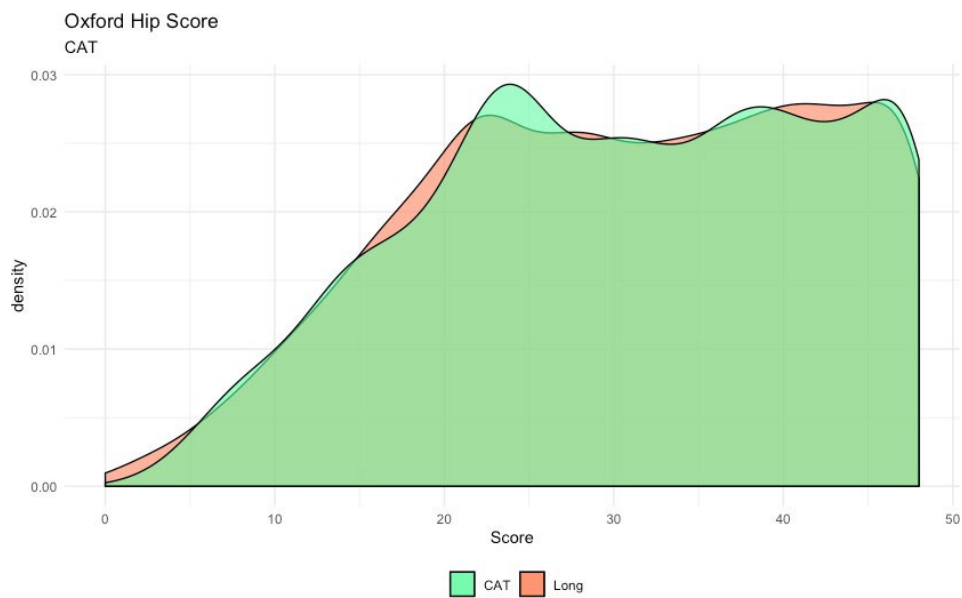




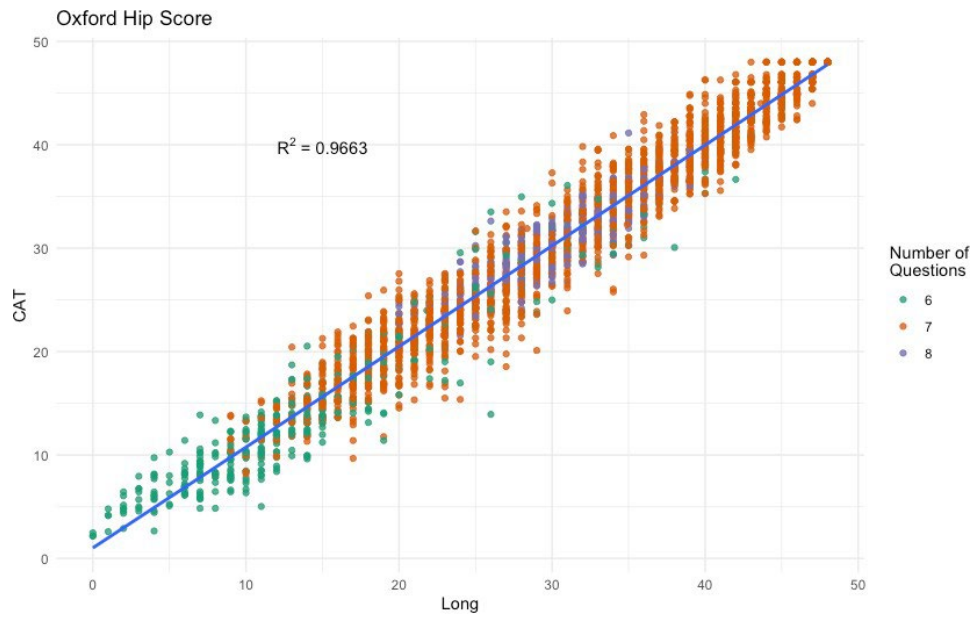
## Supplementary Material

10.1302/2633-1462.310.BJO-2022-0073.R1

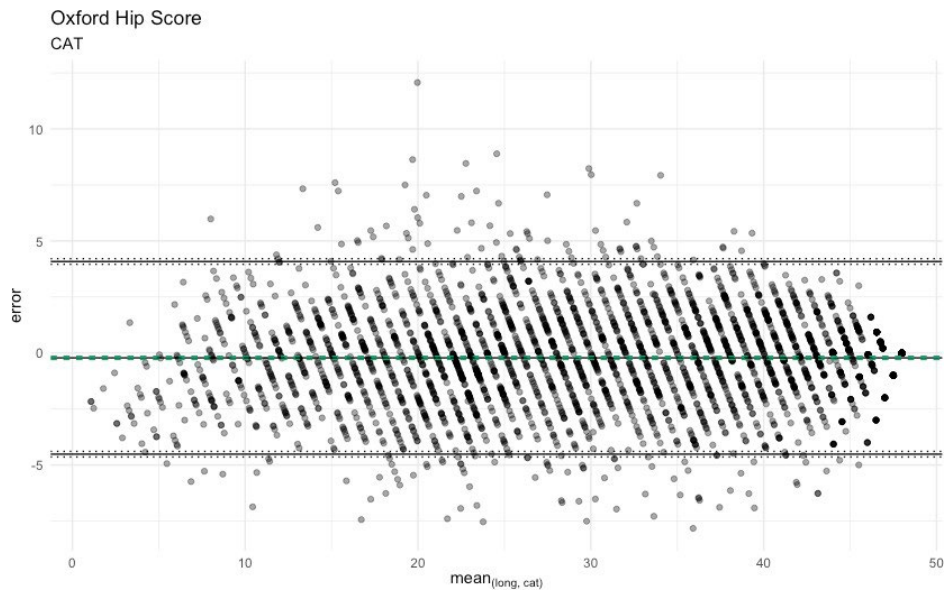
### Oxford Hip Score



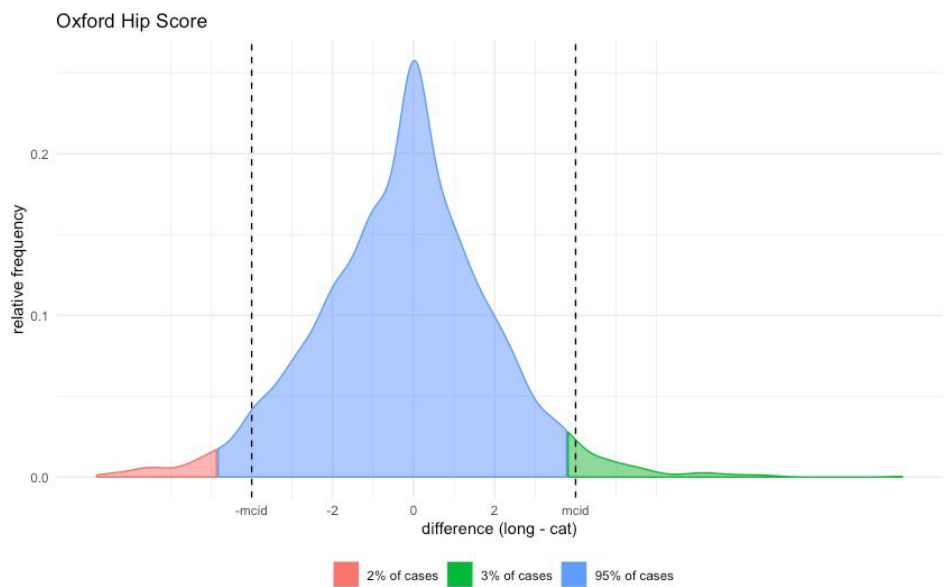
**Fig a.** Score distributions for the full-length ('Long') Oxford Hip Score and computerized adaptive testing (CAT) version of the Oxford Hip Score.



**Fig b.** Correlation of scores from the full-length ('Long') Oxford Hip Score and computerized adaptive testing (CAT) version of the Oxford Hip Score. Point colour indicates the number of questions used by the CAT algorithm in that instance.

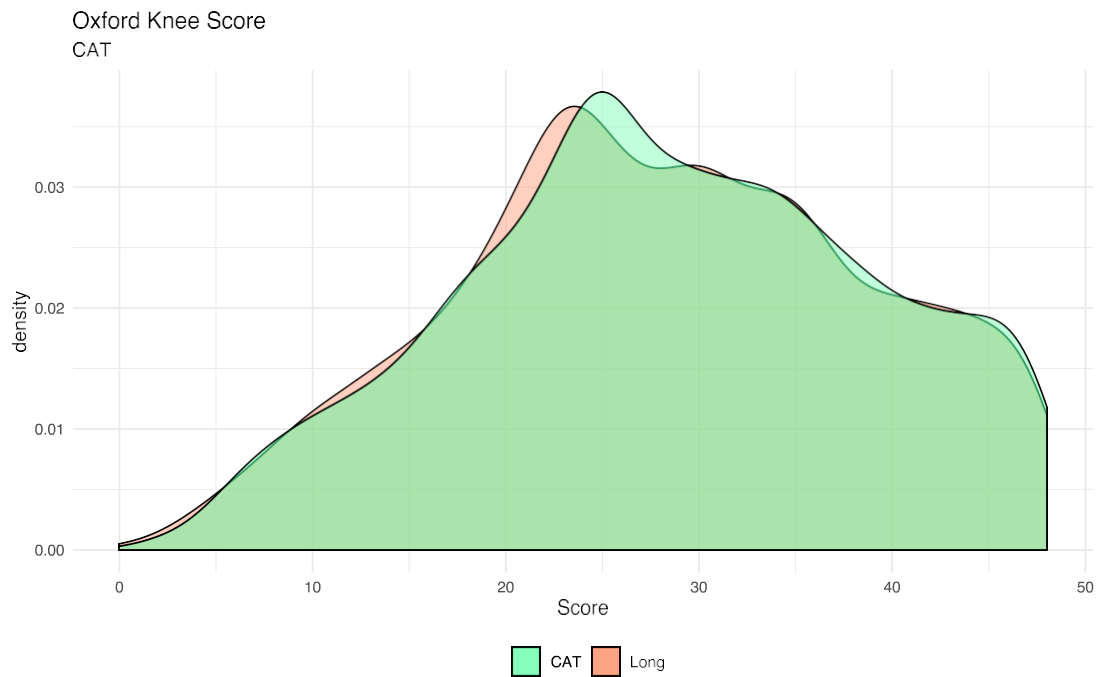


**Fig c.** Bland Altman plot demonstrating the concordance of scores from the full-length ('long') and computerized adaptive testing (CAT) versions of the Oxford Hip Score. For each pair of scores, the x-axis represents the mean of both scores and the y-axis represents the difference between them. The outermost, horizontal solid lines represent the 95% limits of agreement and the surrounding dotted lines represent their 95% confidence intervals.

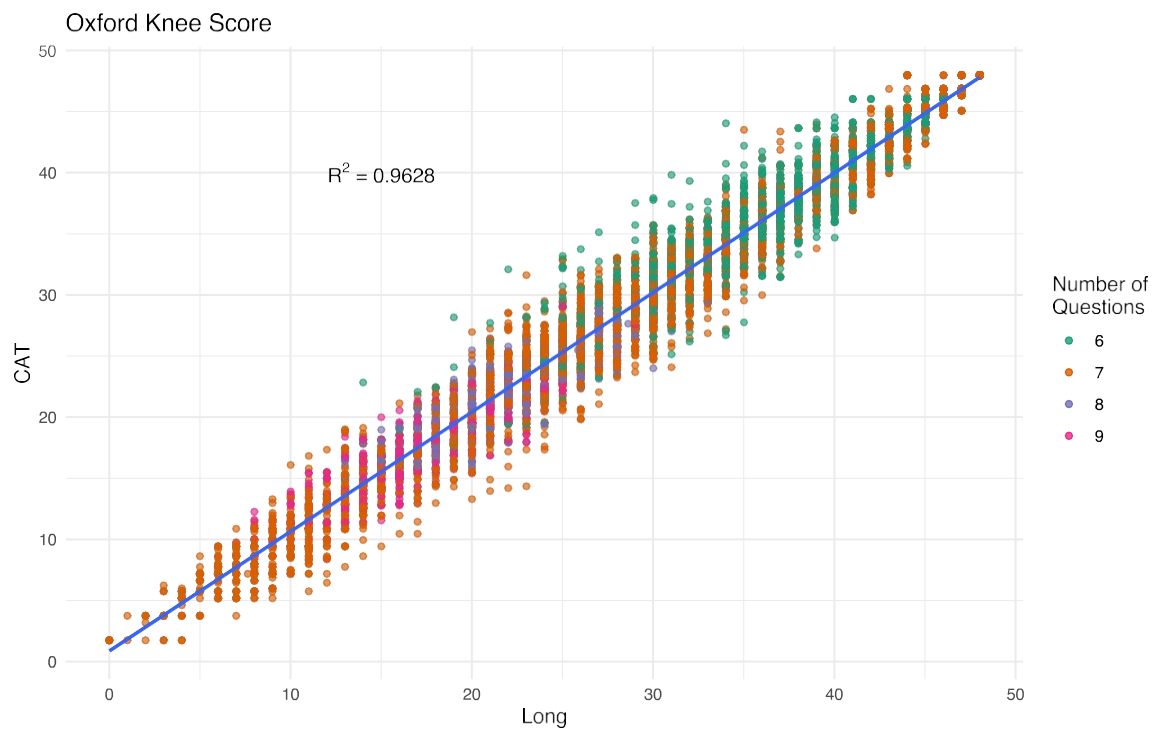


**Fig d.** Difference in score between the full-length ('long') Oxford Hip Score and its computerized adaptive testing (CAT) counterpart. Dashed horizontal lines represent the instrument's minimal clinically important difference (MCID).

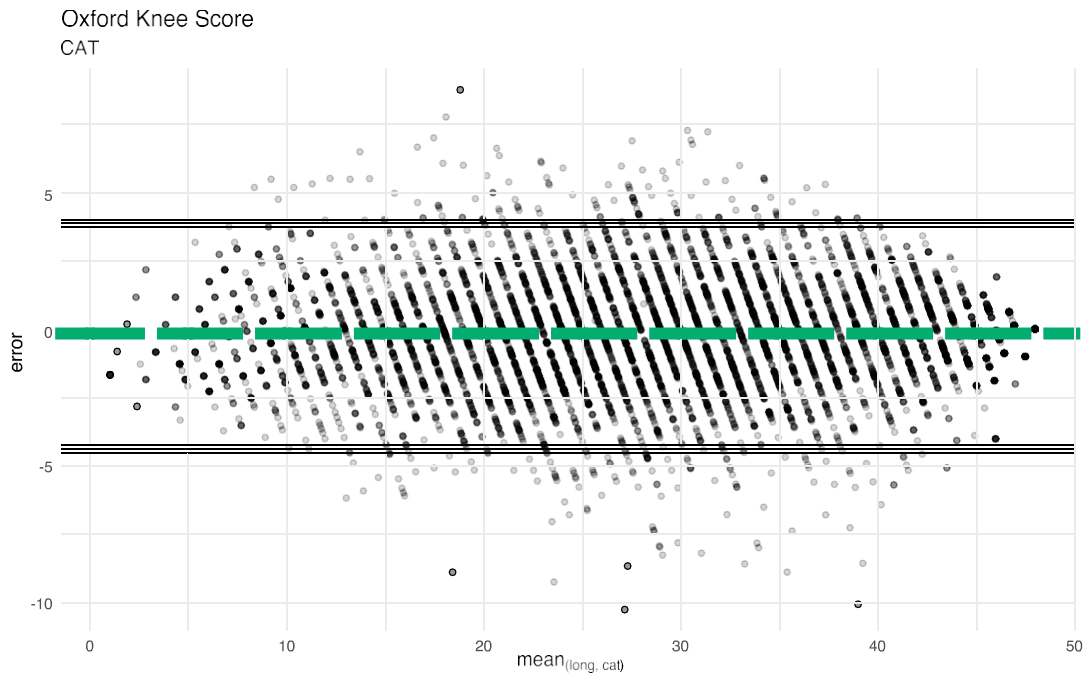
## Oxford Knee Score



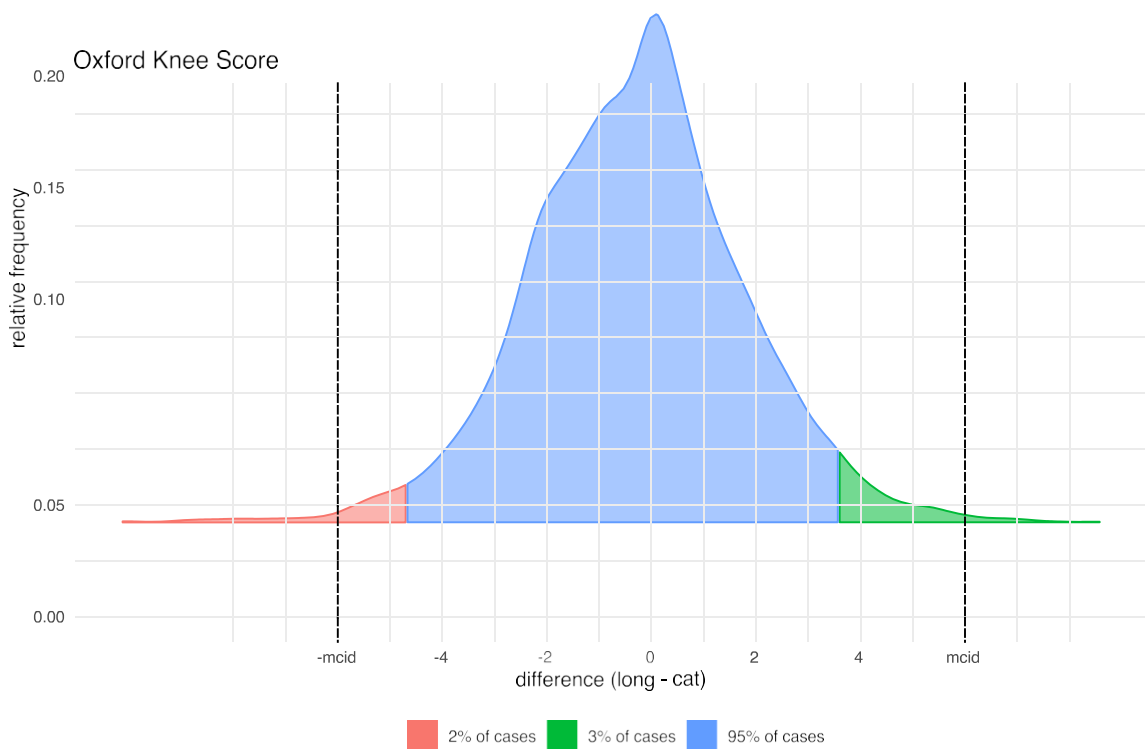
**Fig e.** Score distributions for the full-length ('Long') Oxford Knee Score and computerized adaptive testing (CAT) version of the Oxford Knee Score.



**Fig f.** Correlation of scores from the full-length ('Long') Oxford Knee Score and computerized adaptive testing (CAT) version of the Oxford Knee Score. Point colour indicates the number of questions used by the CAT algorithm in that instance.

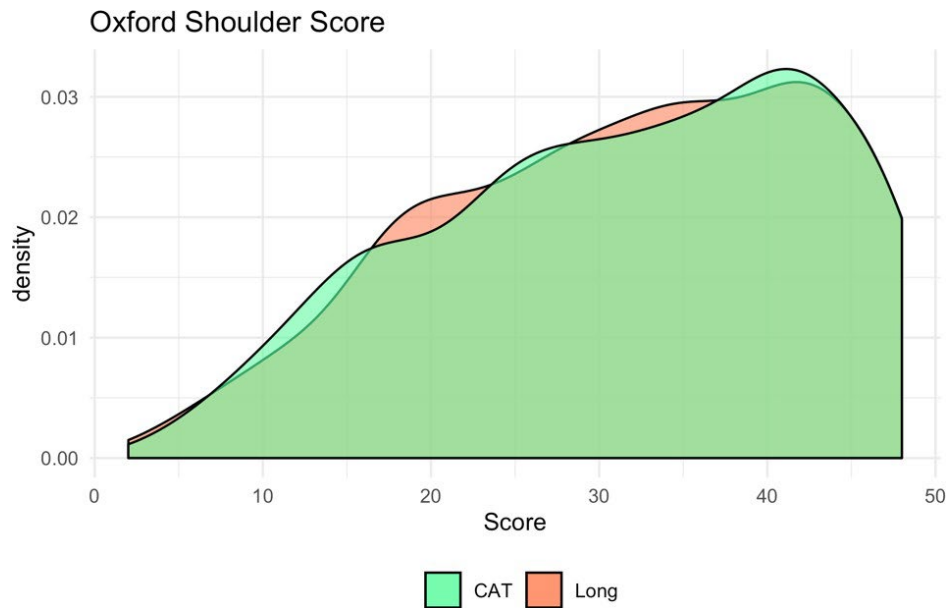


**Fig g.** Bland Altman plot demonstrating the concordance of scores from the full-length ('long') and computerized adaptive testing (CAT) versions of the Oxford Knee Score. For each pair of scores, the x-axis represents the mean of both scores and the y-axis represents the difference between them. The outermost, horizontal solid lines represent the 95% limits of agreement and the surrounding dotted lines represent their 95% confidence intervals.

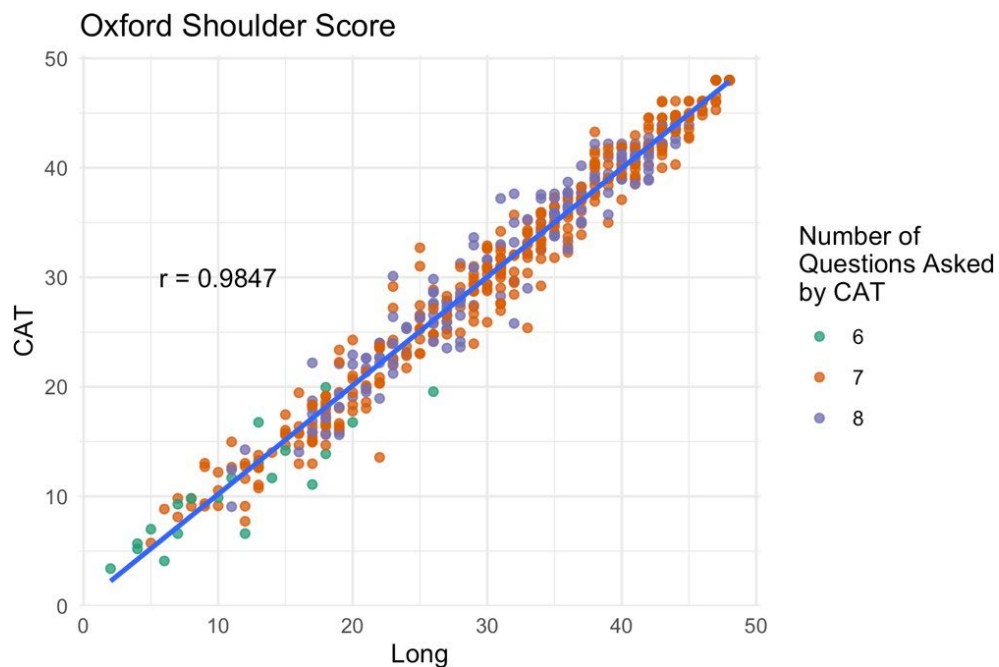


**Fig h.** Difference in score between the full-length ('long') Oxford Knee Score and its computerized adaptive testing (CAT) counterpart. Dashed horizontal lines represent the instrument's minimal clinically important difference (MCID).

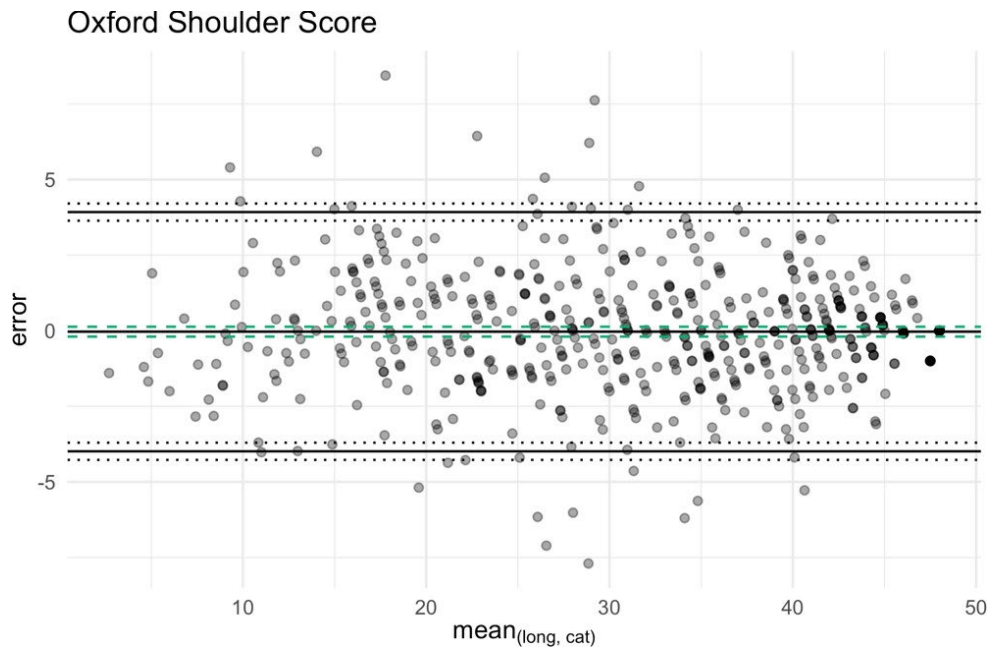
## Oxford Shoulder Score



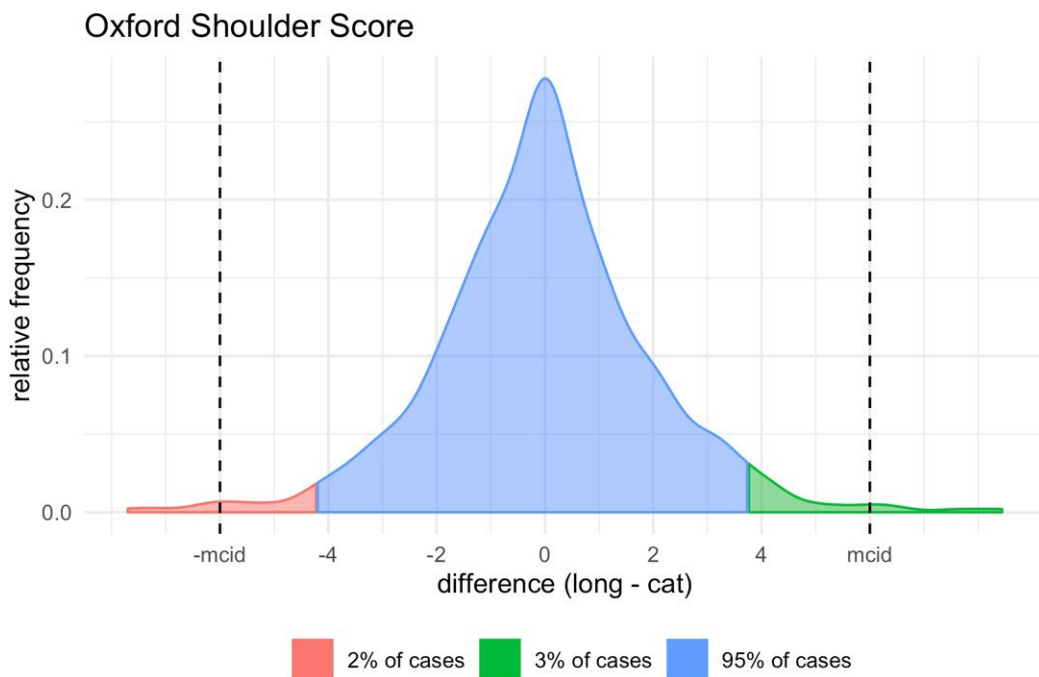
**Fig i.** Score distributions for the full-length ('Long') Oxford Shoulder Score and computerized adaptive testing (CAT) version of the Oxford Shoulder Score.



**Fig j.** Correlation of scores from the full-length ('Long') Oxford Shoulder Score and computerized adaptive testing (CAT) version of the Oxford Shoulder Score. Point colour indicates the number of questions used by the CAT algorithm in that instance.

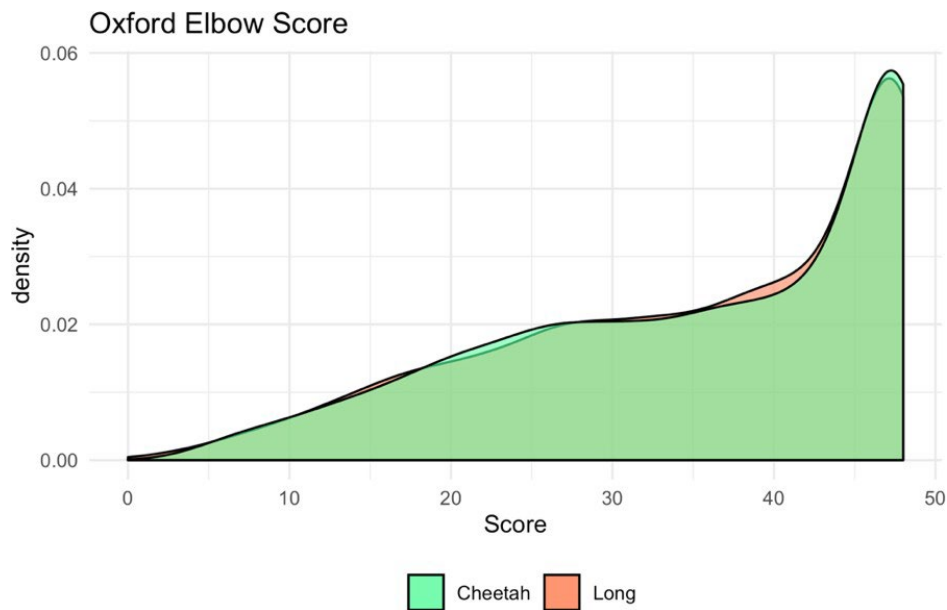


**Fig k.** Bland Altman plot demonstrating the concordance of scores from the full-length ('long') and computerized adaptive testing (CAT) versions of the Oxford Shoulder Score. For each pair of scores, the x-axis represents the mean of both scores and the y-axis represents the difference between them. The outermost, horizontal solid lines represent the 95% limits of agreement and the surrounding dotted lines represent their 95% confidence intervals.

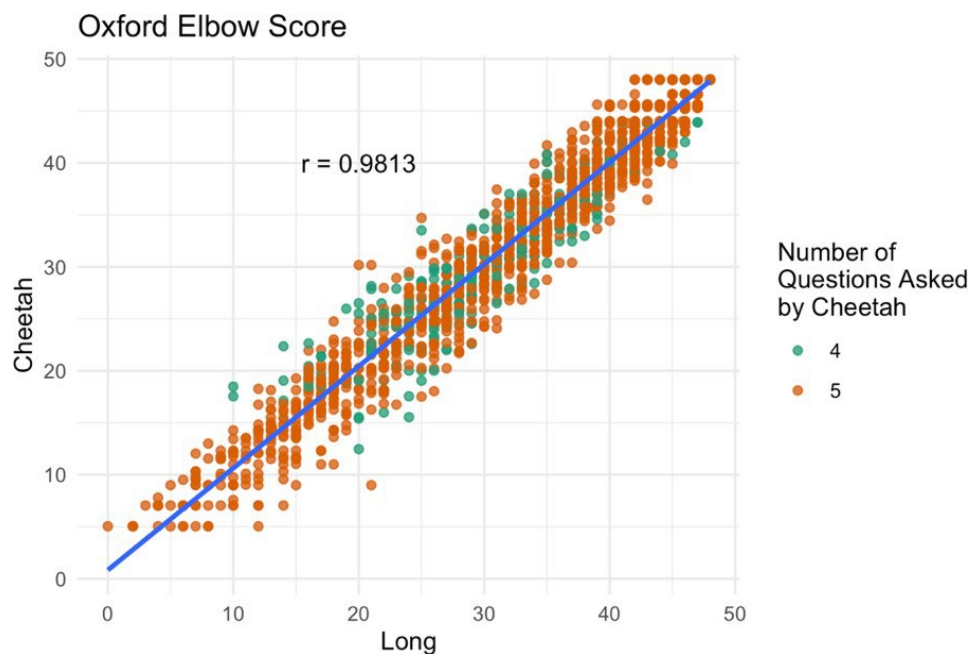


**Fig I.** Difference in score between the full-length ('long') Oxford Shoulder Score and its computerized adaptive testing (CAT) counterpart. Dashed horizontal lines represent the instrument's minimal clinically important difference (MCID).

## Oxford Elbow Score (total combined score)

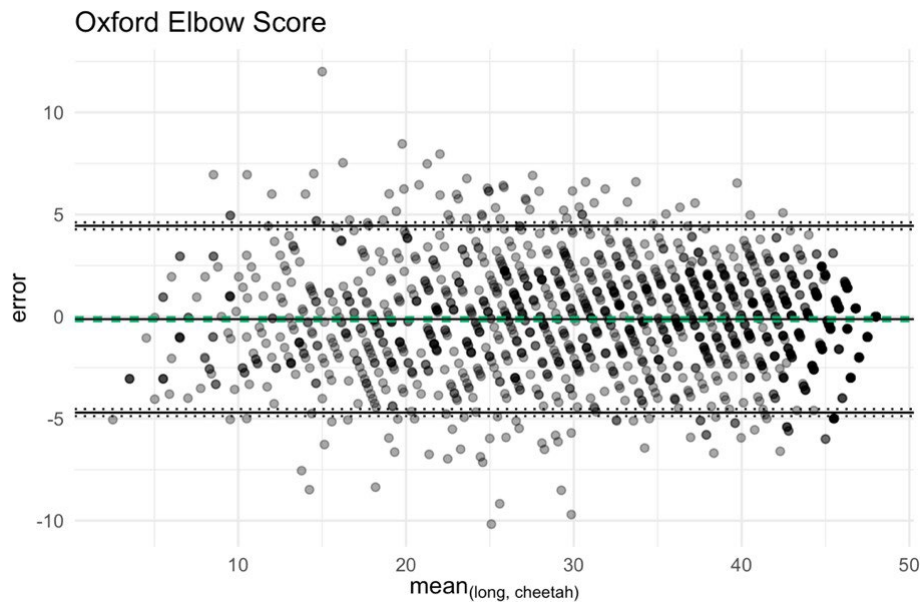


**Fig m.** Score distributions for the full-length ('Long') Oxford Elbow Score (total combined score of all three subscales) and computerized adaptive testing (CAT) version of the Oxford Elbow Score. Cheetah, CAT version (all three subscales combined).

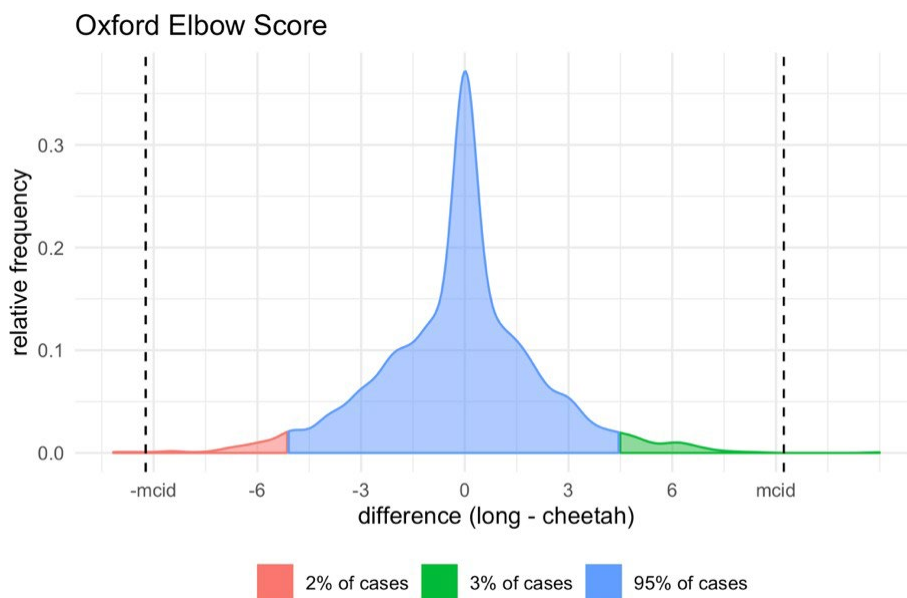


**Fig n.** Correlation of scores from the full-length ('Long') Oxford Elbow Score (total combined score of all three subscales) and computerized adaptive testing (CAT) version of the Oxford Elbow Score. Point colour indicates the number of questions used by the CAT algorithm in that instance. Cheetah, CAT version (all three subscales combined).



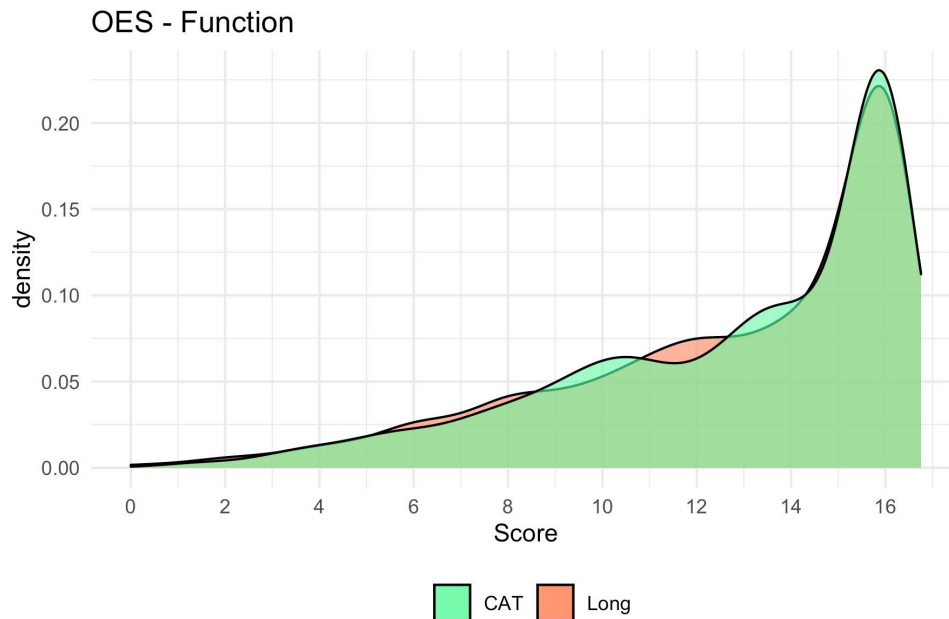


**Fig o.** Bland Altman plot demonstrating the concordance of scores from the full-length ('long') and computerized adaptive testing ('cheetah') versions of the Oxford Elbow Score (total combined score of all three subscales). For each pair of scores, the x-axis represents the mean of both scores and the y-axis represents the difference between them. The outermost, horizontal solid lines represent the 95% limits of agreement and the surrounding dotted lines represent their 95% confidence intervals.

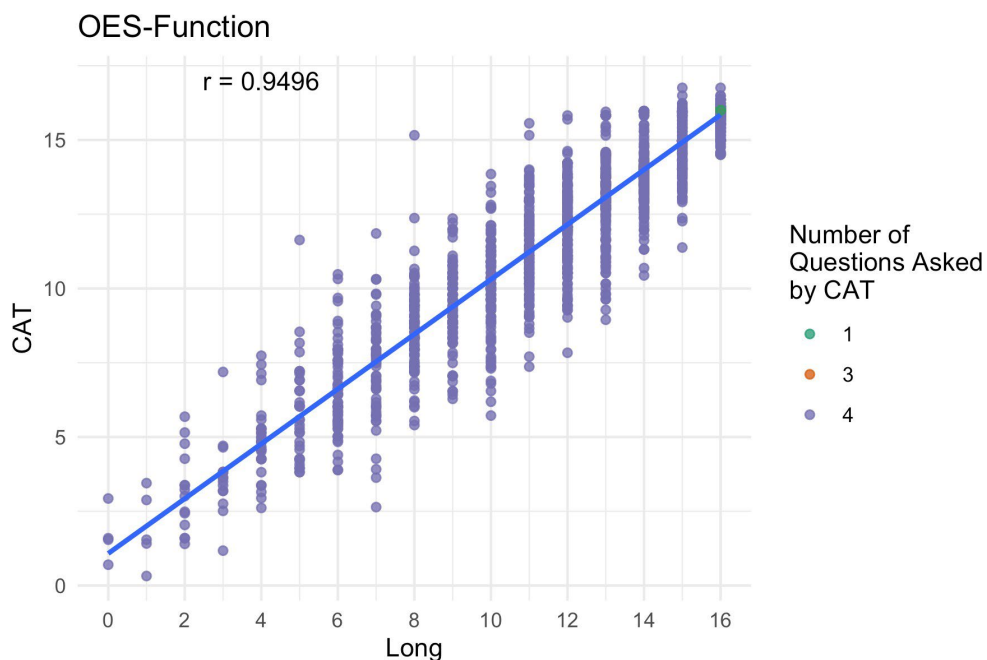


**Fig p.** Difference in score between the full-length ('long') Oxford Elbow Score (total combined score of all three subscales) and its computerized adaptive testing counterpart (cheetah). Dashed horizontal lines represent the instrument's minimal clinically important difference (MCID).

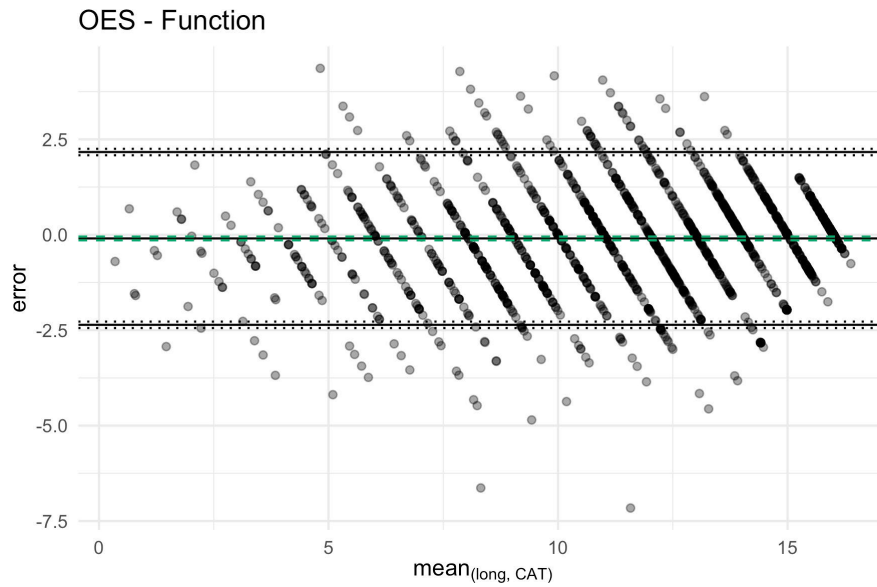
## Oxford Elbow Score (Function subscale)



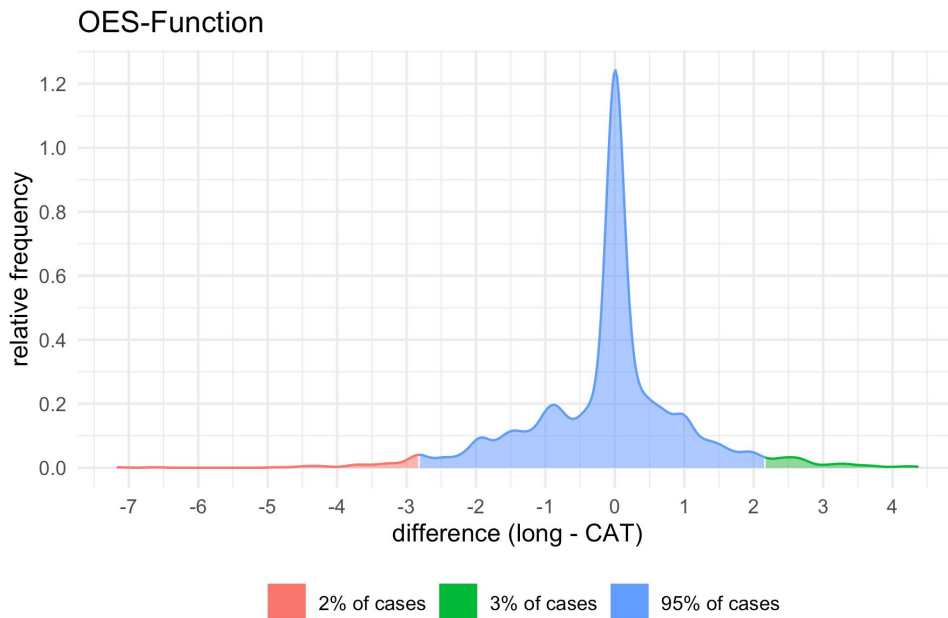
**Fig q.** Score distributions for the full-length ('Long') Oxford Elbow Score (OES) (Function subscale) and computerized adaptive testing (CAT) version of the OES (Function subscale).



**Fig r.** Correlation of scores from the full-length ('Long') Oxford Elbow Score (OES) (Function subscale) and computerized adaptive testing (CAT) version of the OES (Function subscale). Point colour indicates the number of questions used by the CAT algorithm in that instance.

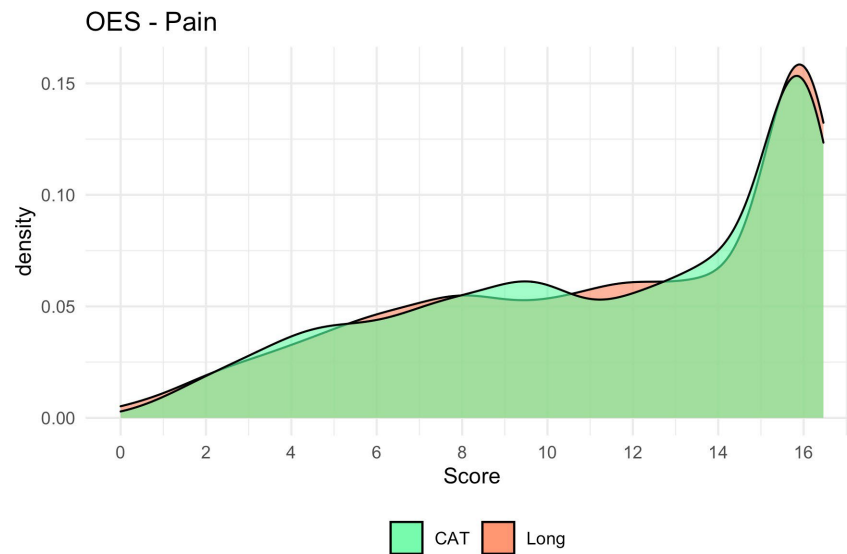


**Fig s.** Bland Altman plot demonstrating the concordance of scores from the full-length ('long') and computerized adaptive testing (CAT) versions of the Oxford Elbow Score (OES) (Function subscale). For each pair of scores, the x-axis represents the mean of both scores and the y-axis represents the difference between them. The outermost, horizontal solid lines represent the 95% limits of agreement and the surrounding dotted lines represent their 95% confidence intervals.

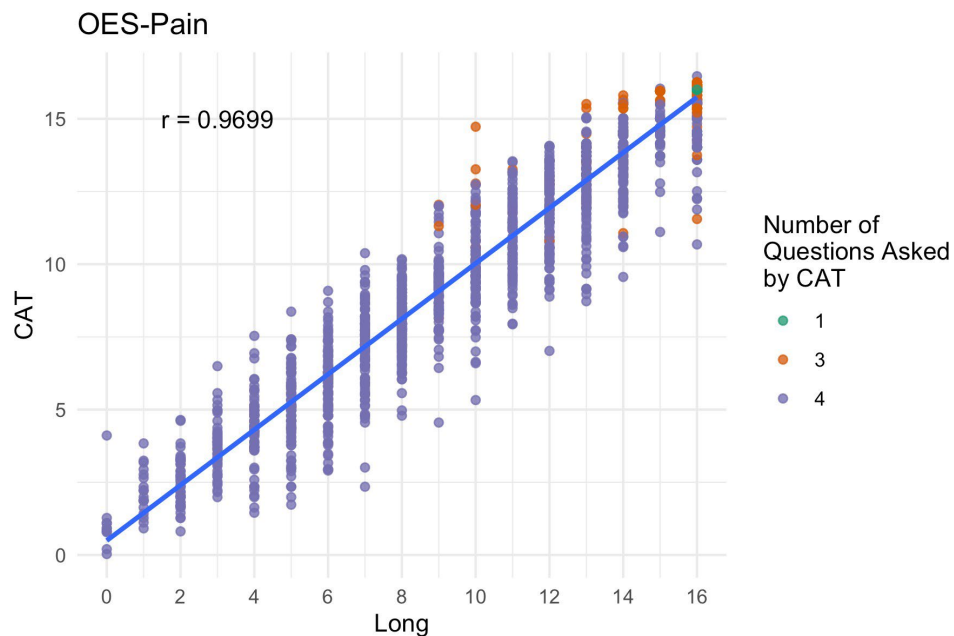


**Fig t.** Difference in score between the full-length ('long') Oxford Elbow Score (OES) (Function subscale) and its computerized adaptive testing (CAT) counterpart.

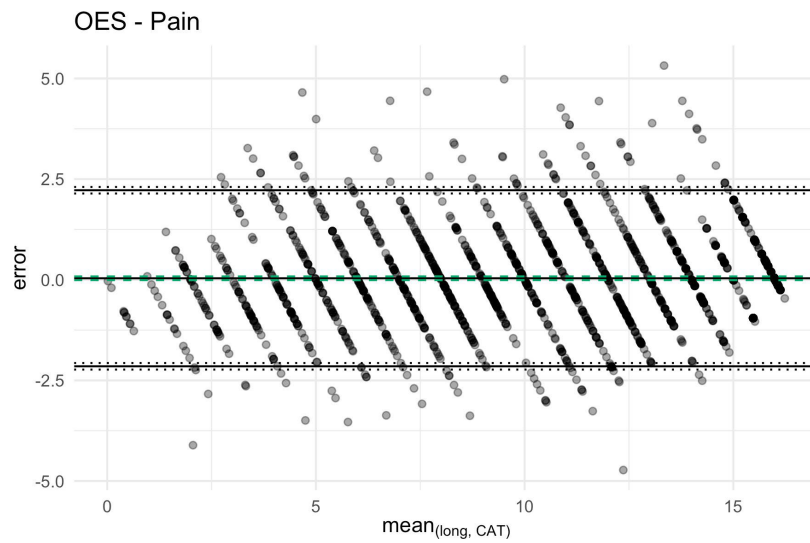
## Oxford Elbow Score (Pain subscale)



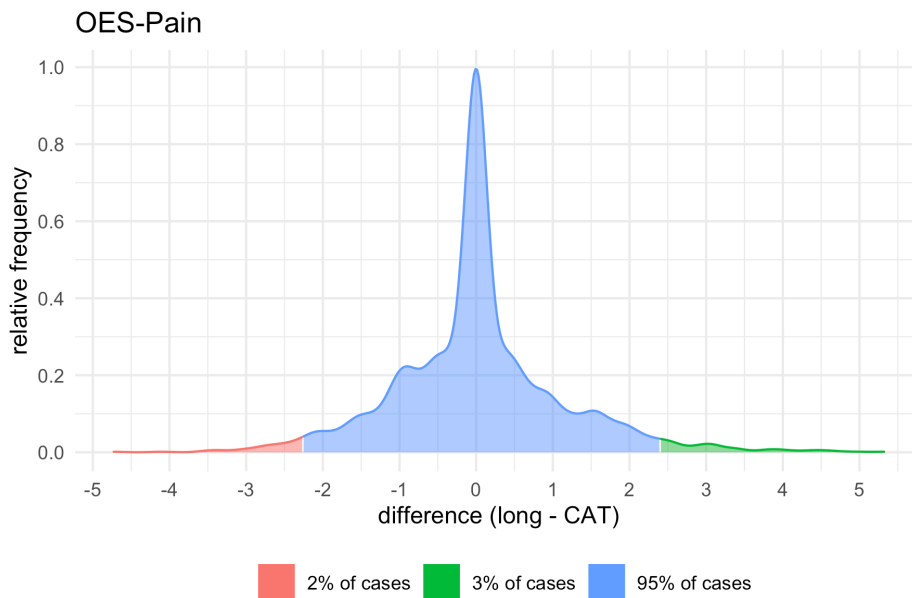
**Fig u.** Score distributions for the full-length ('Long') Oxford Elbow Score (OES) (Pain subscale) and computerized adaptive testing (CAT) version of the OES (Pain subscale).



**Fig v.** Correlation of scores from the full-length ('Long') Oxford Elbow Score (OES) (Pain subscale) and computerized adaptive testing (CAT) version of the OES (Pain subscale). Point colour indicates the number of questions used by the CAT algorithm in that instance.

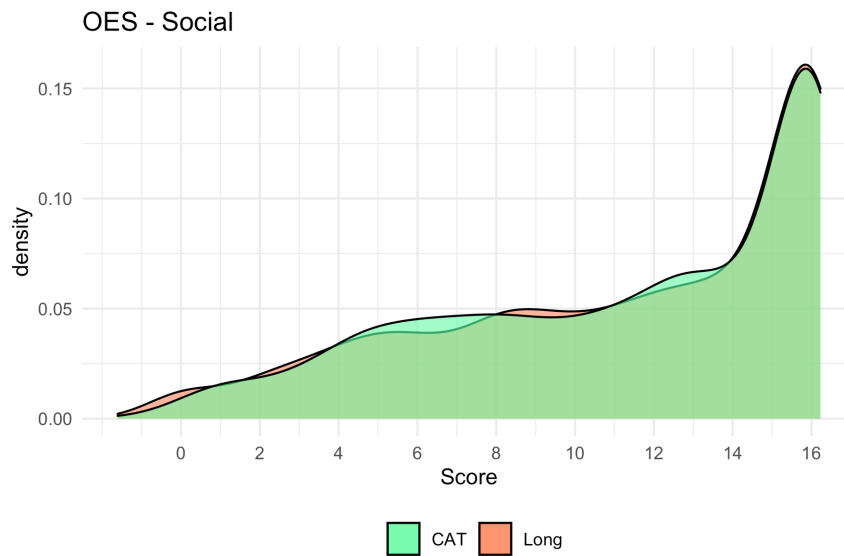


**Fig w.** Bland Altman plot demonstrating the concordance of scores from the full-length ('long') and computerized adaptive testing (CAT) versions of the Oxford Elbow Score (OES) (Pain subscale). For each pair of scores, the x-axis represents the mean of both scores and the y-axis represents the difference between them. The outermost, horizontal solid lines represent the 95% limits of agreement, and the surrounding dotted lines represent their 95% confidence intervals.

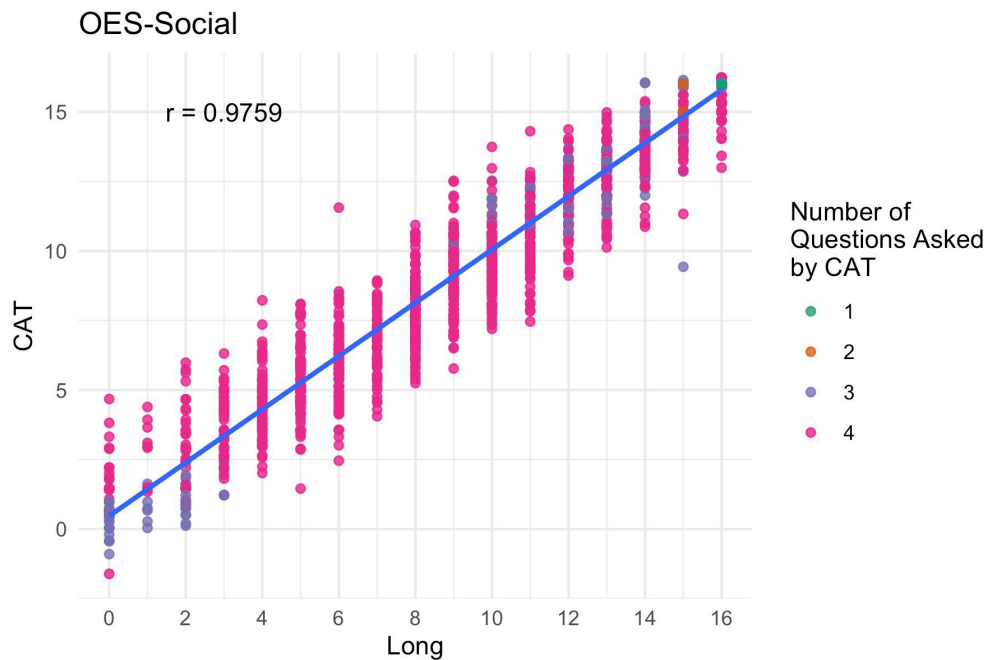


**Fig x.** Difference in score between the full-length ('long') Oxford Elbow Score (OES) (Pain subscale) and its computerized adaptive testing (CAT) counterpart.

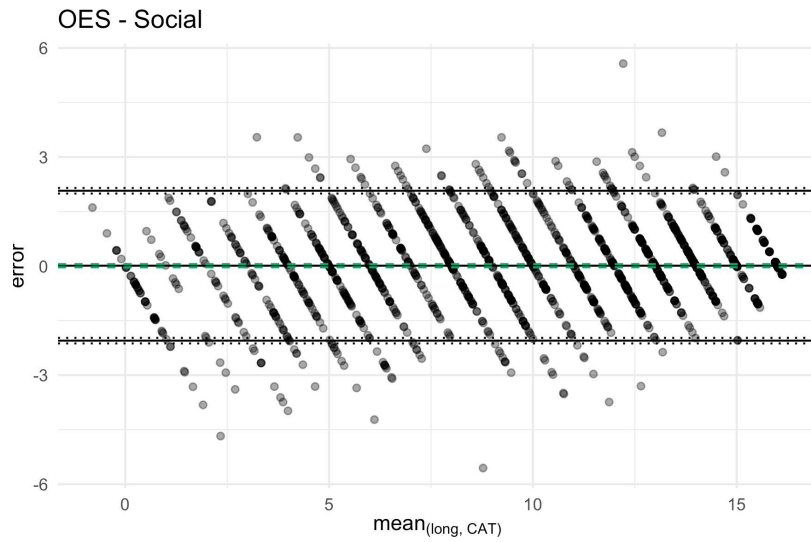
## Oxford Elbow Score (Social subscale)



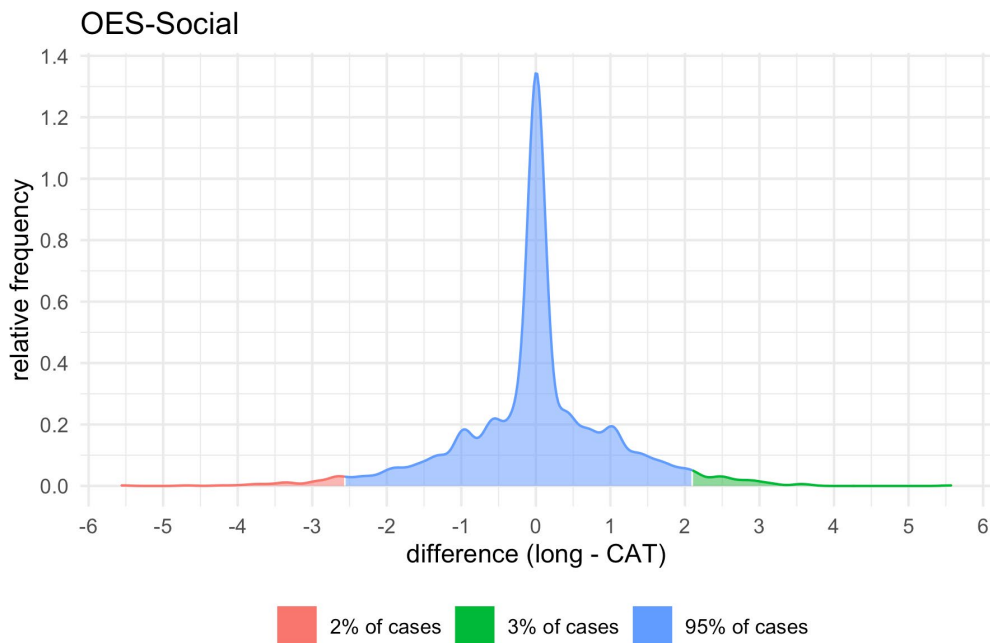
**Fig y.** Score distributions for the full-length ('Long') Oxford Elbow Score (OES) (Social subscale) and computerized adaptive testing version of the OES (Social subscale).



**Fig z.** Correlation of scores from the full-length ('Long') Oxford Elbow Score (OES) (Social subscale) and computerized adaptive testing (CAT) version of the OES (Social subscale). Point colour indicates the number of questions used by the CAT algorithm in that instance.



**Fig aa.** Bland Altman plot demonstrating the concordance of scores from the full-length ('long') and computerized adaptive testing (CAT) versions of the Oxford Elbow Score (OES) (Social subscale). For each pair of scores, the x-axis represents the mean of both scores and the y-axis represents the difference between them. The outermost, horizontal solid lines represent the 95% limits of agreement, and the surrounding dotted lines represent their 95% confidence intervals.



**Fig ab.** Difference in score between the full-length ('long') Oxford Elbow Score (OES) (Social subscale) and its computerized adaptive testing (CAT) counterpart.