What will questionnaire measure?

What should questionnaire include?

Suitability for potential respondents?

Does questionnaire measure desired construct?

Will questionnaire be required for use in new language or culture?

- What construct should be measured?
- How is construct defined?

- How should content be produced (item generation)?
- How is best item set selected (item reduction)?

- Does it appear to measure construct appropriately (face validity)?
- Is content a reasonable representation of patient experience (content validity)?

- Do items measure intended construct (construct validity)?
- Can item scores be added together validly (scalability)?
- Is level of measurement error acceptable (reliability)?
- Is scale able to measure real change in construct (responsiveness)?
- Are data collected free from biases unrelated to construct (differential item functioning, DIF)?

- Have appropriate translation methods been employed?
- Have new language versions been tested to ensure suitability for local patients?
- Do adaptations have adequate psychometric and scaling properties?