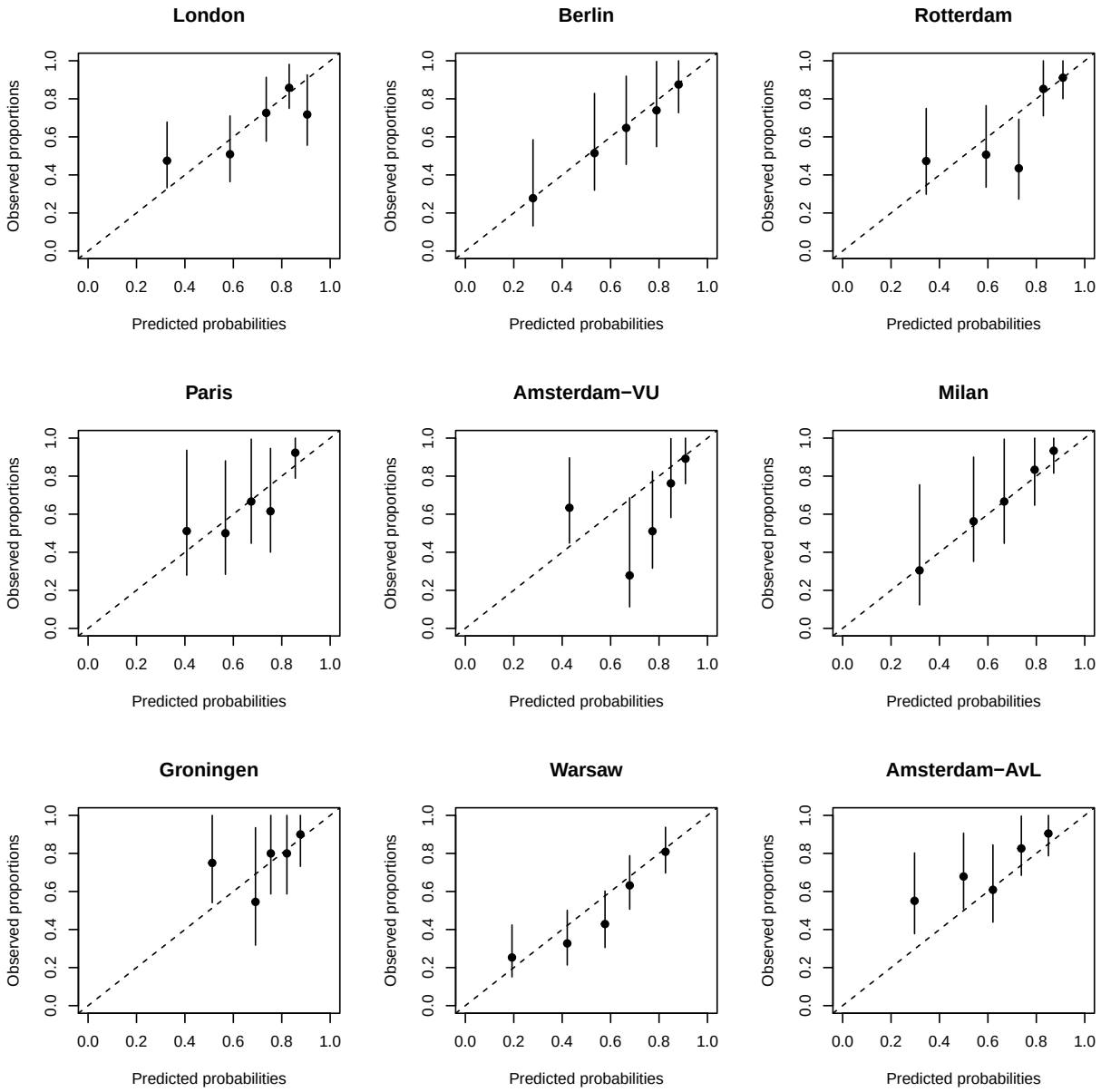

Figure 12: Leave-one-center-out cross validation for prediction of overall mortality in the merged dataset

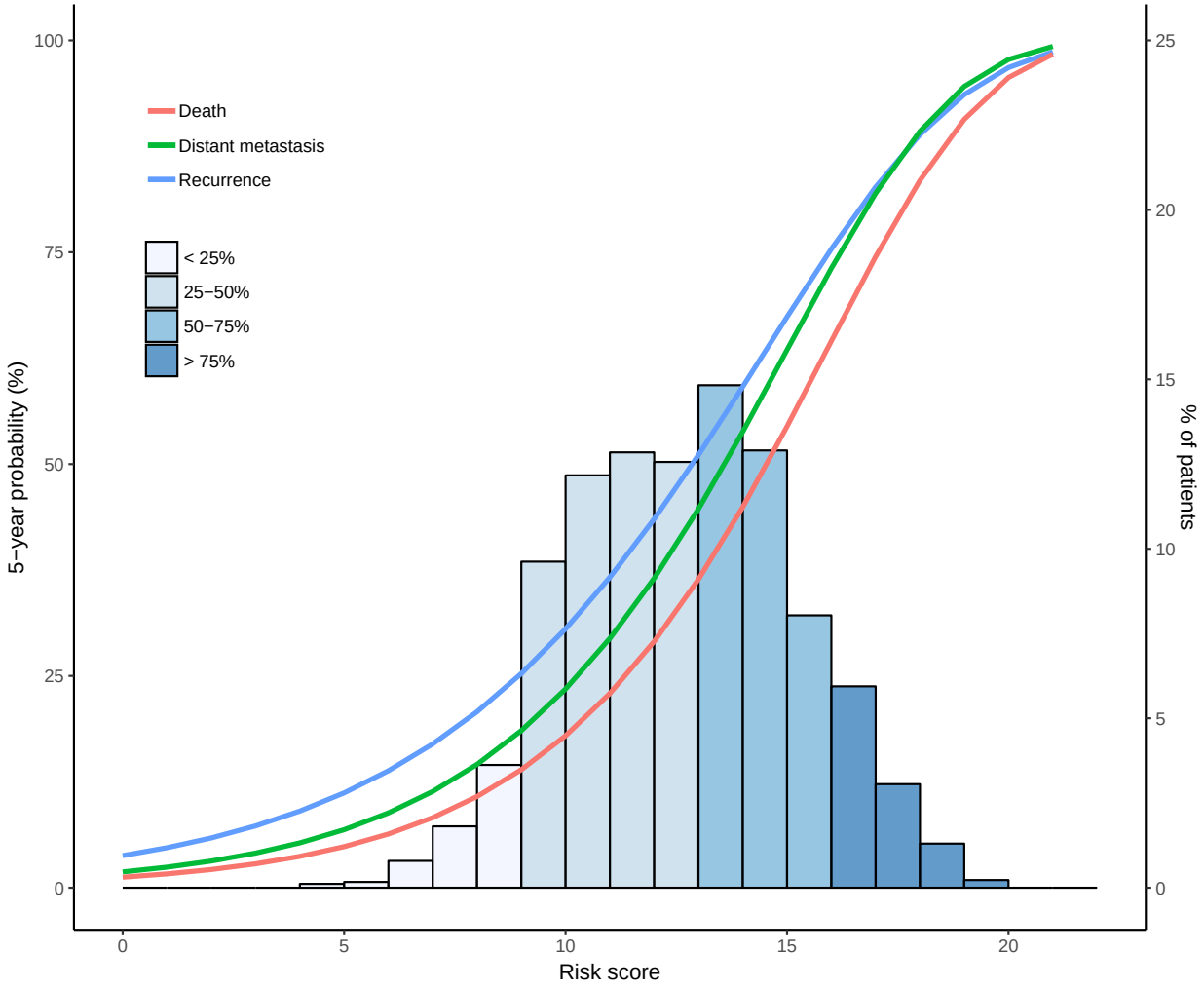

Figure 13: Absolute risks along with risk distribution using the merged dataset