Statistical analyses for peer review uncertainty study

Adrian Barnett

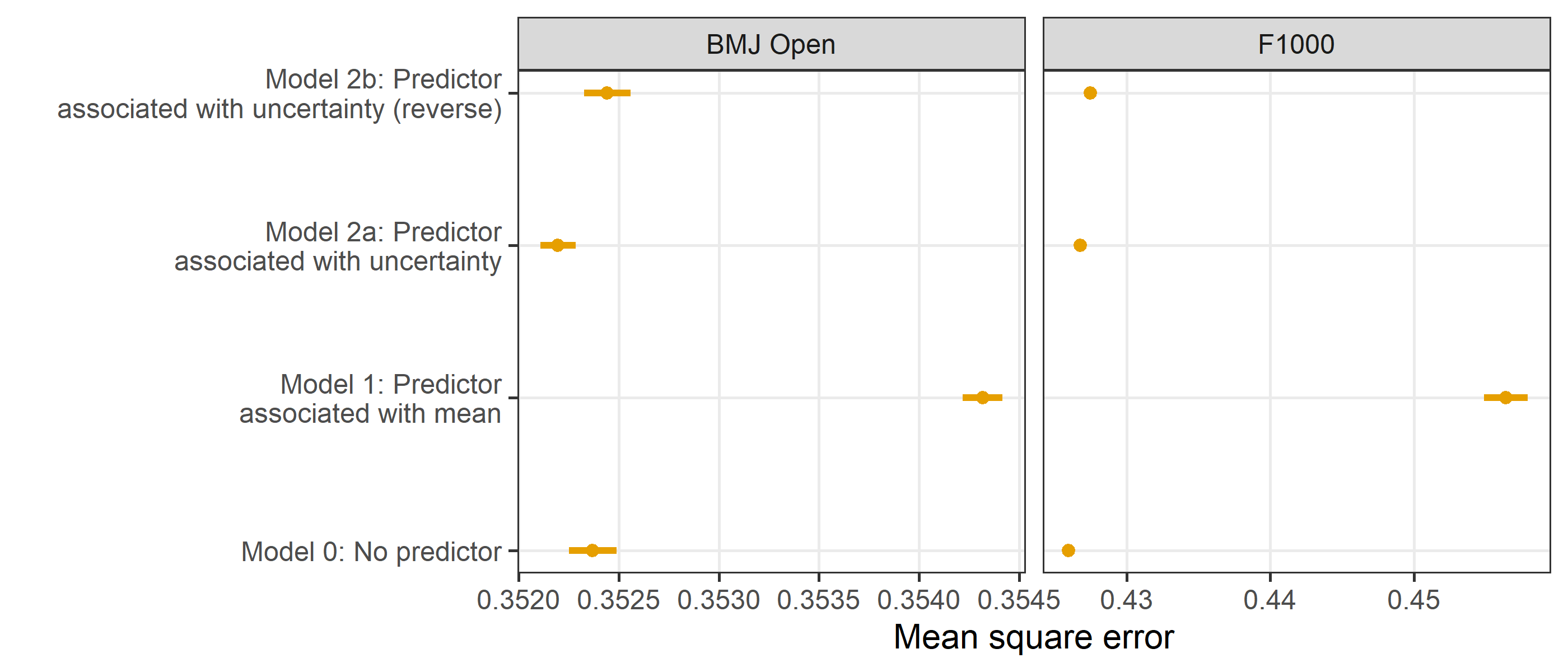
02 March, 2024

Here we show the results of the regression models that examined if uncertainty was associated with reviewer characteristics.

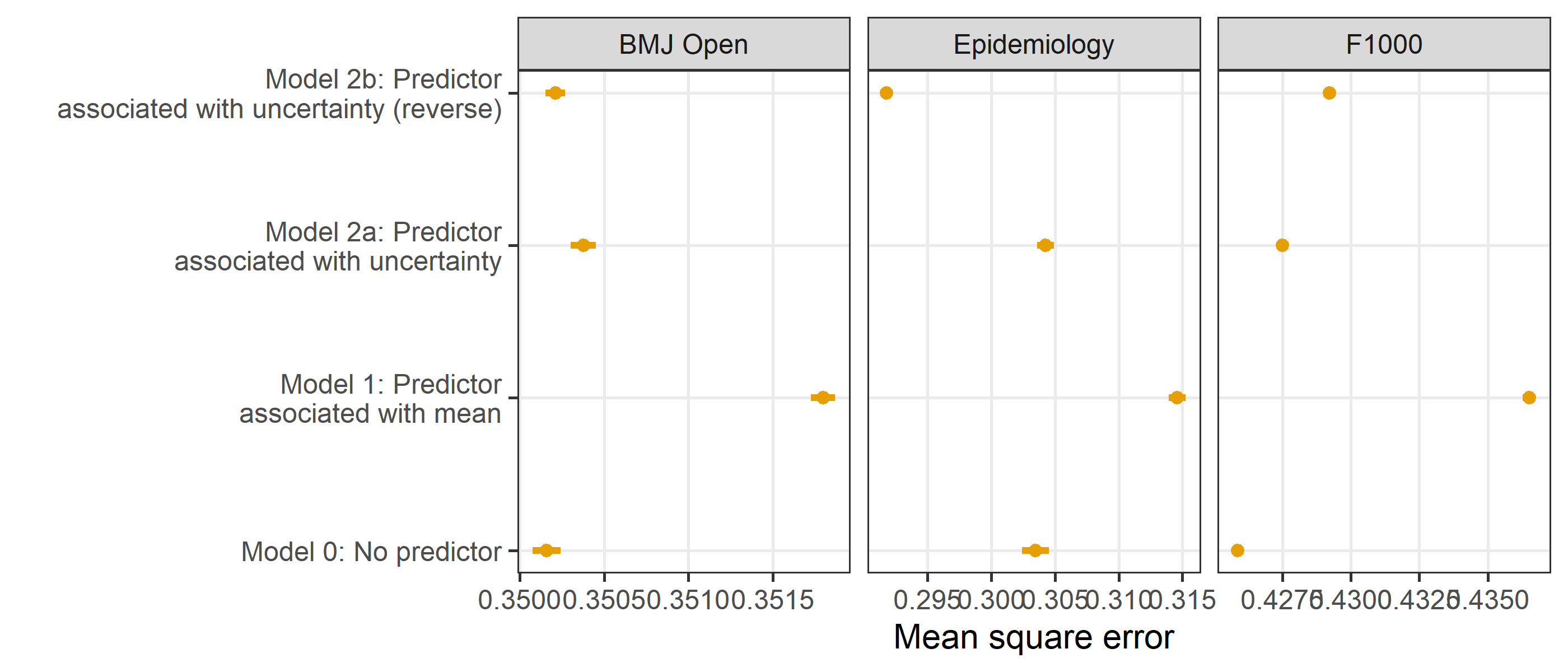
## Model fit

The model fit was assessed using ten-fold cross validation and the mean square error with a 95% confidence interval for the mean. Separate results are shown for each predictor. Results are shown for the complete case and imputed data. The key comparison is whether the mean square errors are smaller than the baseline model with no predictors (model 0).

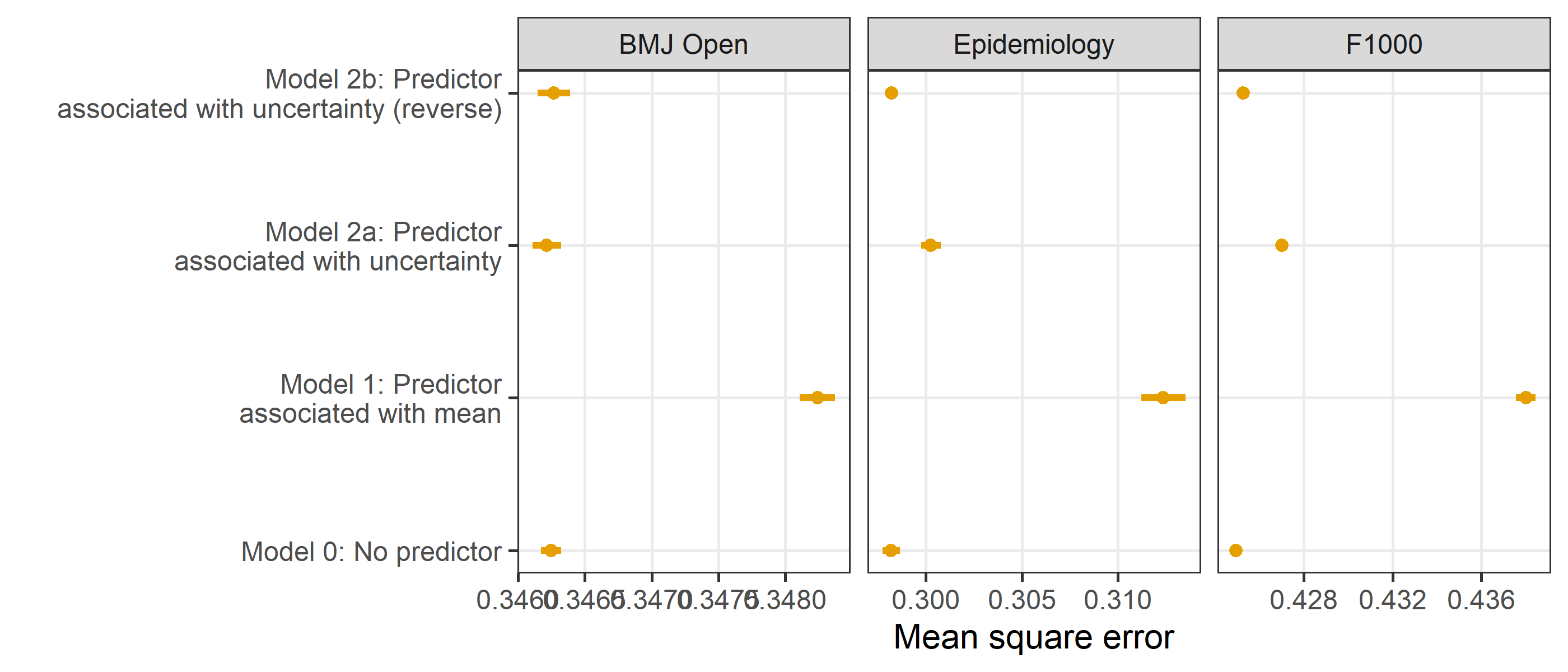
### Protocol (model fit)



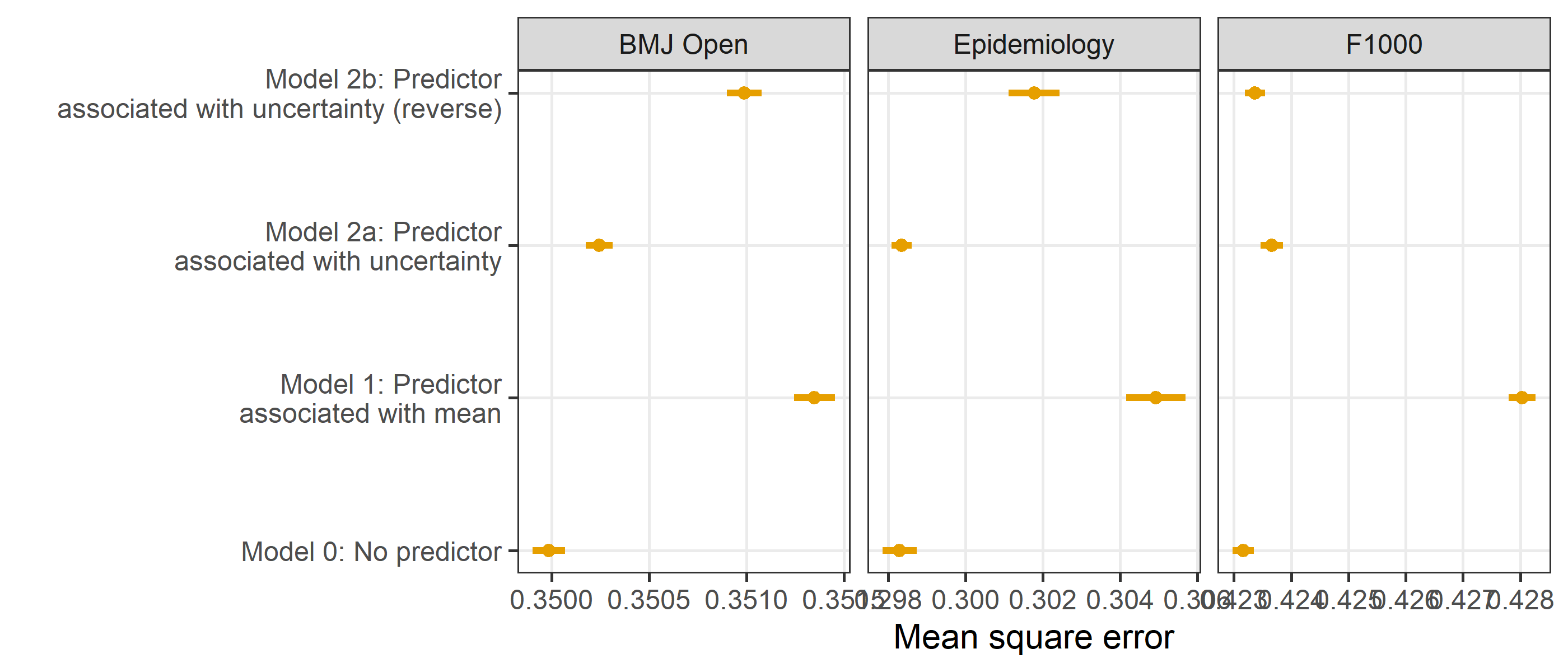
### Reviewer’s gender (model fit)



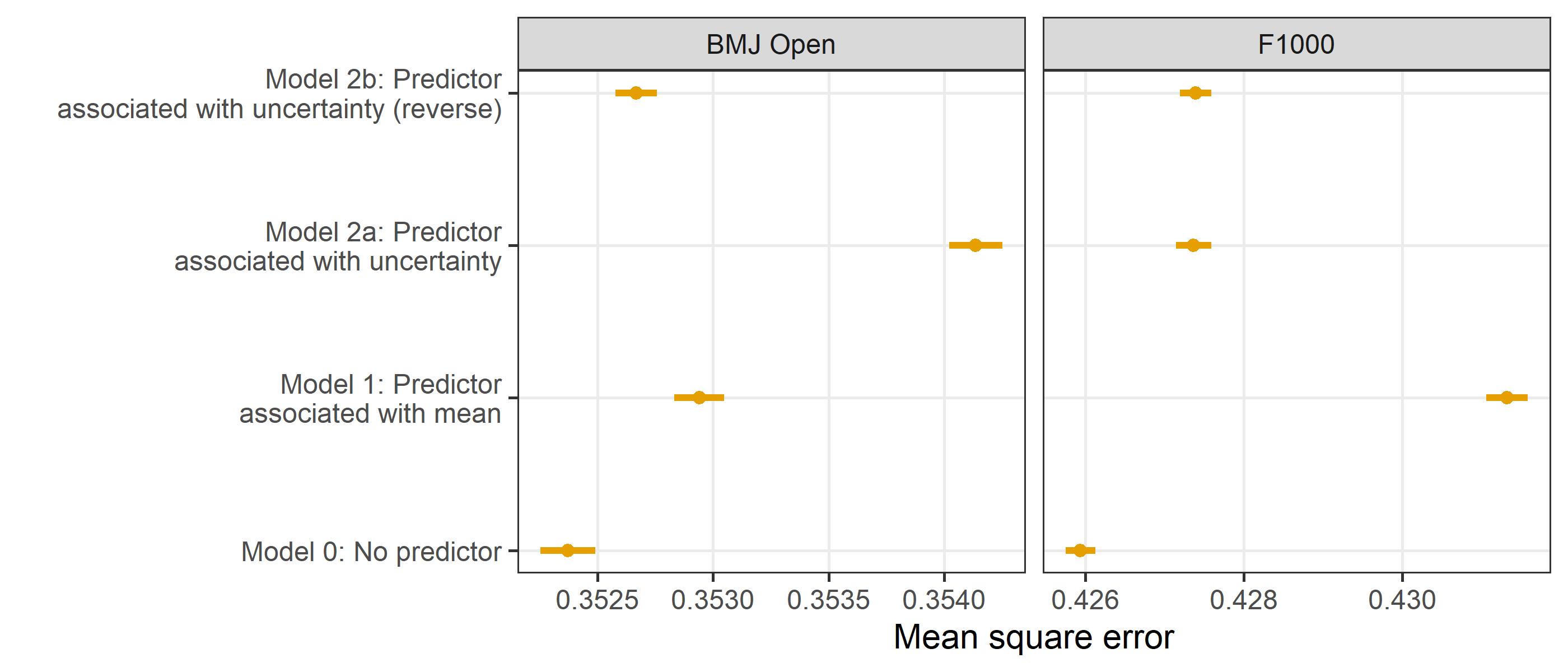
### Reviewer’s time spent (model fit)



### Reviewer’s experience (model fit)

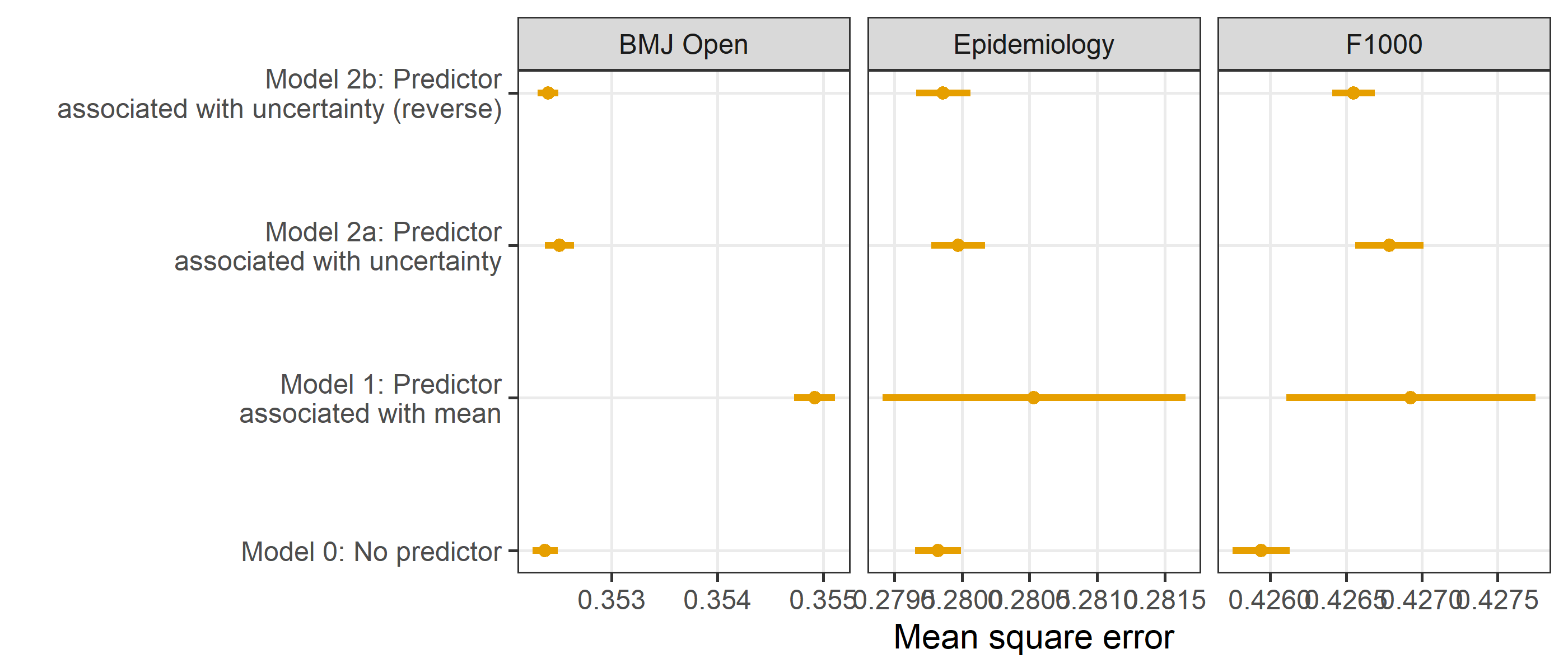


### Version (model fit)



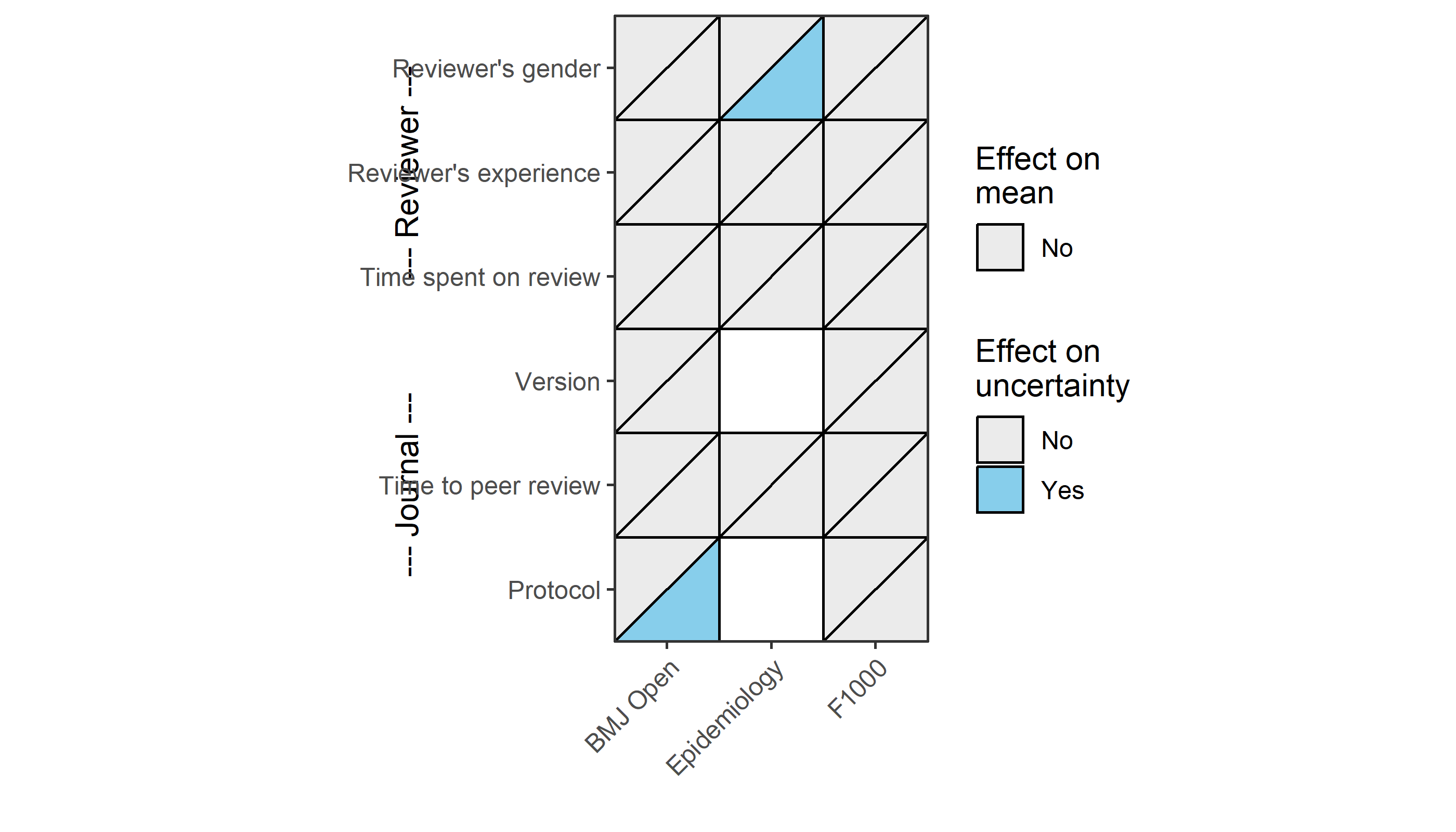
All the submissions at Epidemiology were the first version and so this journal was not included in this analysis.

### Time difference from submission to review (model fit)



## Model fit summary

The figure below summarises the model fit results as shows for each journal and predictor whether the predictor was associated with changes in the mean and/or uncertainty.



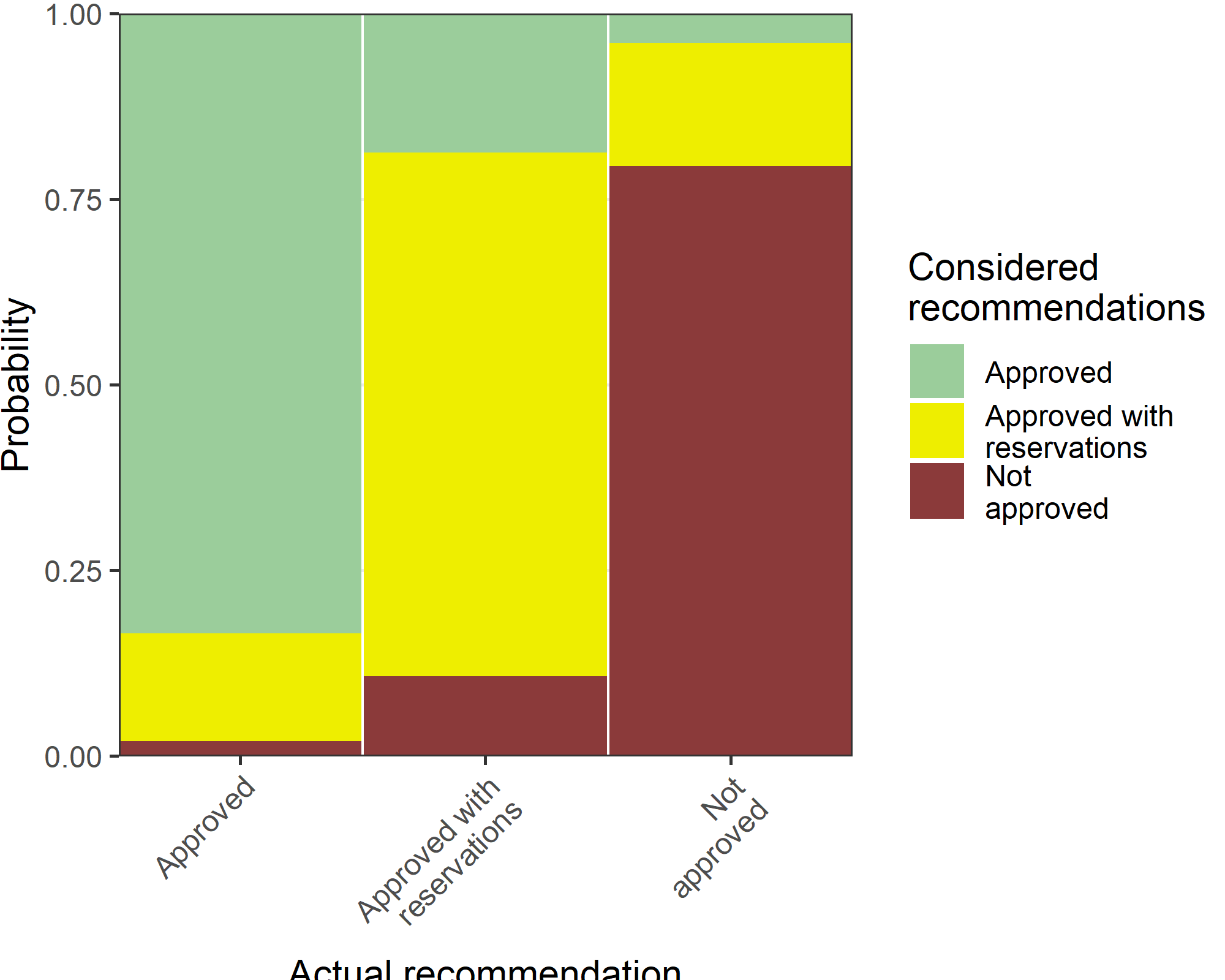
At Epidemiology there were no protocols and all papers were the first version.

## Plots of fitted values

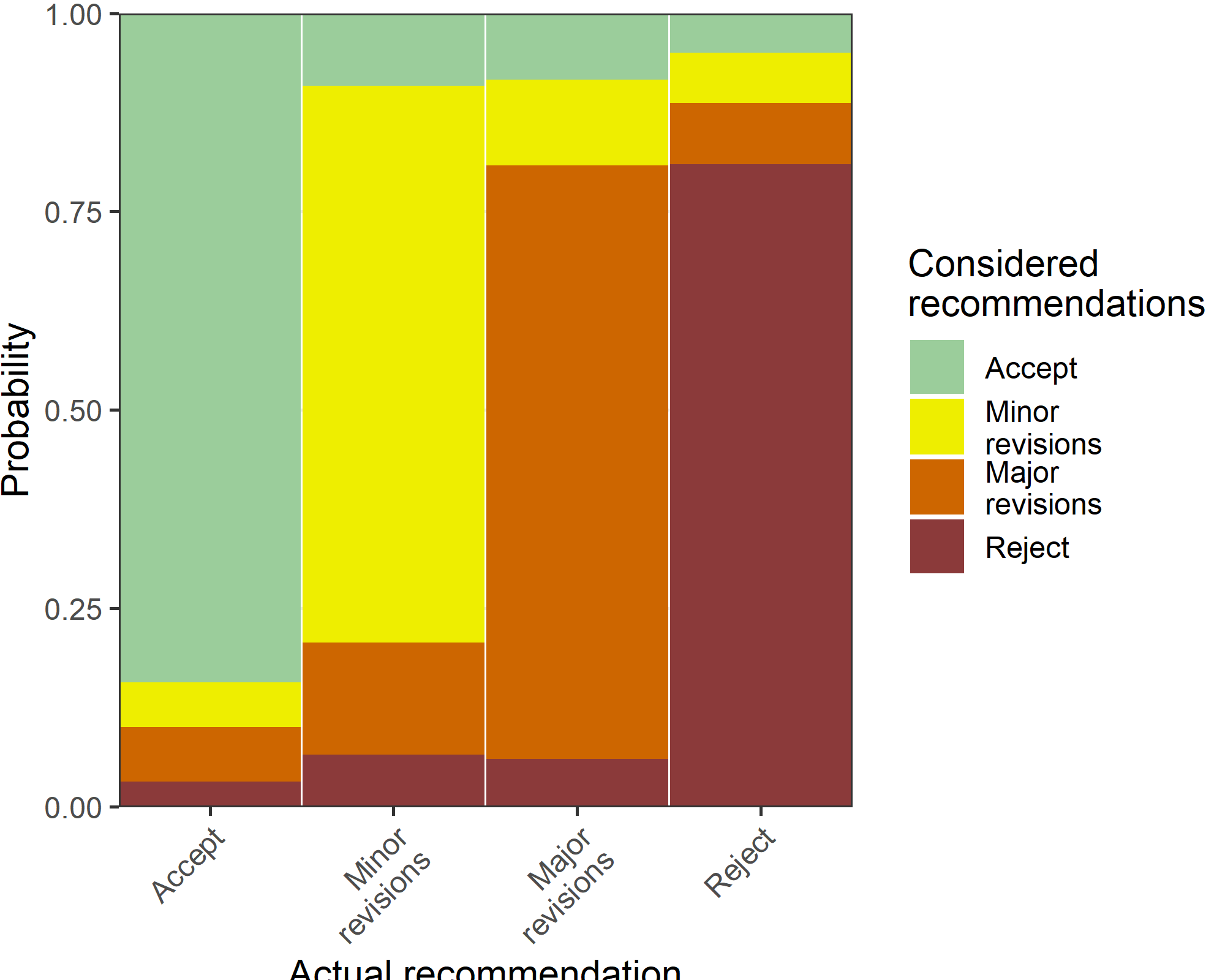
### Model without predictors

This is the baseline model (Model 0) where the only predictor is the reviewers’ recommendations.

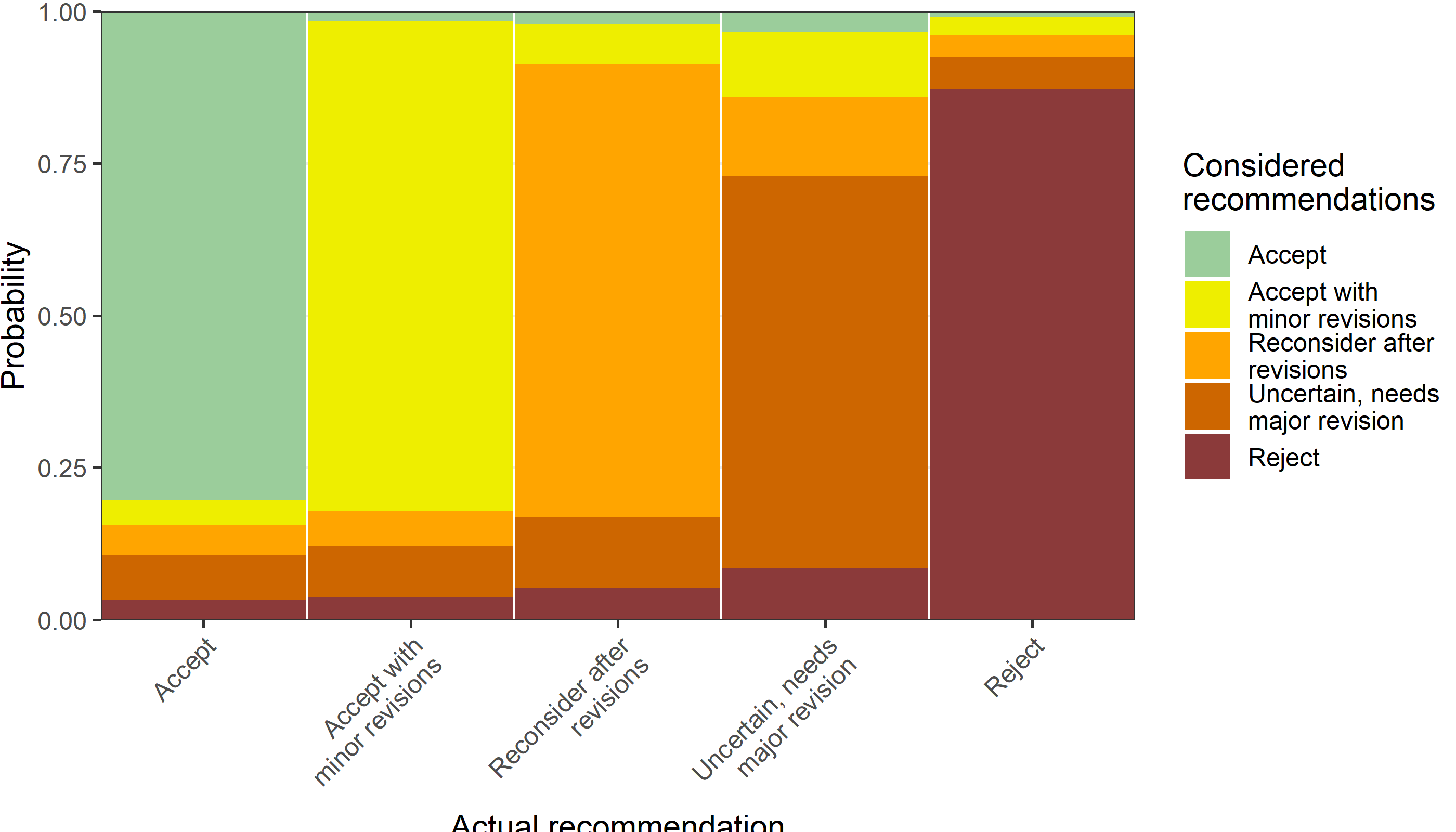
#### F1000, No predictors



#### BMJ Open, No predictors

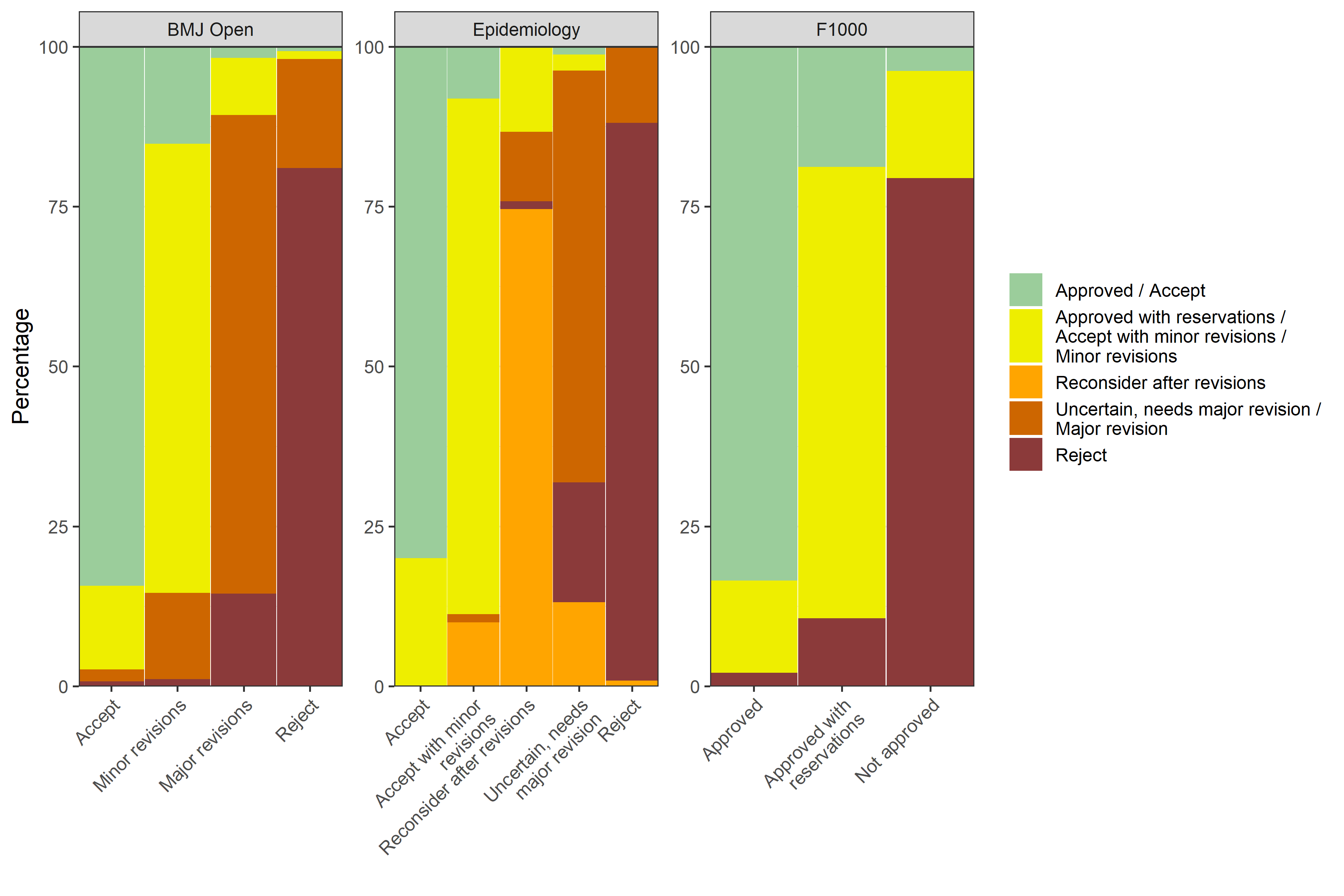


#### Epidemiology, No predictors



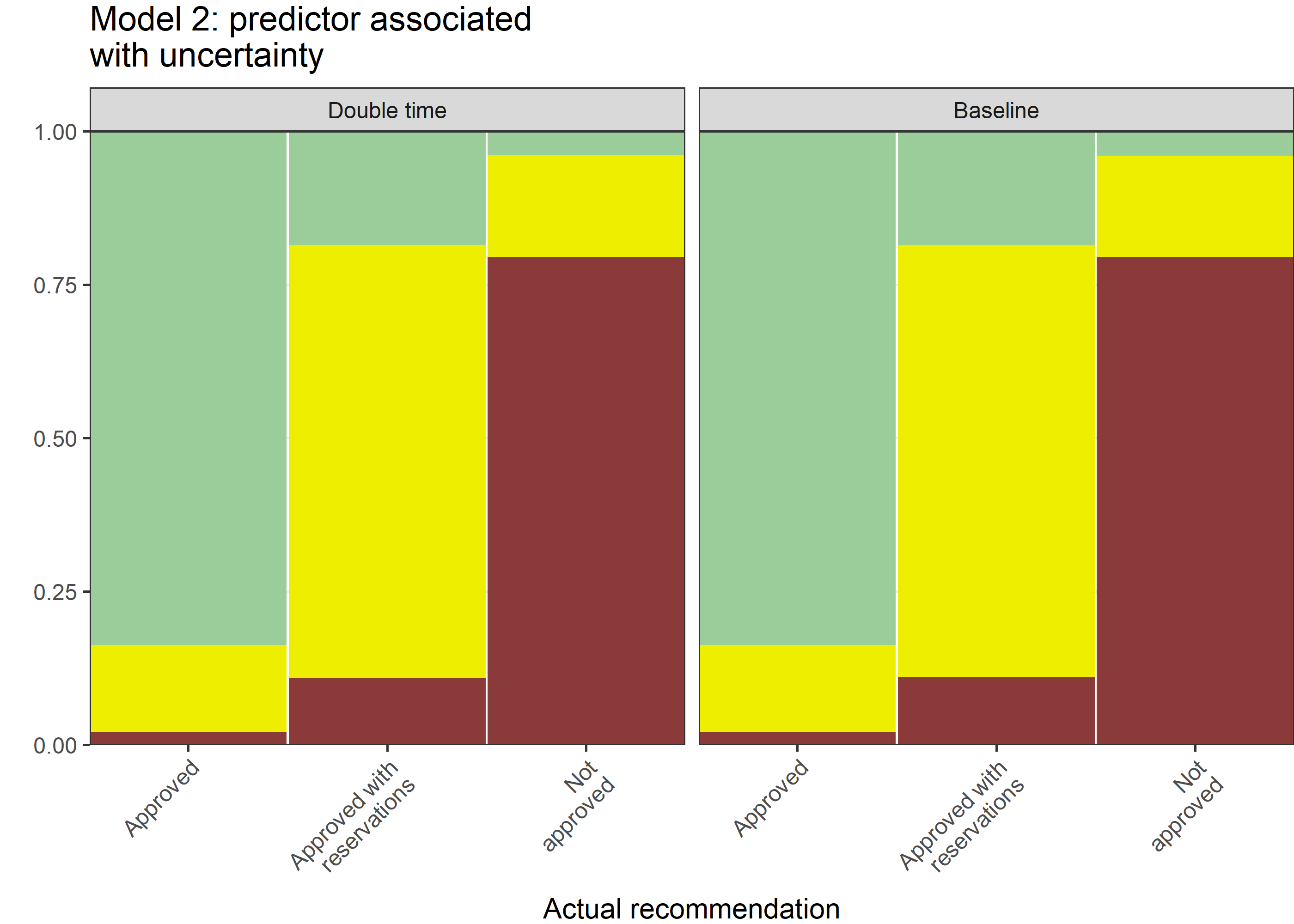
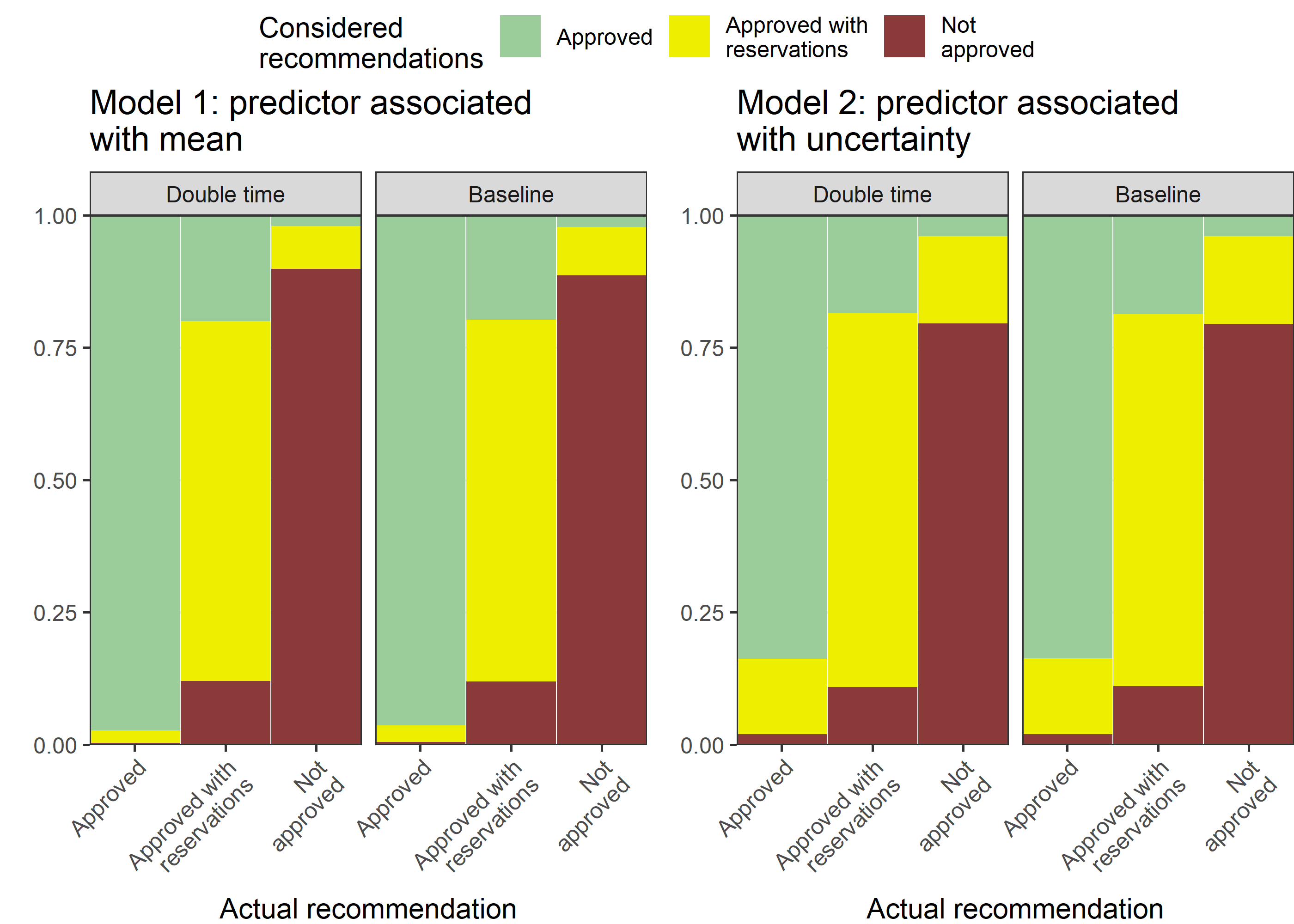
#### Summary statistics

Here we use summary statistics to verify the fitted values from the models above. We include the modal percentage in each recommendation.

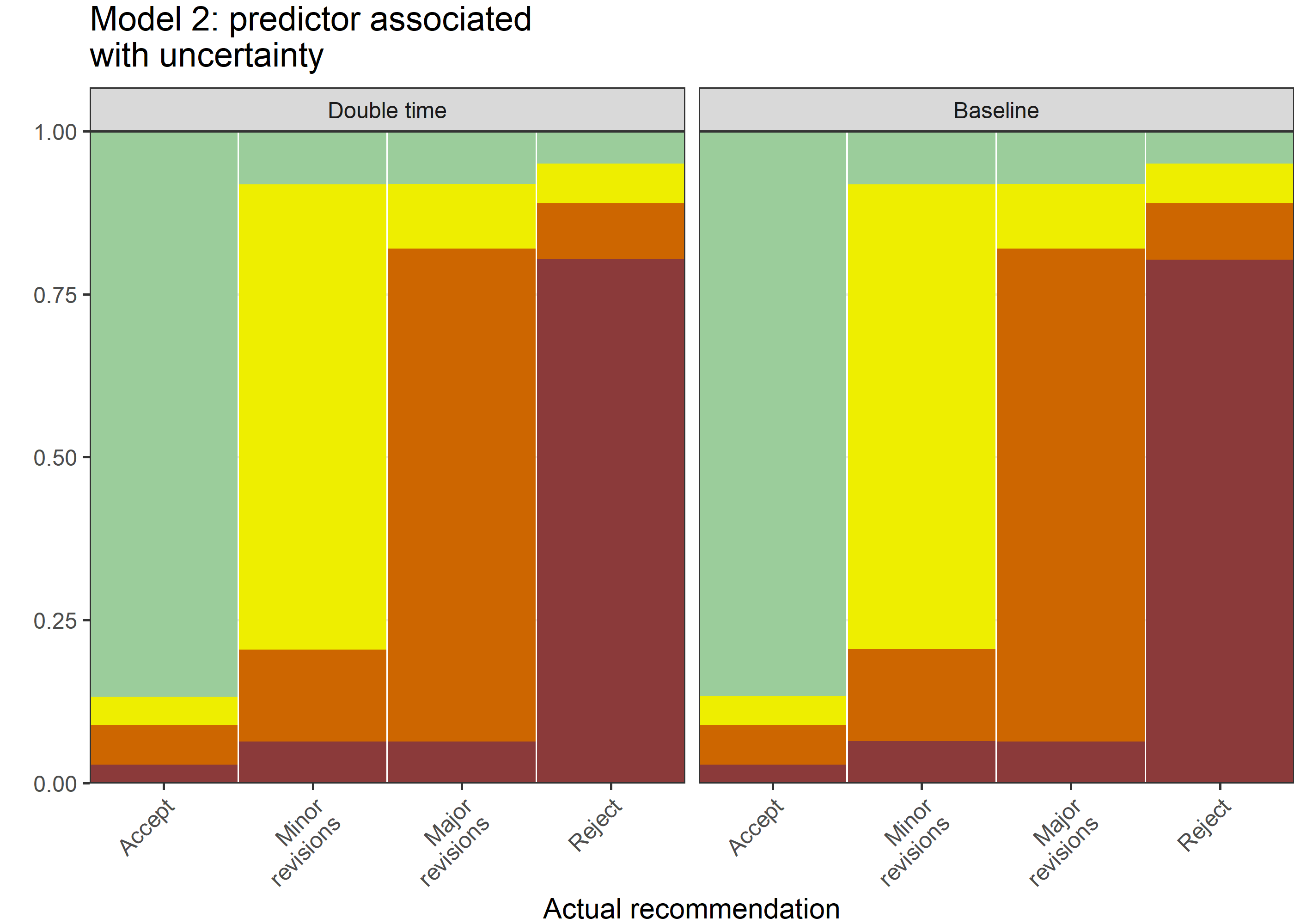
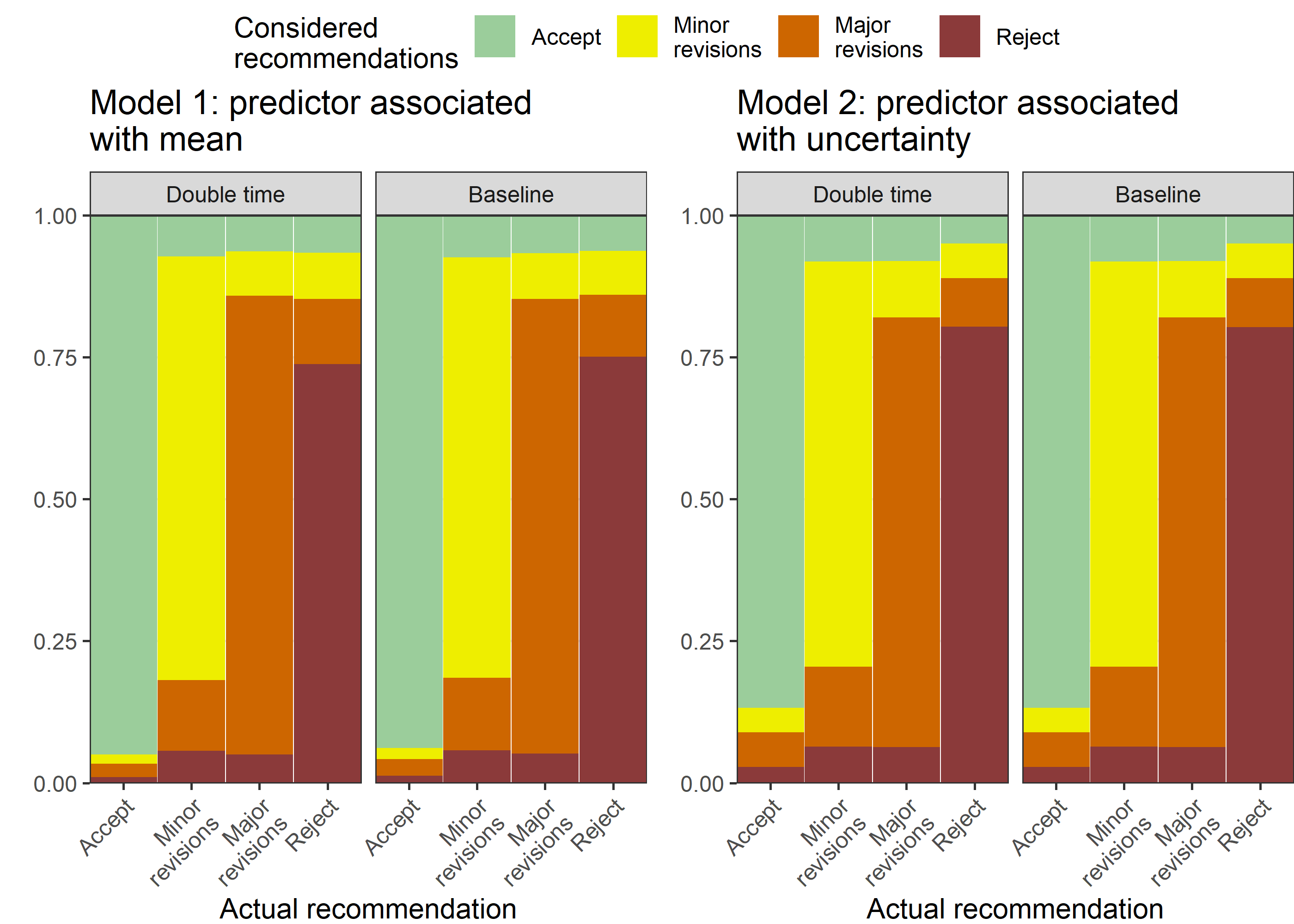


### Time to peer review

#### F1000, Time to peer review

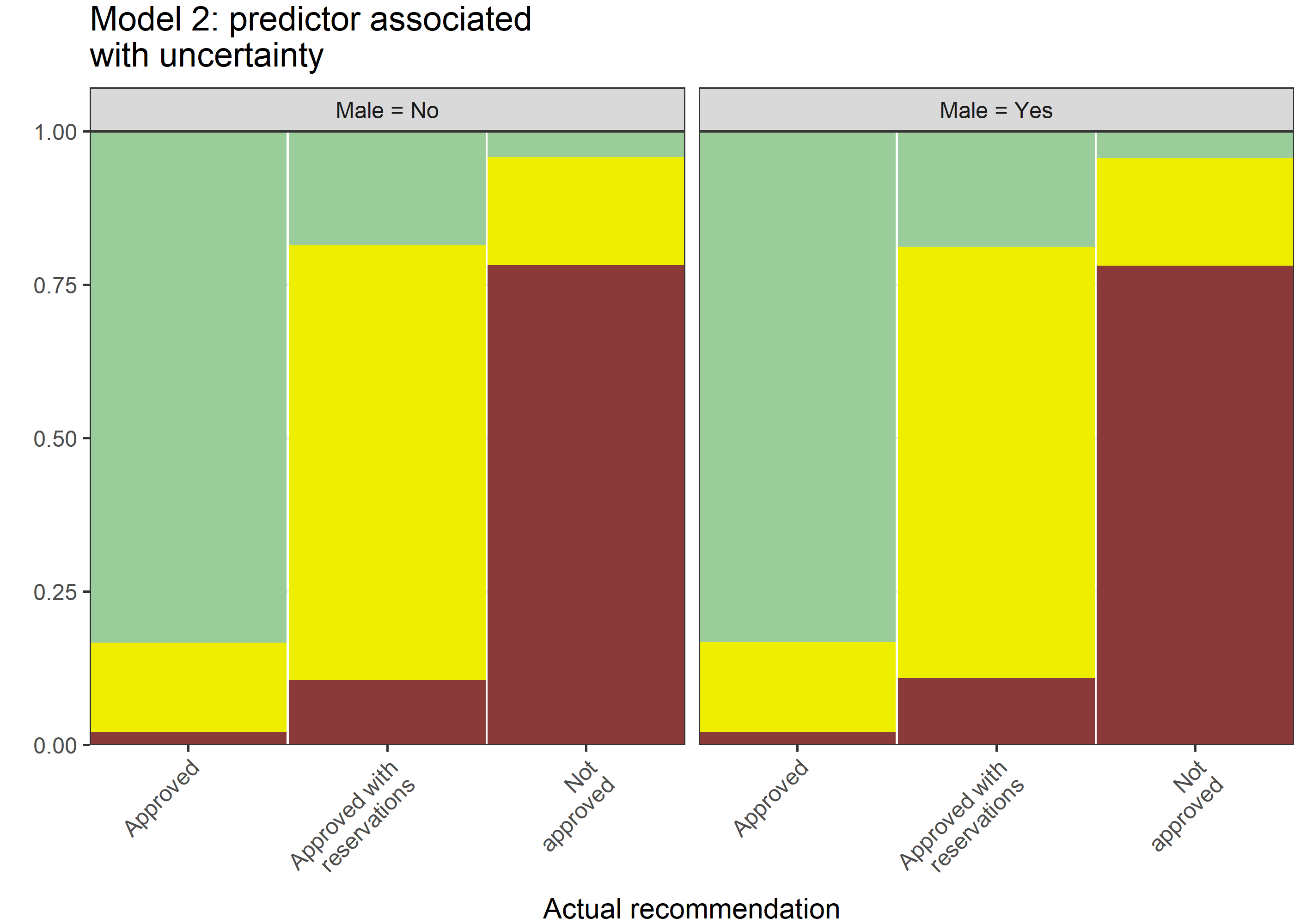
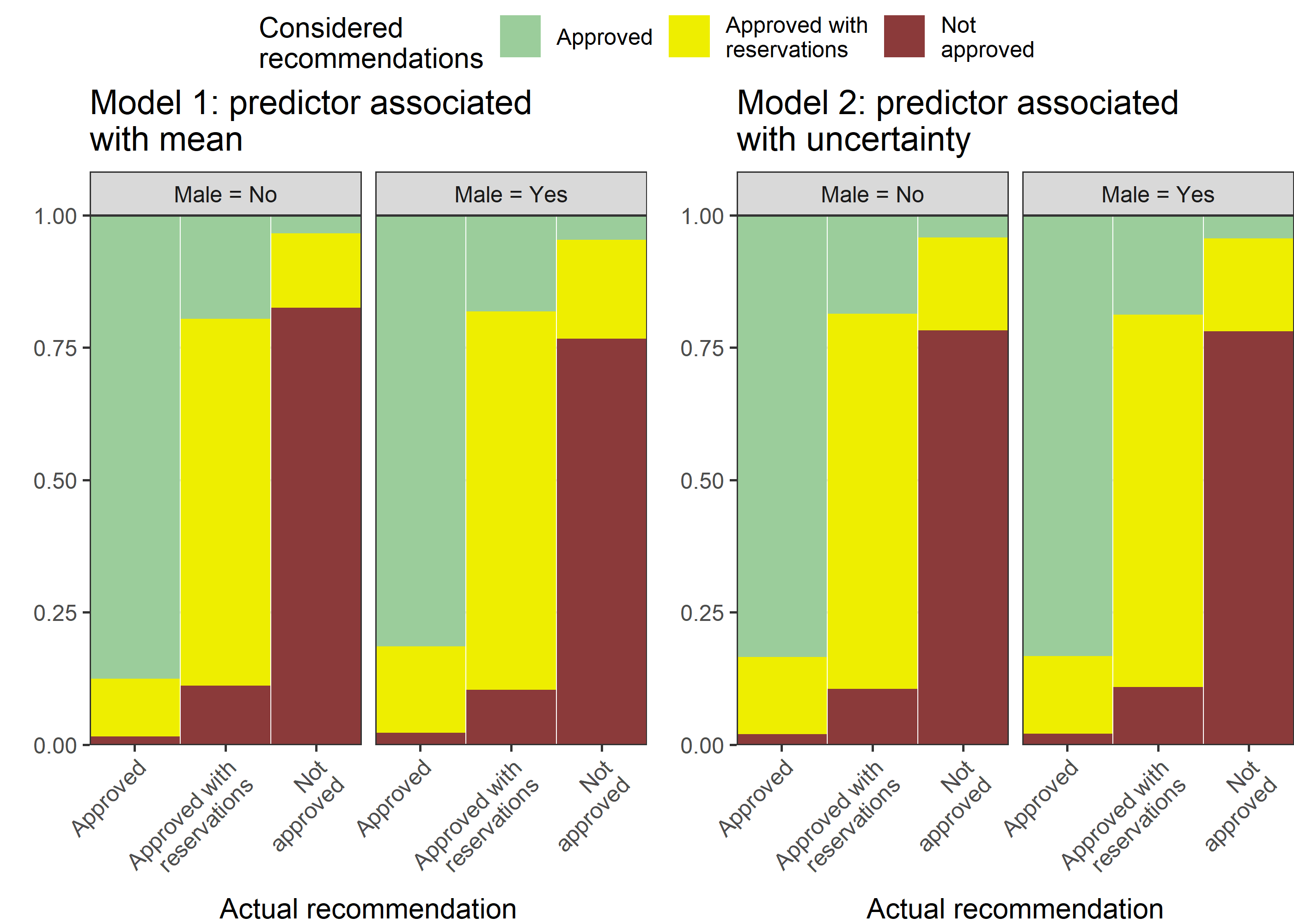


#### BMJ Open, Time to peer review

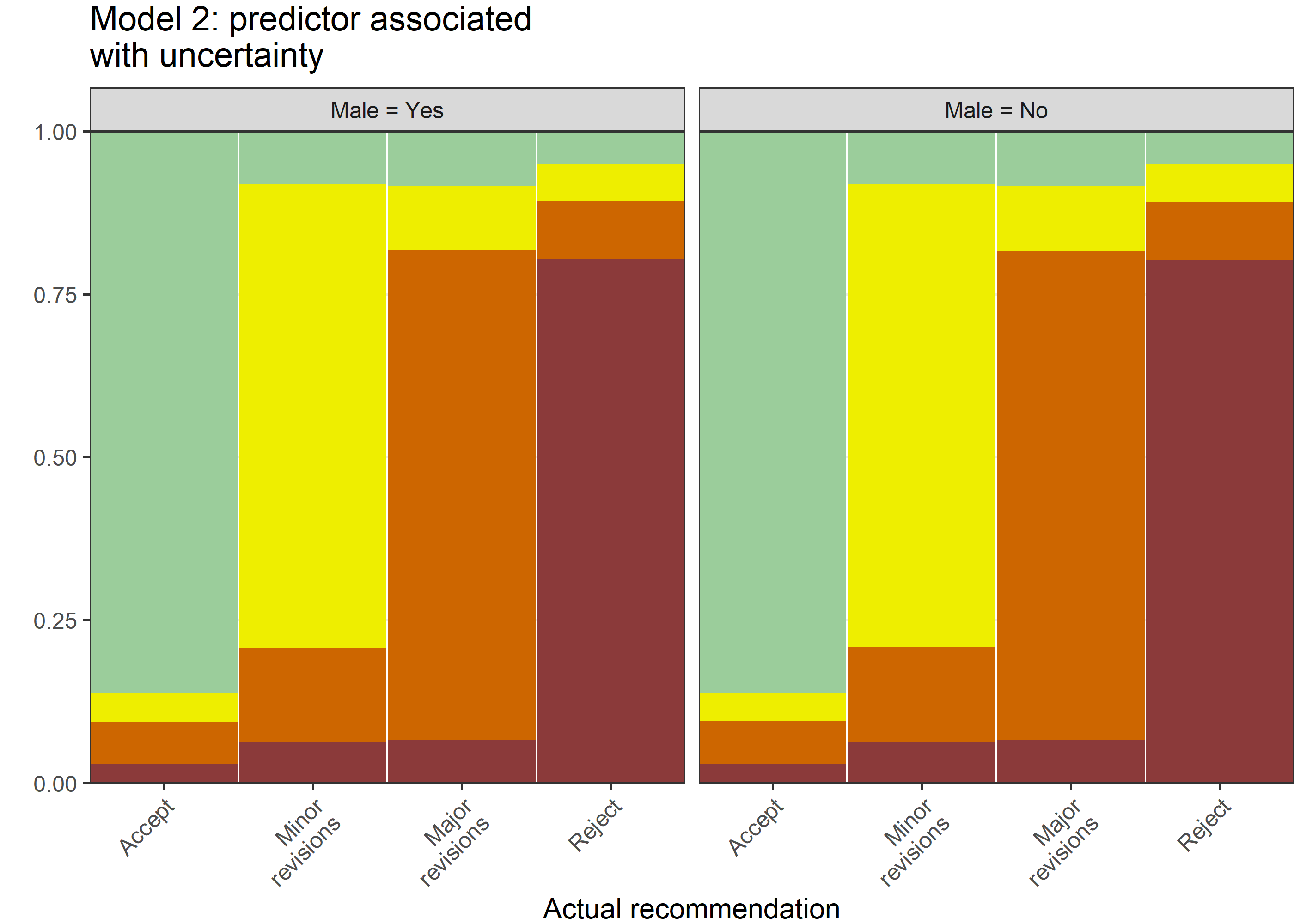
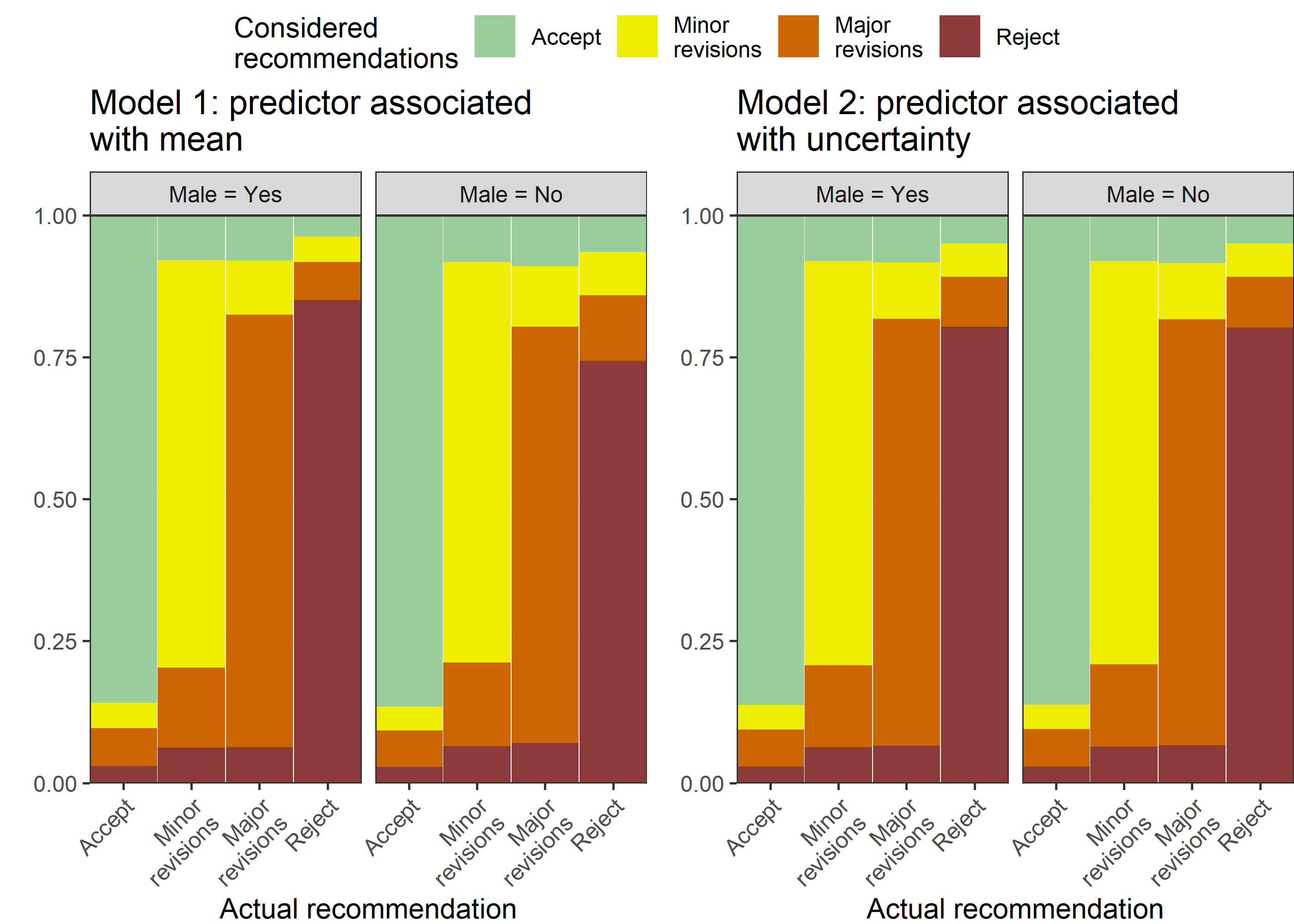


### Gender

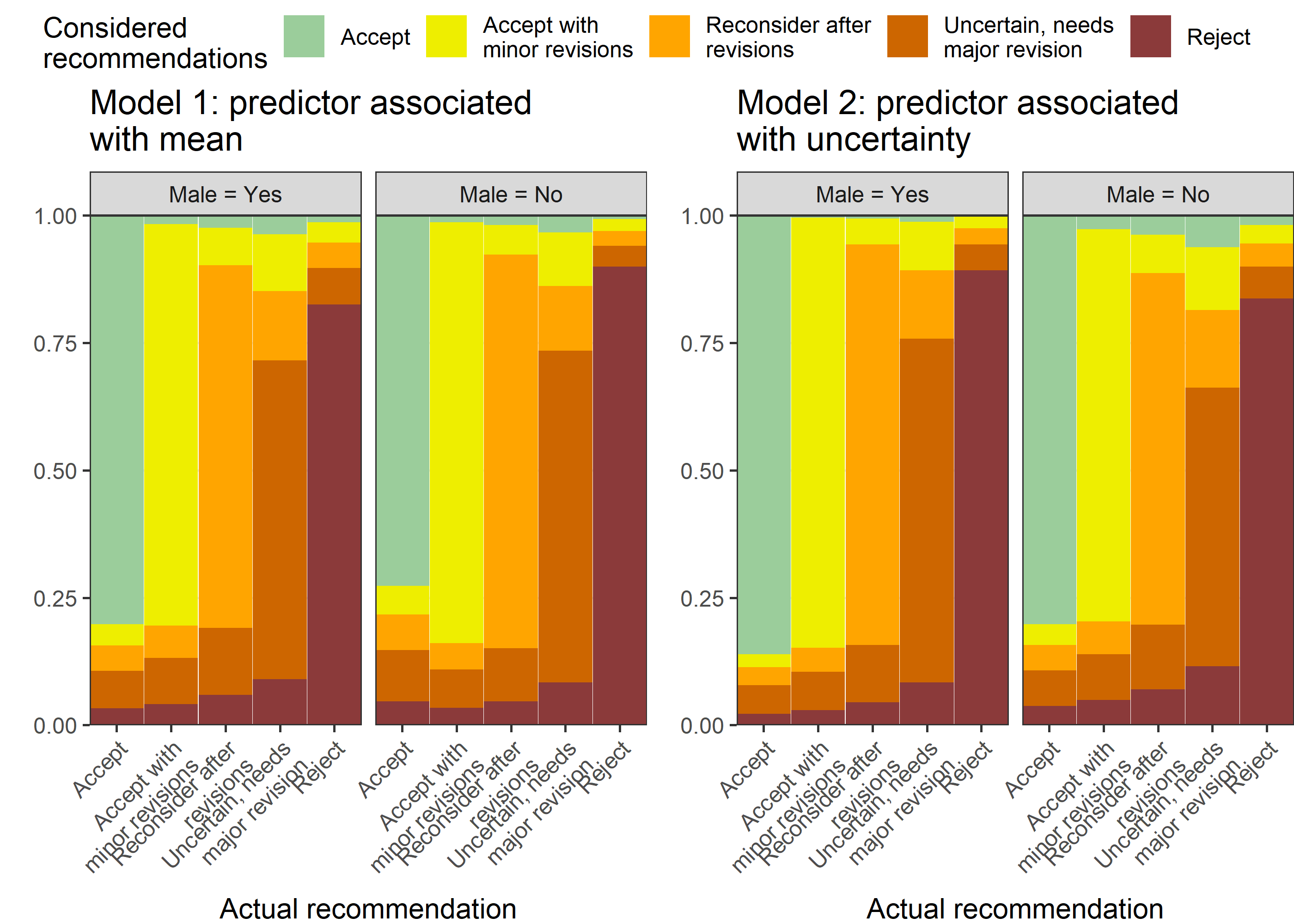
#### F1000, gender



#### BMJ Open, gender

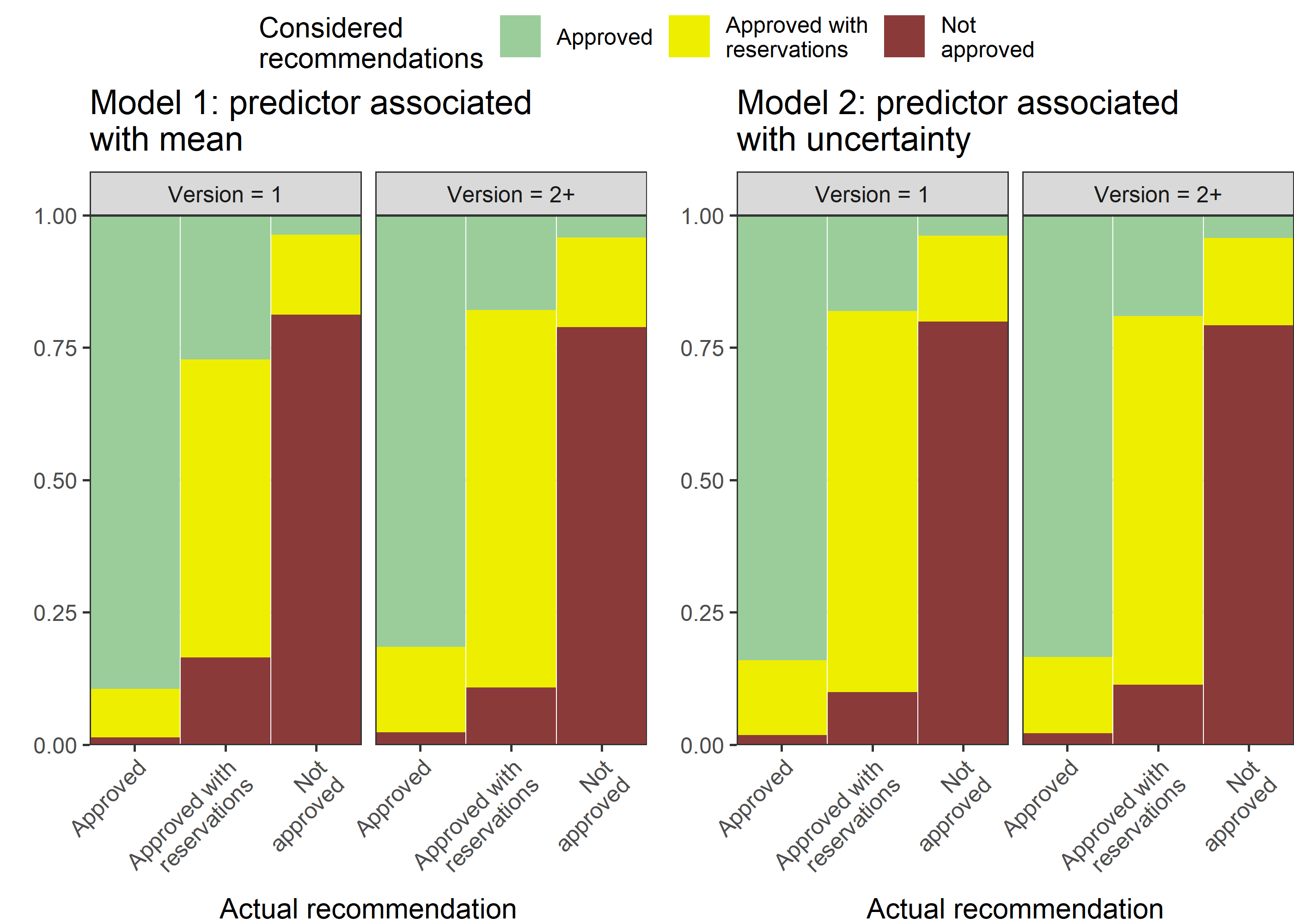


#### Epidemiology, gender

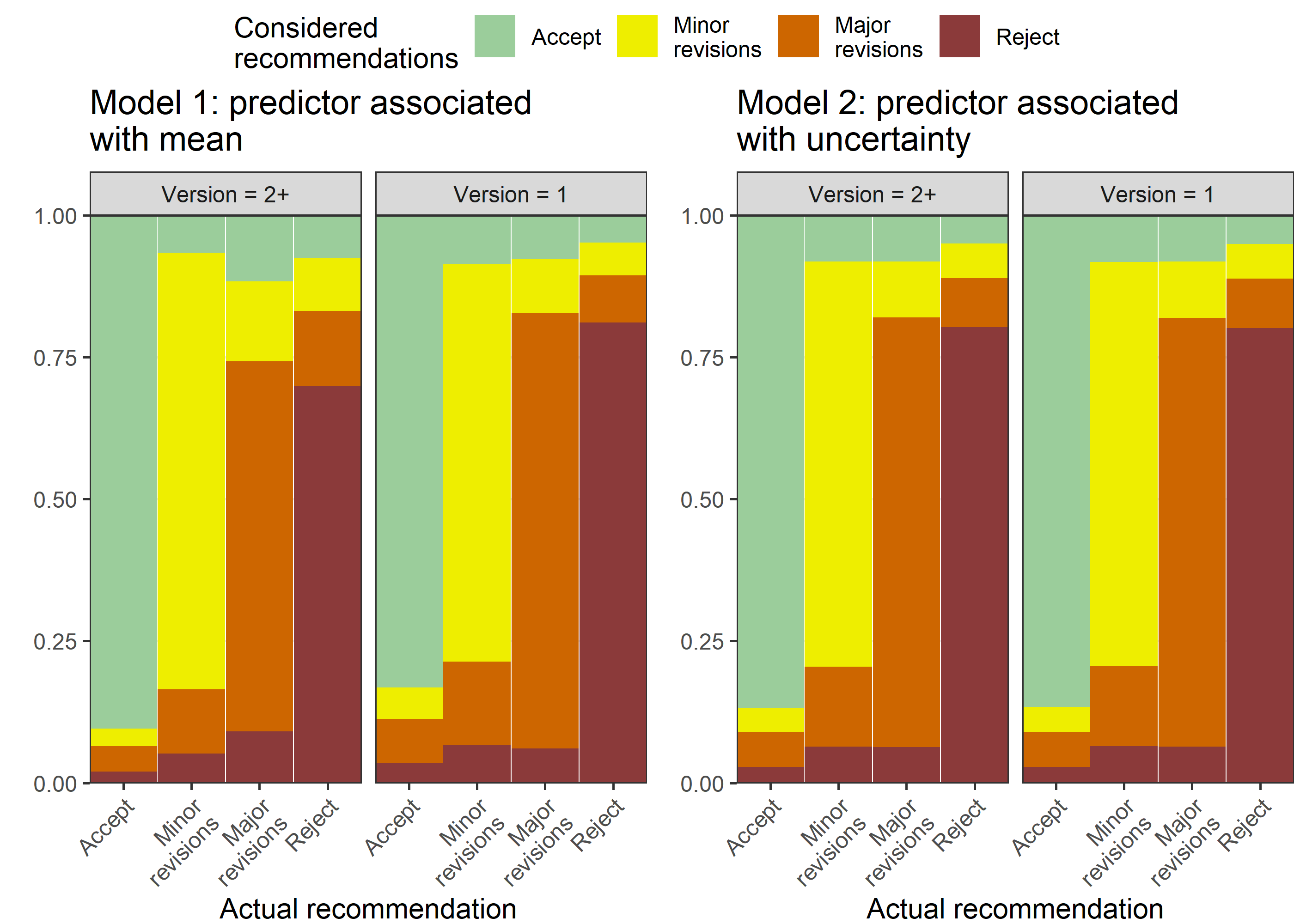


### Version

#### F1000, version

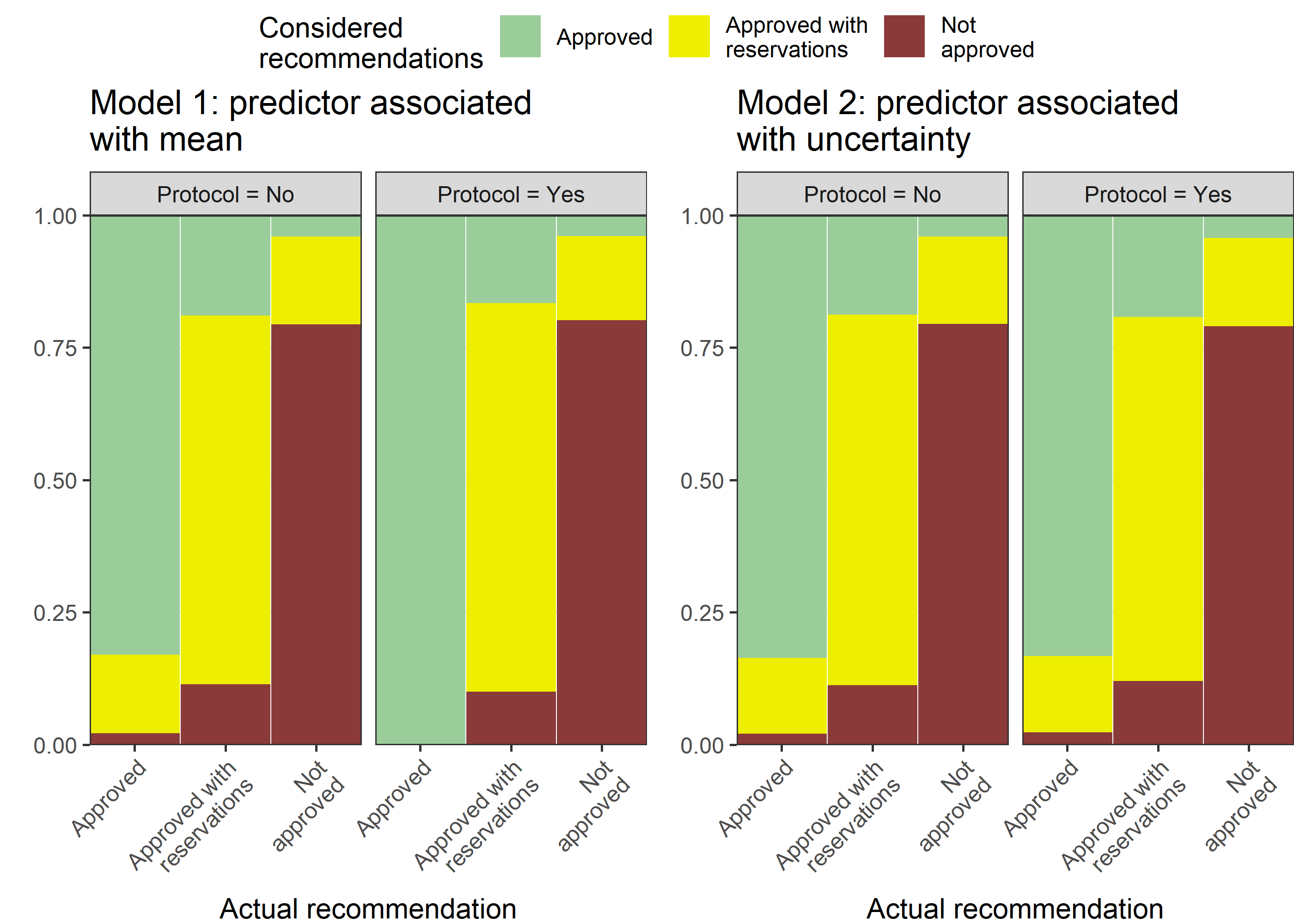


#### BMJ Open, version

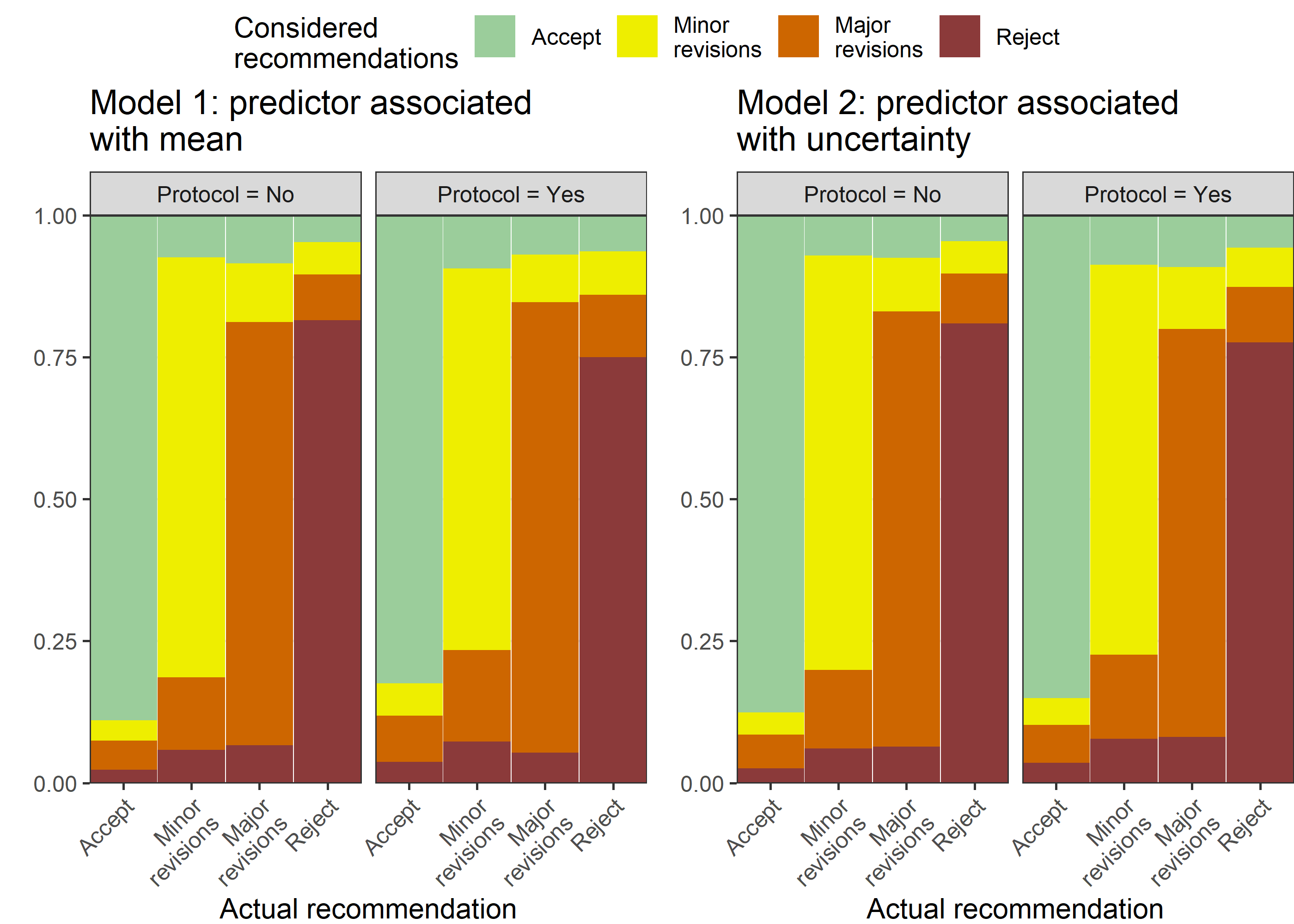


### Protocol

#### F1000, protocol

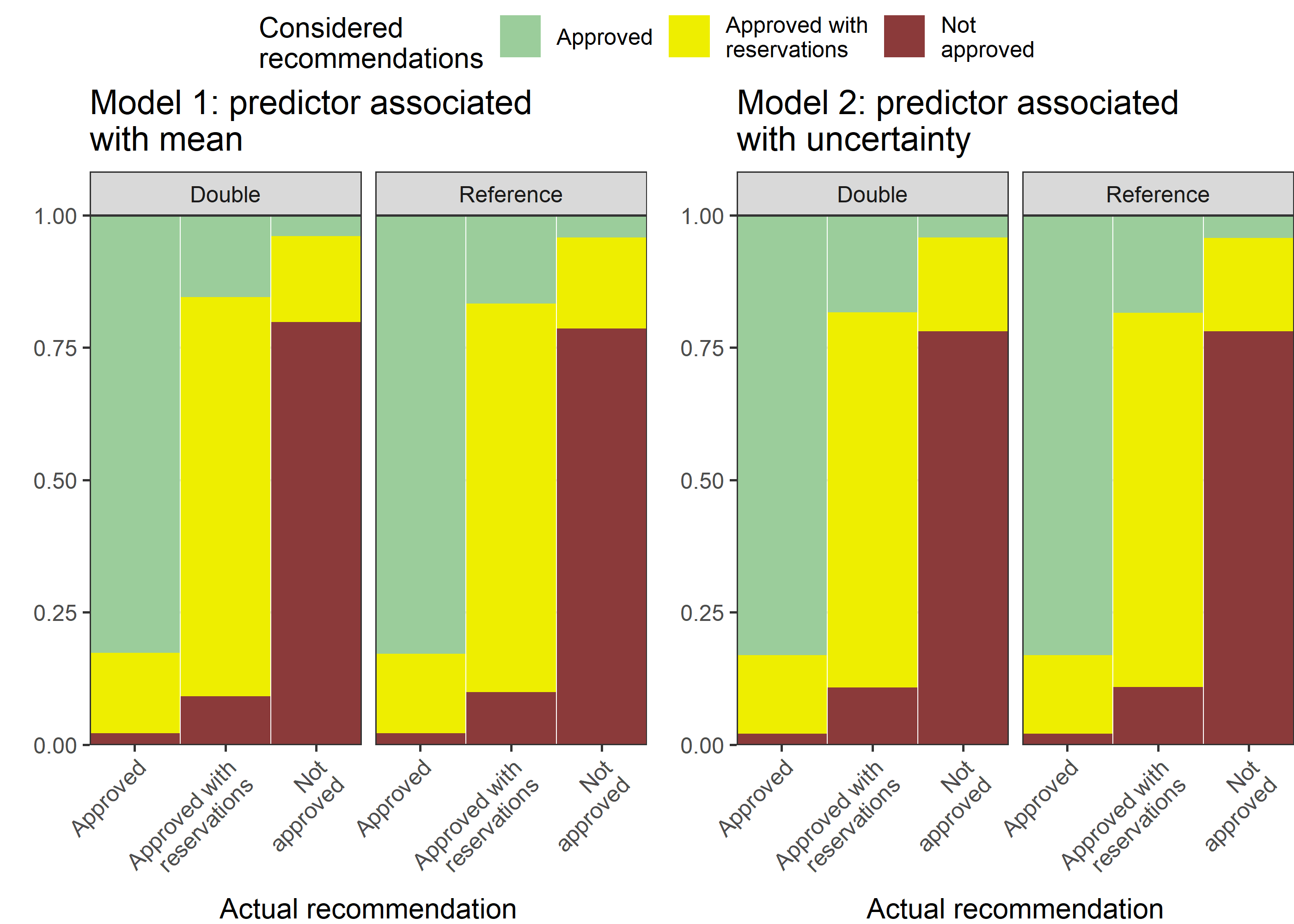


#### BMJ Open, protocol

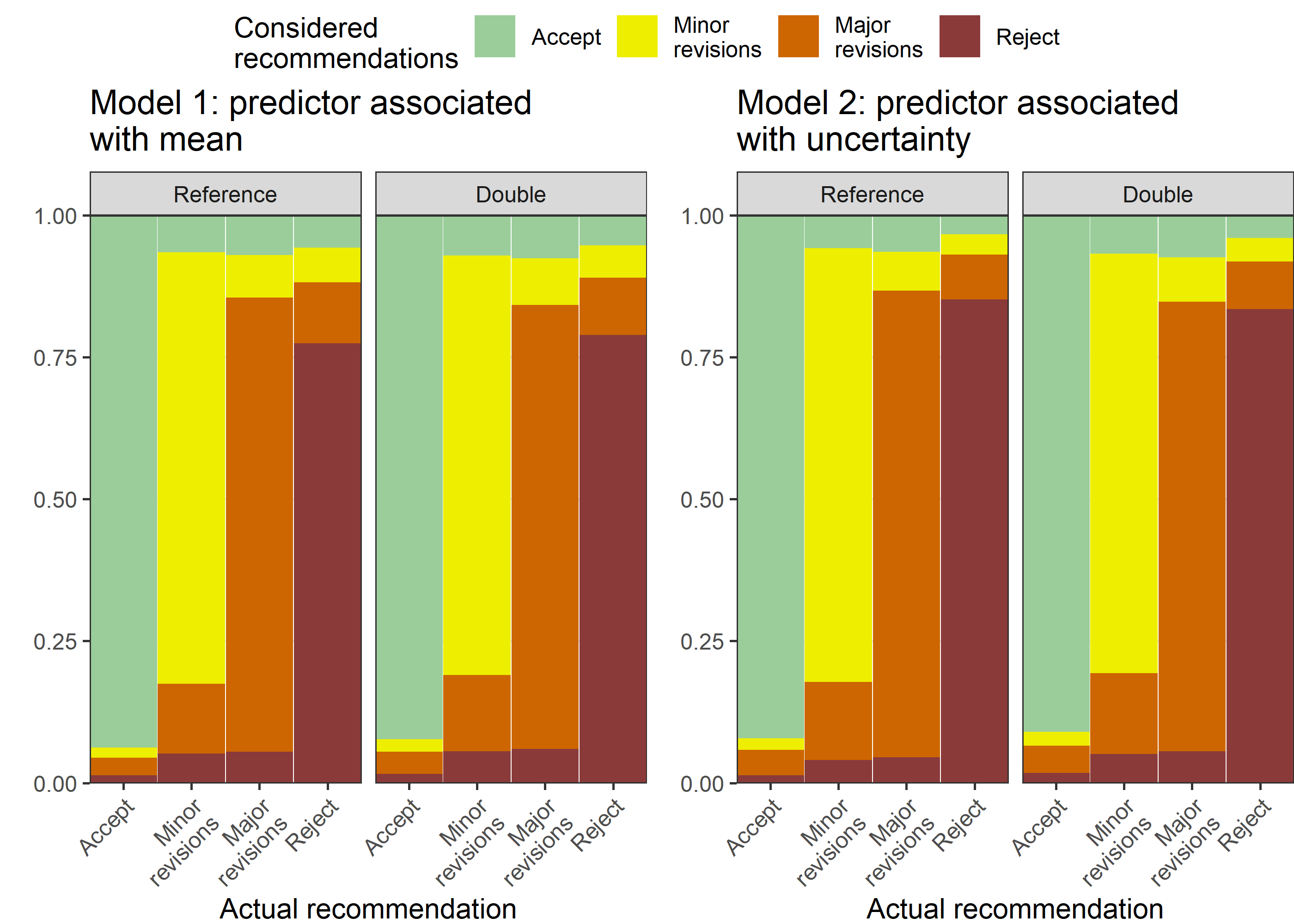


### Reviewer’s time spent

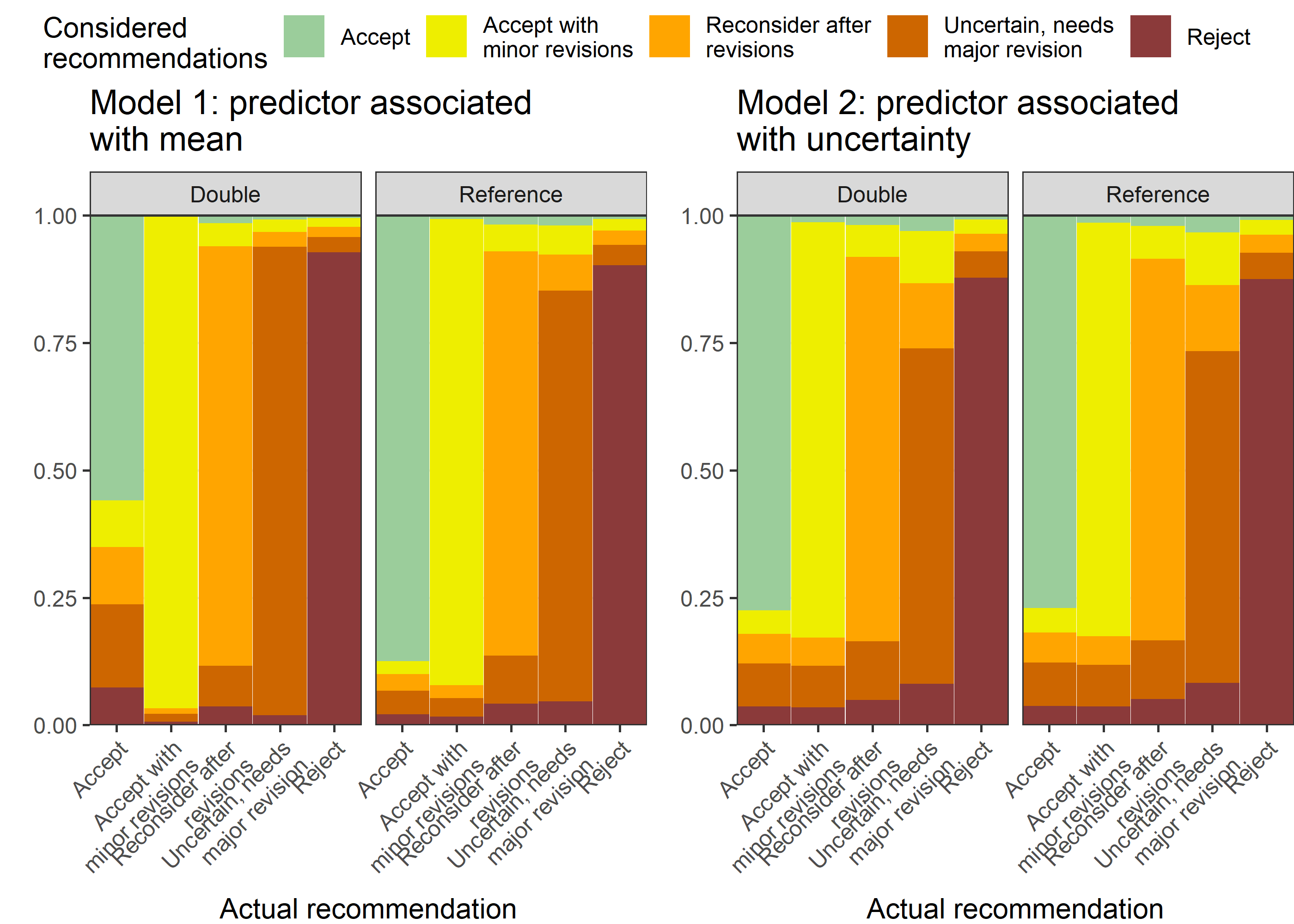
#### F1000, reviewer’s time spent



#### BMJ Open, reviewer’s time spent

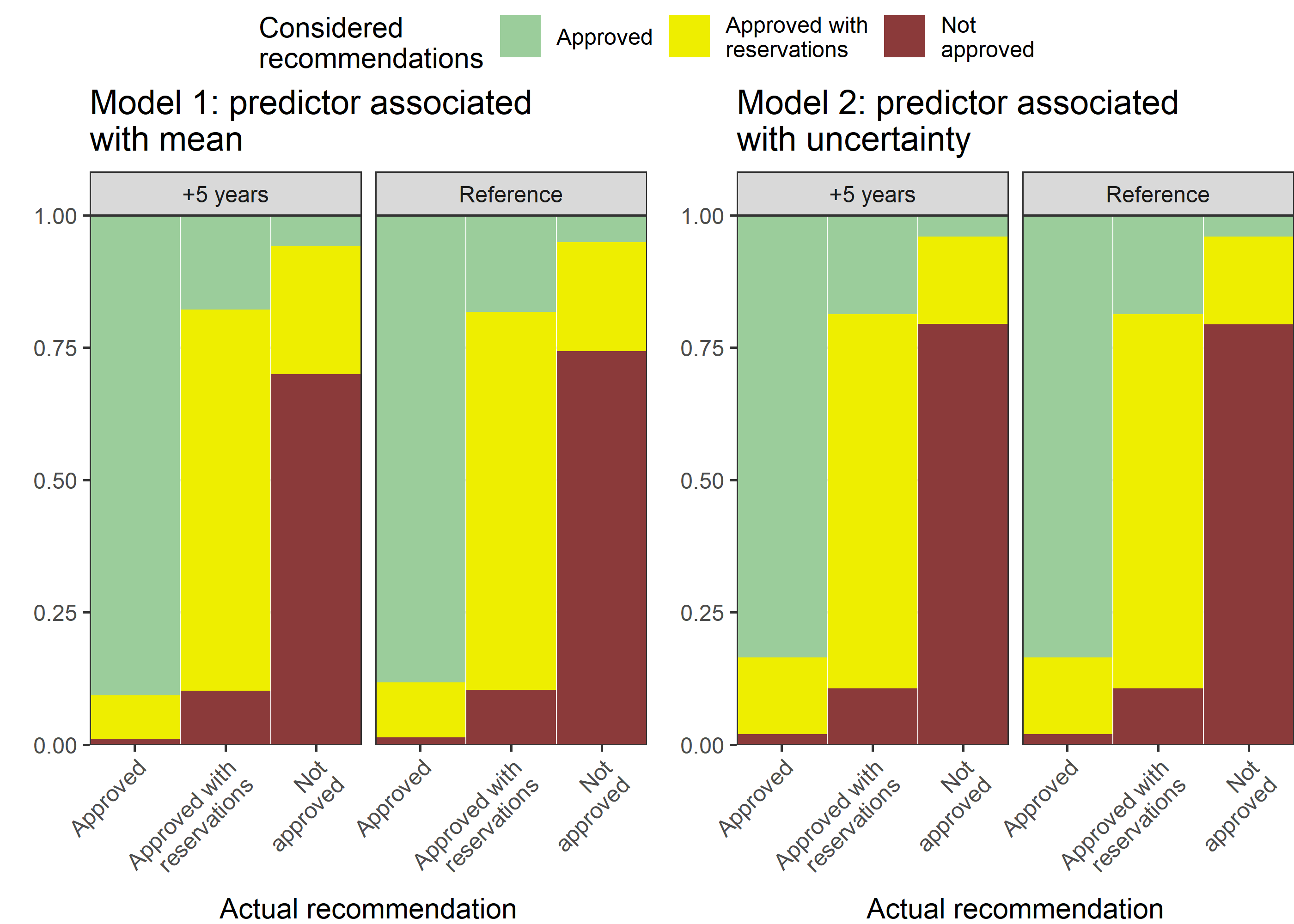


#### Epidemiology, reviewer’s time spent

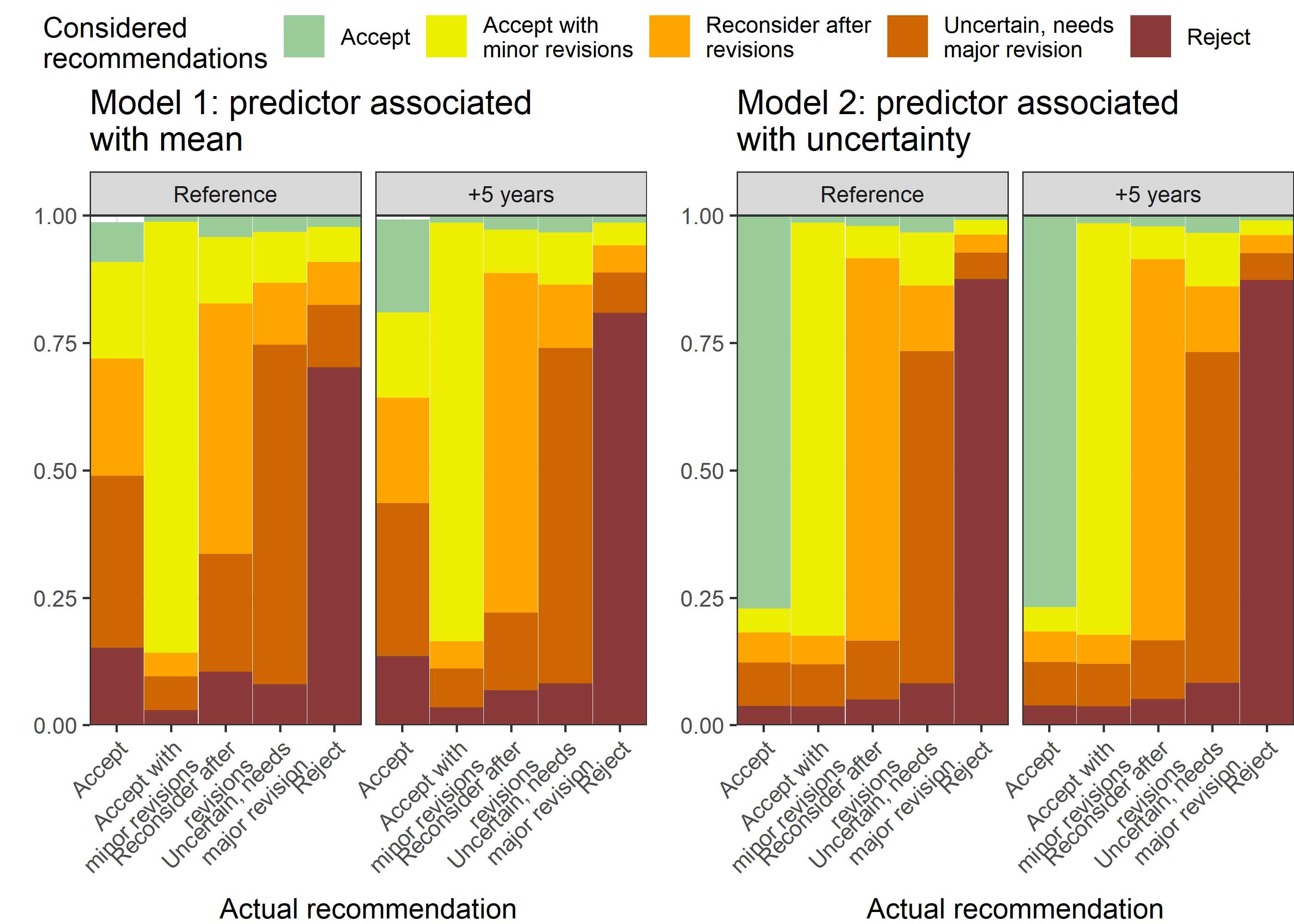


### Reviewer’s experience

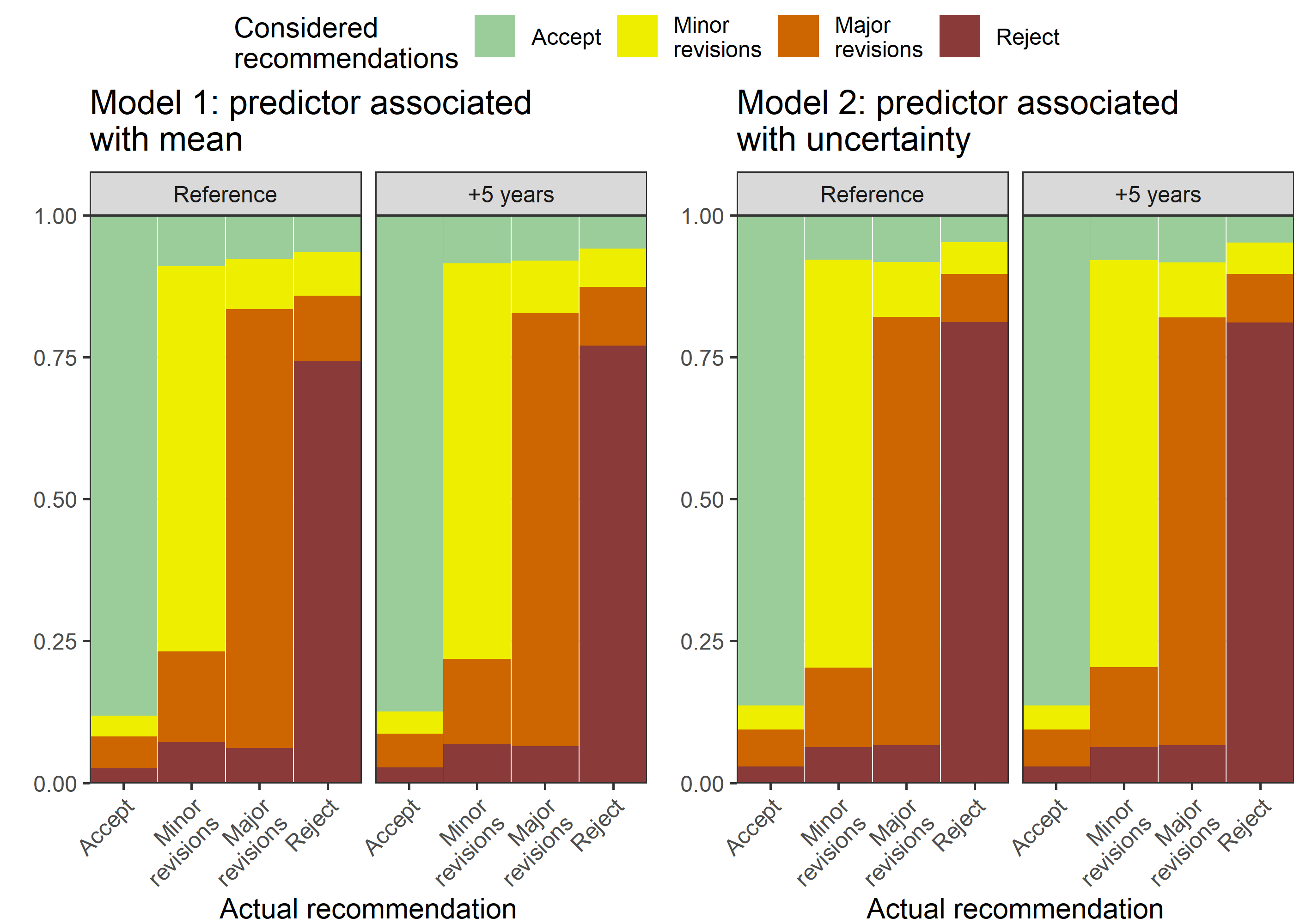
#### F1000, reviewer’s experience



#### Epidemiology, reviewer’s experience



#### BMJ Open, reviewer’s experience



## Parameter estimates

The tables below show the parameter estimates with the 95% confidence intervals and Bayesian posterior probability.

Selected models only. Complete case estimates.

### F1000, version

| **model** | **parameter** | **mean** | **lower** | **upper** | **post\_prob** |
| --- | --- | --- | --- | --- | --- |
| Model 0: No predictor | Intercept, recommendation 2 | 1.40 | 1.21 | 1.59 | <0.0001 |
| Intercept, recommendation 3 | -0.50 | -0.61 | -0.40 | <0.0001 |
| Slope, recommendation 1 | 3.17 | 2.97 | 3.36 | <0.0001 |
| Slope, recommendation 2 | -0.09 | -0.28 | 0.11 | 0.4026 |
| Slope, recommendation 3 | 3.48 | 3.28 | 3.68 | <0.0001 |
| Model 1: Predictor associated with mean | Intercept, recommendation 2 | 1.41 | 1.22 | 1.60 | <0.0001 |
| Intercept, recommendation 3 | -0.50 | -0.61 | -0.40 | <0.0001 |
| Slope, recommendation 1 | 3.68 | 3.40 | 3.96 | <0.0001 |
| Slope, recommendation 2 | -0.68 | -0.96 | -0.42 | <0.0001 |
| Slope, recommendation 3 | 3.60 | 3.31 | 3.90 | <0.0001 |
| Predictor, recommendation 2 | 0.66 | 0.45 | 0.87 | <0.0001 |
| Predictor, recommendation 3 | -0.15 | -0.44 | 0.13 | 0.2906 |
| Model 2a: Predictor associated with uncertainty | Intercept, recommendation 2 | 1.45 | 1.24 | 1.66 | <0.0001 |
| Intercept, recommendation 3 | -0.62 | -0.86 | -0.44 | <0.0001 |
| Slope, recommendation 1 | 3.23 | 3.02 | 3.45 | <0.0001 |
| Slope, recommendation 2 | -0.06 | -0.27 | 0.14 | 0.6001 |
| Slope, recommendation 3 | 3.66 | 3.38 | 4.01 | <0.0001 |
| Uncertainty | 0.13 | 0.01 | 0.33 | <0.0001 |
| Model 2b: Predictor associated with uncertainty (reverse) | Intercept, recommendation 2 | 1.44 | 1.24 | 1.64 | <0.0001 |
| Intercept, recommendation 3 | -0.49 | -0.60 | -0.38 | <0.0001 |
| Slope, recommendation 1 | 3.21 | 3.01 | 3.42 | <0.0001 |
| Slope, recommendation 2 | -0.11 | -0.31 | 0.09 | 0.2712 |
| Slope, recommendation 3 | 3.50 | 3.30 | 3.70 | <0.0001 |
| Uncertainty | 0.08 | 0.00 | 0.25 | <0.0001 |

### BMJ Open, reviewer’s time spent

| **model** | **parameter** | **mean** | **lower** | **upper** | **post\_prob** |
| --- | --- | --- | --- | --- | --- |
| Model 0: No predictor | Intercept, recommendation 2 | 0.08 | -0.01 | 0.16 | 0.10022 |
| Intercept, recommendation 3 | 0.64 | 0.57 | 0.72 | < 0.0001 |
| Intercept, recommendation 4 | -0.23 | -0.31 | -0.15 | < 0.0001 |
| Slope, recommendation 1 | 3.49 | 3.36 | 3.62 | < 0.0001 |
| Slope, recommendation 2 | 2.15 | 2.05 | 2.24 | < 0.0001 |
| Slope, recommendation 3 | 1.54 | 1.45 | 1.62 | < 0.0001 |
| Slope, recommendation 4 | 3.06 | 2.95 | 3.18 | < 0.0001 |
| Model 1: Predictor associated with mean | Intercept, recommendation 2 | 0.08 | -0.01 | 0.16 | 0.08922 |
| Intercept, recommendation 3 | 0.64 | 0.57 | 0.72 | < 0.0001 |
| Intercept, recommendation 4 | -0.23 | -0.31 | -0.15 | < 0.0001 |
| Slope, recommendation 1 | 4.04 | 3.77 | 4.31 | < 0.0001 |
| Slope, recommendation 2 | 2.39 | 2.25 | 2.52 | < 0.0001 |
| Slope, recommendation 3 | 1.80 | 1.63 | 1.96 | < 0.0001 |
| Slope, recommendation 4 | 2.85 | 2.56 | 3.14 | < 0.0001 |
| Predictor, recommendation 2 | -0.11 | -0.15 | -0.07 | < 0.0001 |
| Predictor, recommendation 3 | -0.10 | -0.16 | -0.05 | < 0.0001 |
| Predictor, recommendation 4 | 0.09 | -0.02 | 0.19 | 0.10667 |
| Model 2a: Predictor associated with uncertainty | Intercept, recommendation 2 | 0.07 | -0.07 | 0.20 | 0.32978 |
| Intercept, recommendation 3 | 0.86 | 0.74 | 0.99 | < 0.0001 |
| Intercept, recommendation 4 | -0.36 | -0.50 | -0.23 | < 0.0001 |
| Slope, recommendation 1 | 3.88 | 3.70 | 4.07 | < 0.0001 |
| Slope, recommendation 2 | 2.51 | 2.34 | 2.69 | < 0.0001 |
| Slope, recommendation 3 | 1.69 | 1.58 | 1.80 | < 0.0001 |
| Slope, recommendation 4 | 3.59 | 3.37 | 3.83 | < 0.0001 |
| Uncertainty | 0.22 | 0.14 | 0.31 | < 0.0001 |
| Model 2b: Predictor associated with uncertainty (reverse) | Intercept, recommendation 2 | 0.08 | -0.01 | 0.17 | 0.10222 |
| Intercept, recommendation 3 | 0.65 | 0.57 | 0.73 | < 0.0001 |
| Intercept, recommendation 4 | -0.23 | -0.32 | -0.15 | < 0.0001 |
| Slope, recommendation 1 | 3.50 | 3.37 | 3.63 | < 0.0001 |
| Slope, recommendation 2 | 2.15 | 2.06 | 2.25 | < 0.0001 |
| Slope, recommendation 3 | 1.54 | 1.45 | 1.63 | < 0.0001 |
| Slope, recommendation 4 | 3.08 | 2.96 | 3.20 | < 0.0001 |
| Uncertainty | 0.00 | 0.00 | 0.01 | < 0.0001 |

### Epidemiology, gender

| **model** | **parameter** | **mean** | **lower** | **upper** | **post\_prob** |
| --- | --- | --- | --- | --- | --- |
| Model 0: No predictor | Intercept, recommendation 2 | 1.13 | 0.88 | 1.40 | < 0.0001 |
| Intercept, recommendation 3 | 1.34 | 1.08 | 1.61 | < 0.0001 |
| Intercept, recommendation 4 | 1.71 | 1.46 | 1.97 | < 0.0001 |
| Intercept, recommendation 5 | 0.92 | 0.65 | 1.19 | < 0.0001 |
| Slope, recommendation 1 | 4.11 | 3.58 | 4.68 | < 0.0001 |
| Slope, recommendation 2 | 2.85 | 2.61 | 3.09 | < 0.0001 |
| Slope, recommendation 3 | 2.24 | 2.03 | 2.46 | < 0.0001 |
| Slope, recommendation 4 | 1.23 | 1.01 | 1.44 | < 0.0001 |
| Slope, recommendation 5 | 3.60 | 3.36 | 3.85 | < 0.0001 |
| Model 1: Predictor associated with mean | Intercept, recommendation 2 | 1.14 | 0.88 | 1.42 | < 0.0001 |
| Intercept, recommendation 3 | 1.34 | 1.08 | 1.63 | < 0.0001 |
| Intercept, recommendation 4 | 1.72 | 1.46 | 1.99 | < 0.0001 |
| Intercept, recommendation 5 | 0.92 | 0.66 | 1.20 | < 0.0001 |
| Slope, recommendation 1 | 4.12 | 3.59 | 4.68 | < 0.0001 |
| Slope, recommendation 2 | 2.74 | 2.45 | 3.03 | < 0.0001 |
| Slope, recommendation 3 | 2.06 | 1.80 | 2.32 | < 0.0001 |
| Slope, recommendation 4 | 1.15 | 0.82 | 1.48 | < 0.0001 |
| Slope, recommendation 5 | 3.24 | 2.93 | 3.55 | < 0.0001 |
| Predictor, recommendation 2 | 0.24 | -0.12 | 0.59 | 0.1870000 |
| Predictor, recommendation 3 | 0.32 | 0.06 | 0.58 | 0.0148889 |
| Predictor, recommendation 4 | 0.11 | -0.23 | 0.43 | 0.5283889 |
| Predictor, recommendation 5 | 0.64 | 0.28 | 0.99 | 0.0002222 |
| Model 2a: Predictor associated with uncertainty | Intercept, recommendation 2 | 1.16 | 0.89 | 1.43 | < 0.0001 |
| Intercept, recommendation 3 | 1.36 | 1.08 | 1.63 | < 0.0001 |
| Intercept, recommendation 4 | 1.73 | 1.47 | 2.01 | < 0.0001 |
| Intercept, recommendation 5 | 0.93 | 0.66 | 1.22 | < 0.0001 |
| Slope, recommendation 1 | 4.13 | 3.60 | 4.70 | < 0.0001 |
| Slope, recommendation 2 | 2.85 | 2.61 | 3.09 | < 0.0001 |
| Slope, recommendation 3 | 2.24 | 2.02 | 2.46 | < 0.0001 |
| Slope, recommendation 4 | 1.23 | 1.01 | 1.45 | < 0.0001 |
| Slope, recommendation 5 | 3.61 | 3.36 | 3.86 | < 0.0001 |
| Uncertainty | 0.03 | 0.00 | 0.10 | < 0.0001 |
| Model 2b: Predictor associated with uncertainty (reverse) | Intercept, recommendation 2 | 2.10 | 1.57 | 2.68 | < 0.0001 |
| Intercept, recommendation 3 | 2.45 | 1.91 | 3.03 | < 0.0001 |
| Intercept, recommendation 4 | 2.91 | 2.38 | 3.50 | < 0.0001 |
| Intercept, recommendation 5 | 1.98 | 1.44 | 2.56 | < 0.0001 |
| Slope, recommendation 1 | 5.65 | 4.92 | 6.42 | < 0.0001 |
| Slope, recommendation 2 | 3.23 | 2.95 | 3.51 | < 0.0001 |
| Slope, recommendation 3 | 2.40 | 2.16 | 2.64 | < 0.0001 |
| Slope, recommendation 4 | 1.15 | 0.92 | 1.39 | < 0.0001 |
| Slope, recommendation 5 | 3.81 | 3.54 | 4.08 | < 0.0001 |
| Uncertainty | 6.71 | 3.22 | 12.53 | < 0.0001 |