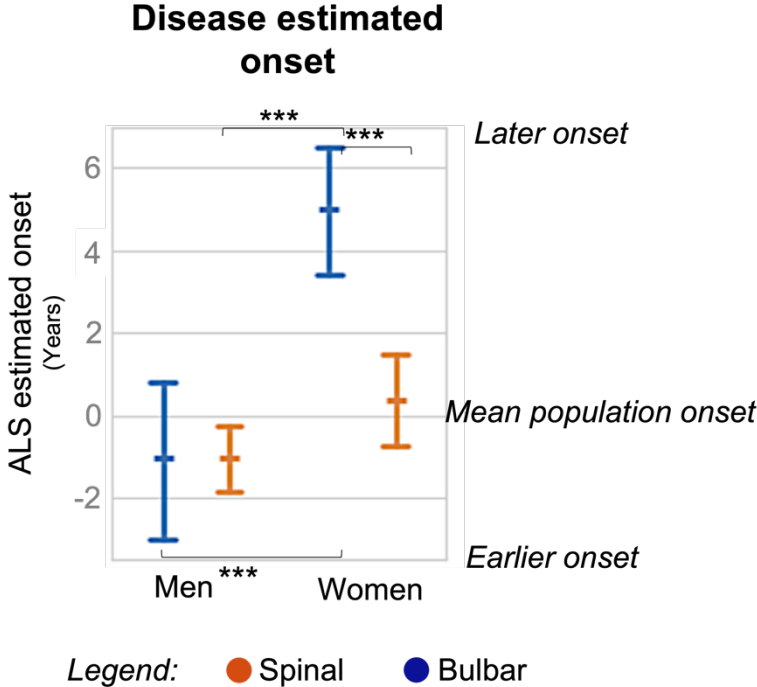


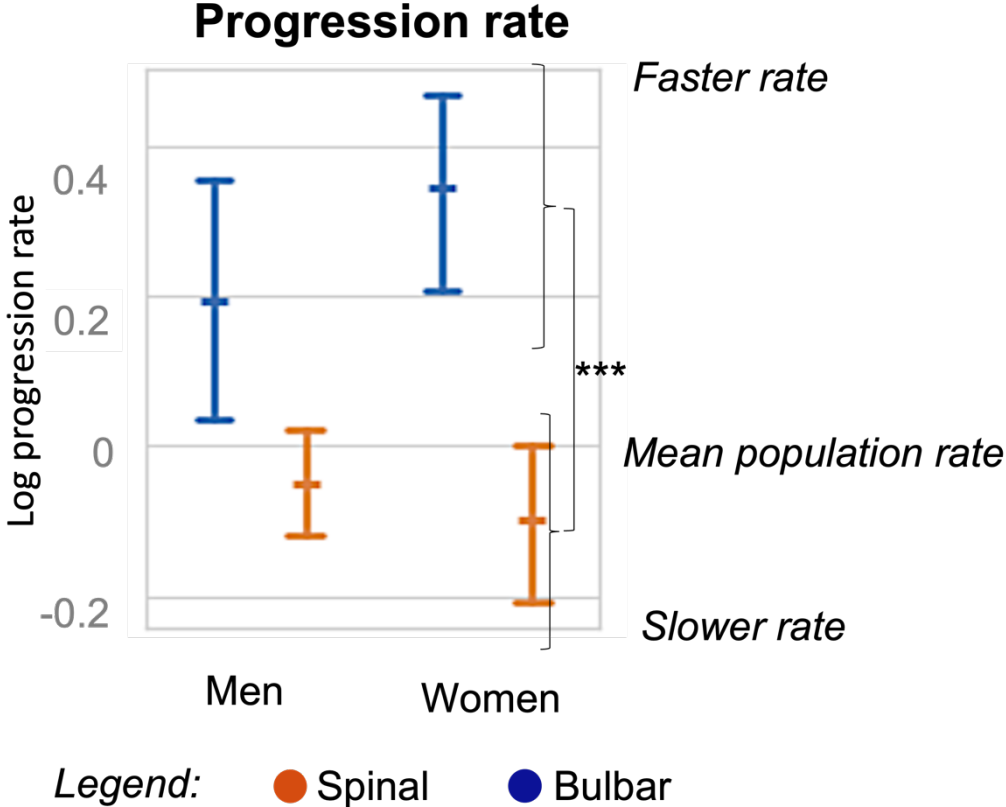
eFigure 1: ALS estimated onset depending on sex and onset site

**Legend:** The vertical axis presents the delay of ALS impairment onset in years compared to the mean onset of the whole population. The horizontal axis is for sex with estimation for the bulbar patients in blue and the spinal patients in orange. The graph displays the mean of this delay with its confidence interval. Results are corrected for the group rate. ANOVA interaction p-value is equal to 0.002. \*\*\*: p-value for pairwise comparison <0.001



eFigure 2: ALS rate depending on sex and onset site

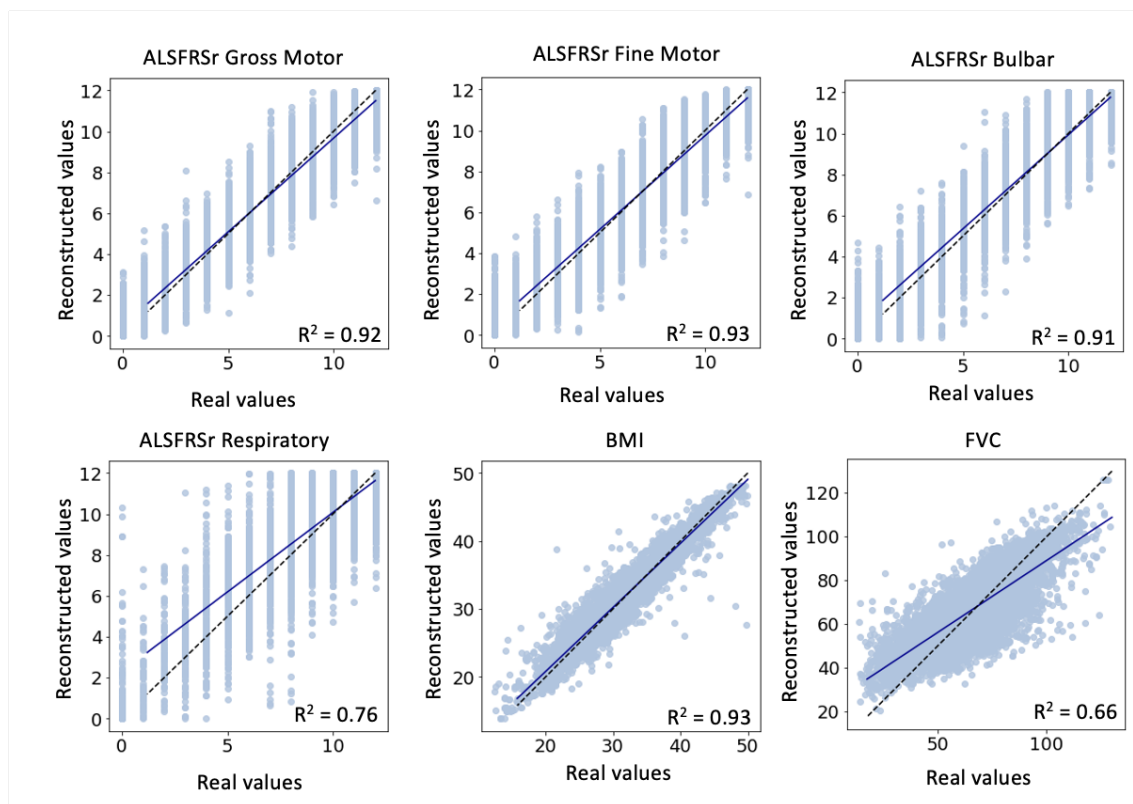
**Legend:** The vertical axis presents the log rate compared to the mean log rate of the whole population. The horizontal axis is for sex with estimation for the bulbar patients in blue and the spinal patients in orange. The graph displays the mean of this delay with its confidence interval. Results are corrected for estimated age at onset. ANOVA interaction p-value is equal to 0.14. \*\*\*: p-value for pairwise comparison <0.001



eFigure 3: True against reconstructed values,

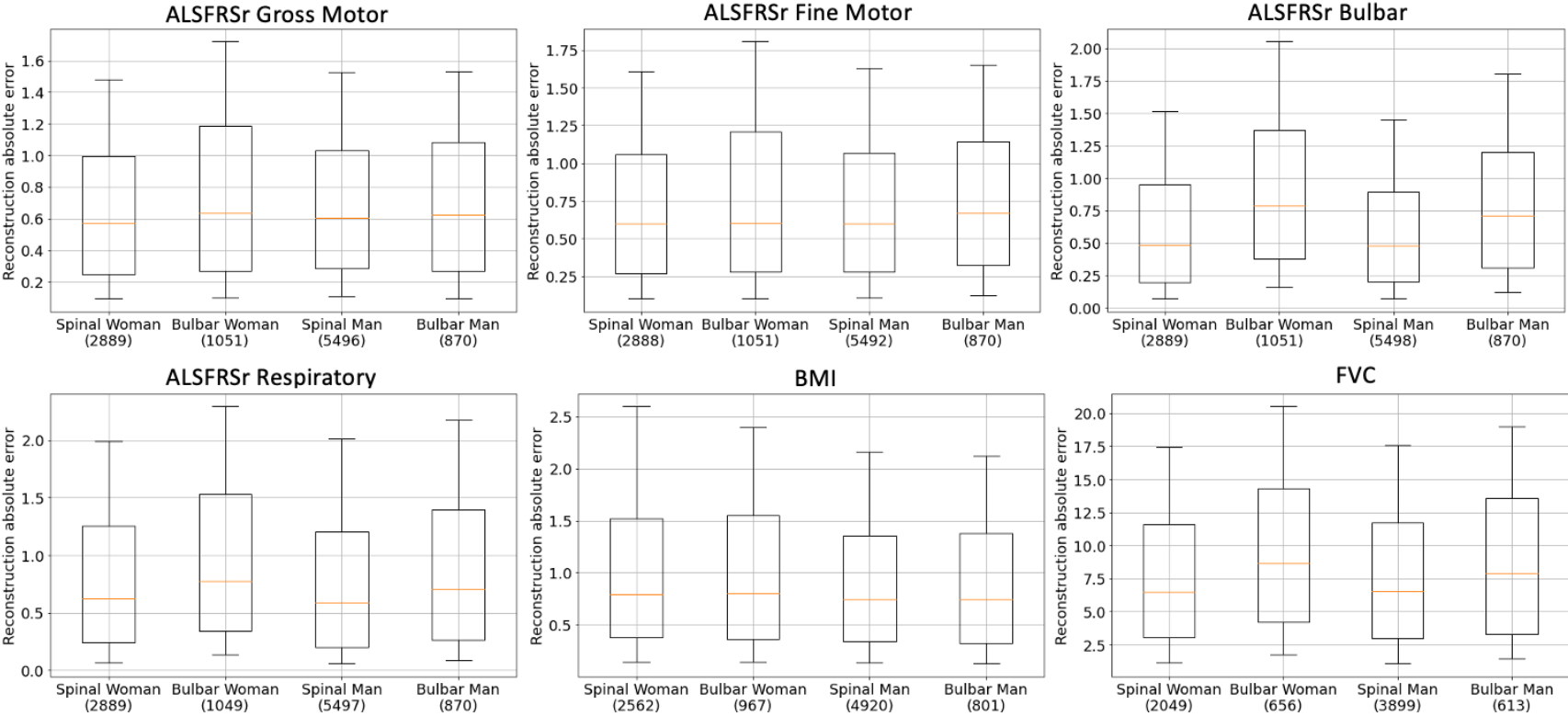
Legend: dashed black line: bisector ( $y=x$ ), blue continuous line: regression line

Interpretation: We can observe that the dots are around the dashed line, which means that the reconstructions are correct. Thanks to the regression line, we can study bias depending on the value of the outcomes. We found that the model tends to overestimate ALSFRSr respiratory (the regression line is above the dashed line). We also found that the model tends to underestimate high values of FVC and overestimate small values of FVC.



eFigure 4: Reconstruction error distribution

Interpretation: We can observe small variations of the median (orange line), smaller than the range of variation of the outcomes considered: maximum variation of median between groups of 2.18 for FVC, 0.05 for BMI, 0.19 respiratory, 0.31 bulbar, 0.08 fine motor and 0.07 gross motor.



## *eReproducibility*

The code developed to conduct this study is publicly available on GitLab at [https://gitlab.com/JulietteOrtholand/als\\_sex\\_subtype\\_article](https://gitlab.com/JulietteOrtholand/als_sex_subtype_article). The models were computed with 20,000 iterations. Note that it uses a private library leaspype that contains a wrapper of a public library available online at <https://gitlab.com/icm-institute/aramislab/leaspy> as well as functions to format PRO-ACT database. Restricted access to the private library could be asked by email to [juliette.ortholand@icm-institute.org](mailto:juliette.ortholand@icm-institute.org).