Challenges in validating cost (RU) questionnaires

Joanna Coast

DIRUM, University of Birmingham, October 2011
“Our review identified 95 studies which included an economic analysis; of which 85 recorded patient-level resource use. Alarmingly, only 21 of these reported piloting their patient health care resource use questionnaires and only 28 reported validating their health care resource use data collection methods.” (DIRUM website, re. HTA systematic review of resource use data collection in trial settings)
Is this ‘alarming’?
Validity:

- the extent to which the instrument measures what it purports to represent
- and the degree to which it might be helpful in answering a particular question
Why do few studies consider validity?

- Are there some intrinsic conceptual challenges about validating resource use measures?
  - Nature of information being collected
  - Cognitive processes involved in responding
  - Important types of validity for RU questionnaires
# Nature of information

## Outcome for EE
- Perceptive
  - No gold standard
- Generic

## RU for EE
- Factual
  - Gold standard
- Specific
  - To condition
  - To intervention
  - To study period
  - To health system
- Complete
  - Across whole trial period

---

**Joanna Coast: Challenges in validating cost questionnaires**

[Logo: MRC Hubs for Trials Methodology Research] [Logo: University of Birmingham]
Examples of EQ-5D questions

- **Mobility**
  - I have no problems in walking about
  - I have some problems in walking about
  - I am confined to bed

- **Pain/Discomfort**
  - I have no pain or discomfort
  - I have moderate pain or discomfort
  - I have extreme pain or discomfort

Joanna Coast: Challenges in validating cost questionnaires
Since you agreed to take part in the study, have you visited a GP, because of your skin condition?

No  □1 → *please go to question 9*

Yes □2 → how many visits?  ..........
Cognitive response model

Comprehension → Retrieval

Response ← Judgement
Validity

- Content validity
- Criterion validity
- Construct validity
Validity

- Content validity
- Criterion validity
- Construct validity
Content validity

- “Systematic & documented approach to ensure that the instrument assesses the desired construct” (Cook, American Journal of Medicine, 2006)
  - Where the construct is total resource use associated with the intervention for the condition over the study period
“Evidence of content validity since a number of health economists brought their own questionnaires from previous studies to provide a checklist of resource use items” (annotated cost questionnaire)

“Evidence of content validity since protocol referred to a focus group made up of healthcare professionals who developed the programme content.” (PD REHAB questionnaire)

“Evidence of content validity since questionnaire content developed from expert panel review and previous pilot.” (PD MED questionnaire)
- “Evidence of content validity since resources were defined prior to study” (RUQ-D questionnaire)
- “Evidence of content validity, resource use identified by reference to similar studies” (Stroke Receipt of Services questionnaire)
Criterion validity

- **Criterion validity** examines a measure’s performance in relation to an appropriate ‘gold-standard’ instrument.
  - Assumption that routine databases provide a gold standard
  - Area for discussion
Evidence of criterion validity as questionnaires were compared to patient databases (BRUM questionnaire)

Evidence of face, content and criterion validity, for example validation performed to external databases, resource use identified through pilot study and consultation with healthcare professionals (CESAR questionnaire)

Evidence of content and criterion validity: for example extensive expert review of previous literature and use of access database to cross-validate dates of admission for length of stays. (HOPEFUL study questionnaire)
Critierion validity & routine databases

- A gold standard?

- Examples of rules required with a database for identifying GP consultations associated with patients receiving alternative dermatological treatments
  - If a skin condition was mentioned, the consultation was recorded, even if this was not the primary reason for the consultation;
  - If a patient had a skin condition, but the consultation was about other problems and the skin condition was not mentioned, this was not recorded;
  - If a repeat prescription was issued on the same day as the consultation but the skin condition was not mentioned, the consultation was not recorded, but the prescription was.
Construct validity

- represents whether relationships between the measure and other factors are those that would be anticipated \textit{a priori} both in terms of observing relationships (in the anticipated direction) where they are expected and not observing them when they are not.

  - Is this enough for resource use data?
    - Not just the relationship that is important but also the absolute quantities…
Discussion

- Is it feasible to validate every RU questionnaire for every condition in every context
  - Unlike for outcome measures, every study will require its own validation
  - Is this a good use of resources?
Discussion

☐ Validation of methods rather than valuation of questionnaires?
  - E.g. optimal recall times for different groups?
  - E.g. optimal methods of asking about recall (length of time, salient events)

☐ Extent to which the questionnaire is a ‘whole’ and ‘the questionnaire’ can/should be validated, rather than individual questions or types of question?