This document helps explain in more detail the mathematical and statistical background of the methods used in the analysis of the data.

Below, we have created several chapters which are linked to the chapters in the main manuscript.

Procedures and statistics:
Rasch analysis 1: pre-IN-QoL
After the data was subjected to a quality control the data was split up into 3 random groups. This was done to improve the item-to-patient ratio, as most measurement models require a higher total number of patients than total number of items, as is the case for the Rasch unidimensional measurement methods (RUMM2030) used. [1, 2]
The 3 subgroups were randomly created by using the commando “random sample” in Stata. Next, the data in these 3 groups were checked for misfitting items, which was defined as either having a fit residuals exceeding ±2.5 or a significant χ²-probability (after Bonferroni adjustment)).

Procedures and statistics
Rasch analysis 2: creating the final IN-QoL
The analyses through Rasch modelling will be explained in more detail here. All item sets with correlations above 0.3 were evaluated starting with the highest correlations (>0.7, >0.6, … up to >0.3). Of each item set, the item with the most over- or under-
discrimination on its category probability curve, eventually combined with less face validity (as judged by ISJM as a clinical expert), was removed. For item bias, we used the following person factors: age categories (18-50 years vs 51-63 years vs 64-90 years), gender (men vs women) and diagnosis type (GBS, CIDP, MGUSP and MMN). The age groups were created to obtain an optimal equal distribution in size between the groups. Misfitting persons were identified as showing fit residuals exceeding ±2.5 or a significant $\chi^2$-probability (after Bonferroni adjustment). We performed a binominal $T$-test to test for multidimensionality. QoL-questionnaires are generally known for either more physically or mentally oriented items. Therefore, if multidimensionality was found, a factor analysis was also performed using Stata 13.0. [3] For additional information and an example on uniform Differential Item Functioning (DIF), see Supplemental Figure 1.

**Procedures and statistics**

Validity, Reliability, and Responsiveness

To determine responsiveness of the IN-QoL, the individual standard errors (SE) were determined for each patient, alongside their locations on the logits-ruler. SEs may vary across the theoretical range of an outcome measure, and therefore the significance of a clinically meaningful change in any particular patient may also vary.[4] The minimal clinically important difference (MCID)-related SE (MCID-SE) score is based on the previously described significant changes.[4] The MCID-SE was calculated for each patient separately. See Supplemental Figure 3 for the formula and classification of the MCID-SE values.

The responsiveness of the EQ-VAS was determined in a similar fashion. As there is no personal SE attainable for a single value, the group SE for each visit was
calculated by Stata 13.0 and was used to determine the MCID-SE for the EQ-VAS.

**REFERENCE LIST**