O19 Pilot study using artificial intelligence in the translation of PROMIS sexual function and satisfaction measures

Lorynn Teela, Tessa van Gastel, Athanasios Angelakis, Lotte Haverman, Caroline Terwee

Amsterdam UMC, Amsterdam, Netherlands

Objective: Translation and linguistic validation of patient-reported outcome measures (PROMs) is a time-consuming and costly process. There is increasing interest in the use of artificial intelligence (AI) for translation, but studies comparing the translation quality of AI supported translations to human translations are scarce. We aimed to explore the potential of AI in the translation of PROMs.

Methods: A total of 58 items from the PROMIS v2.0 SexFS item banks were translated into Dutch through standard PROMIS methodology. In addition, ChatGPT3.5 and DeepL were used to translate the PROMIS SexFS items. Five PROM experts, four sexuality experts, and five language experts rated the quality of the three translations (human, ChatGPT3.5, and DeepL), presented in a blinded and random order. Experts were given three assignments: (1) Rate the quality of the translation. Each expert received the 58 original English items as well as 1 translation per item (randomly selected from the 3 translations). (2) Which item translation is the best? Each expert received the 58 original English items as well as the 3 translations per item (in random order per item and per expert). (3) Which overall translation is the best? Each expert received the 58 original English items as well as the 3 translations per item (same order for each item, but different among experts).

Results: Multiple native Dutch and English speakers were involved in the translation process using standard PROMIS methodology. This translation process also included five independent reviews and pretesting with 20 Dutch adults. The results of the comparison of translations using different methodologies, varying in number of people, time, and costs involved, will be presented at the conference.

Conclusions: Although this study has some limitations (only one language, highly educated raters), the results will contribute to the currently limited evidence on the appropriateness of AI use in PROM translation. More experimental studies are needed, as an improved translation process could significantly enhance the global adoption and use of PROMs.