**Stage 1 - Searching & screening**
- Abstract and full text screening for initial inclusion in Stage 1.

**Stage 2 - Taxonomy development**
- Summary of content of individual instruments:
  - individual items
  - construct definitions
  - additional linked sources
- Development of taxonomy for categorising instruments by content.
  Iterative process involving:
  - comparison of content between instruments to identify a comprehensive set of categories / subcategories
  - collapsing similar categories
  - labelling of categories, where possible, matching to well defined constructs**

**Stage 3 - Categorization of instrument content**
- Categorization of content of instruments according to the taxonomy; tabulation in content matrices

**Stage 4 - Assessment of measurement properties**
- Summary and tabulation of:
  - description of the instrument’s purpose and development (including theoretical basis)
  - methods and main findings of assessment of measurement properties (reliability, validity, responsiveness)
  - assessment of the adequacy of evidence supporting measurement properties, using pre-defined appraisal criteria

**Sources**
- MEDLINE, PsycINFO, HAPI, reference lists of systematic reviews of measurement instruments
- Snowballing: reference lists, citation searching (main papers)

**Papers meeting inclusion criteria for stage 2:**
1. reported development or use of one or more instruments intended to measure a construct a) within the scope of the conceptual framework OR b) identified by the study authors as a potential determinant of successful QI and judged by the review authors as relevant to primary care*;
2. instrument was quantitative, allowing for statistical analysis of its measurement properties;
3. instrument was intended to be self-report;
4. information about the instrument was published in English, in the peer-reviewed literature and sufficient detail was reported to enable assessment of its content

**Stage 4 - Assessment of measurement properties**
- Conceptual framework as initial basis for taxonomy
- Data from content analysis of individual instruments
- Review and conceptual articles identified in search or subsequent purposive search
- Other relevant sources

**Subset of instruments from stage 2 that:**
1. based on content analysis, were confirmed as measuring a construct relevant to the evaluation of CQI in primary care (i.e., items/scales were relevant), OR
2. measure a relevant construct not adequately covered by instruments suitable for a primary care setting.

**Subset of instruments from stage 3 that:**
1. based on content analysis, the instrument appeared to be an adequate measure of a relevant construct;
2. the instrument a) was used in, or developed for, primary care OR b) has potential for use in a primary care setting, no primary care equivalent was identified and the instrument appeared to be one of the best available measures of its type;
3. sufficient detail was reported in the paper, or linked additional files, to enable assessment of the development of the instrument and evidence of its measurement properties;
4. an English-language version was available; and
5. the instrument was not proprietary.

* External factors (e.g., financing, accreditation) were excluded as these are likely to be specific to the local health system

** Extent to which this was possible depended on the existence of consistent construct definitions in multiple included studies or, alternatively, in synthesised sources from the extant literature (i.e., recent or seminal review article or meta-analysis)