

Supplementary Fig. S2.

Examples of two PDQ-39 items (A, item 1/Mobility scale; B, item 19/Emotional well-being scale) displaying statistically significant (following Bonferroni correction) signs of differential item functioning (DIF) between men and women. The item characteristic curves (ICCs; grey curves) represent the expected scores (y-axis) at various levels of the measured construct (x-axis). Superimposed plots represent the observed average item scores of men and women as divided into three class intervals according to their scores on the respective scales. Observed differences indicate that items do not work in the same way and have different meanings across subgroups (e.g., at any given level of mobility impairment, men score higher than women on item 1 of the Mobility scale [A]). For comparison, panel C illustrates an item (item 13/ADL scale) without DIF.

