

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SciVerse ScienceDirect

journal homepage: [www.elsevier.com/locate/jval](http://www.elsevier.com/locate/jval)

## Development of a Database of Instruments for Resource-Use Measurement: Purpose, Feasibility, and Design

Colin H. Ridyard, PhD, Dyfrig A. Hughes, PhD, MRPharms\*, on Behalf of the DIRUM Team

Centre for Health Economics and Medicines Evaluation, Institute of Medical and Social Care Research, Bangor University, Bangor, United Kingdom

### ABSTRACT

**Background:** Health economists frequently rely on methods based on patient recall to estimate resource utilization. Access to questionnaires and diaries, however, is often limited. This study examined the feasibility of establishing an open-access Database of Instruments for Resource-Use Measurement, identified relevant fields for data extraction, and outlined its design. **Methods:** An electronic survey was sent to authors of full UK economic evaluations listed in the National Health Service Economic Evaluation Database (2008–2010), authors of monographs of Health Technology Assessments (1998–2010), and subscribers to the JISCMail health economics e-mailing list. The survey included questions on piloting, validation, recall period, and data capture method. Responses were analyzed and data extracted to generate relevant fields for the database. **Results:** A total of 143 responses to the survey provided data on 54 resource-use instruments for inclusion in the database. All were reliant on patient or carer recall, and a majority

(47) were questionnaires. Thirty-seven were designed for self-completion by the patient, carer, or guardian, and the remainder were designed for completion by researchers or health care professionals while interviewing patients. Methods of development were diverse, particularly in areas such as the planning of resource itemization (evident in 25 instruments), piloting (25), and validation (29). **Conclusion:** On the basis of the present analysis, we developed a Web-enabled Database of Instruments for Resource-Use Measurement, accessible via [www.DIRUM.org](http://www.DIRUM.org). This database may serve as a practical resource for health economists, as well as a means to facilitate further research in the area of resource-use data collection.

**Keywords:** clinical trials, cost analysis, data collection methods, economic evaluation, health technology assessment.

Copyright © 2012, International Society for Pharmacoeconomics and Outcomes Research (ISPOR). Published by Elsevier Inc.

### Background

Health economists use a variety of methods to estimate patient-level resource use for the determination of costs and cost-effectiveness. These typically include a combination of one or more of the following: routine medical records (e.g., patient notes and databases); case report forms completed by trial researchers or health care professionals; and patient- (or carer-) completed logs, questionnaires, or diaries.

In a recent review of the methods of resource-use data collection within UK National Institute for Health Research Health Technology Assessment (HTA) program-funded trials, we identified those based on patient recall as being the most common [1]. We found important differences in whether and how analysts identified important cost drivers *a priori*, piloted their methods (including testing for reliability and validity), and selected recall periods and timing of administration. Of note was the frequency with which questionnaires were reused, recycled, and reinvented. This confirmed anecdotal evidence that health economists rely on a few often-used methods and modify their own methods for differ-

ent settings of care or clinical application, and contrasts with other patient-reported measures—such as health outcomes—whose rigorous testing for reliability, validity, responsiveness, and interpretability is considered essential.

The choice of appropriate health-related quality-of-life and patient-reported outcome instruments is aided by the existence of an Internet-based database, Patient-Reported Outcome and Quality of Life Instruments Database [2]. We hypothesized that a comparable open-access database of resource-use instruments based on patient recall would not only be a valuable resource to health economists but also encourage further research in providing an evidence base for instrument development.

The aims of the present study were to determine the feasibility of accessing resource-use instruments, to determine the appropriate fields of data for extraction, and to develop a Web-enabled database of resource-use instruments. The objectives were 1) to conduct a survey of health economists; 2) to identify a sample of recent resource-use instruments based on patient recall; 3) to extract information on a range of variables, including economists' selection of resource items, evidence on piloting, validation, cost-

DIRUM team: Professor Dyfrig Hughes and Dr. Colin Ridyard (Bangor University), Dr. William Hollingworth, Dr. Sian Noble, and Dr. Joanna Thorn (University of Bristol), Professor Joanna Coast (University of Birmingham), Dr. David Whitehurst (University of British Columbia), and Professor Martin Knapp (London School of Economics and Political Science). DIRUM is funded via the Network of Hubs for Trial Methodology Research (MRC).

\* Address correspondence to: Dyfrig A. Hughes, Centre for Health Economics and Medicines Evaluation, Institute of Medical and Social Care Research, Bangor University, Dean Street, Bangor LL57 1UT, UK.

E-mail: [d.a.hughes@bangor.ac.uk](mailto:d.a.hughes@bangor.ac.uk).

1098-3015/\$36.00 – see front matter Copyright © 2012, International Society for Pharmacoeconomics and Outcomes Research (ISPOR).

Published by Elsevier Inc.

<http://dx.doi.org/10.1016/j.jval.2012.03.004>

ing perspective, frequency of data collection, and period of recall; and 4) to gauge health economists' willingness to provide copies of their resource-use instruments with a view to include in an open-access database.

## Methods

### Sample

Subscribers to the JISCMail health economics e-mailing list [3], whose membership is open to all health economists (and all those interested in health economics) worldwide, were invited to complete an online survey. The mailing list is the recommended discussion list for members of the International Health Economics Association and the UK Health Economists' Study Group and has a membership of more than 750. The survey was also advertised widely via the HealthEconomics.com newsletter, which has an international circulation of 5500.

We also screened abstracts for potential use of resource-use instruments based on patient recall and contacted the corresponding authors (or the health economists, where identified) of 1) UK-based economic evaluations published between January 2008 and March 2010 and listed in the National Health Service (NHS) Economic Evaluation Database ( $n = 159$ ), 2) HTA reports of primary research, published between January 1998 and March 2010 ( $n = 56$ ), and 3) ongoing HTA studies (as of March 2010) ( $n = 99$ ). Screening involved examining the titles and abstracts of articles and excluding any reviews, models, or secondary data analysis. Authors were invited to complete the online survey, and nonresponders were contacted by e-mail with reminders at 3 and 6 weeks.

All survey respondents who agreed to be contacted directly were asked to share a copy of their resource-use instrument(s) and permission for its inclusion in the database. Those who replied, but who did not complete the survey, were invited to share a copy of their resource-use instrument(s) for inclusion in the database.

### Exclusion criteria

All identified resource-use instruments were included except those in which patient recall was not seen as a factor in resource identification (e.g., case report forms used to abstract data directly from hospital notes). We did not screen instruments for their quality, and so they were not excluded on this basis.

### Survey

The survey was administered through SurveyMonkey, an Internet-based system that enables users to create their own Web-based surveys [4]. The survey asked for information on the following:

1. Details on respondents' names, e-mail address, and affiliations
2. Reference for identification purposes (e.g., reference(s), title of trial, authors, and trial registration number such as ISRCTN and NCT)
3. Publications arising from the study that make reference to the resource-use data collection instrument
4. Description of the data collection instrument from a drop-down menu of patient-completed questionnaire, patient-completed Internet-based questionnaire, patient-completed diary, researcher-completed questionnaire based on face-to-face patient interview, researcher-completed questionnaire based on telephone interview, other (with option to specify)
5. Maximum patient recall period (1 day, 2 days to 7 days, >1 week to 1 month, >1 month to 3 months, >3 months to 6 months, >6 months to 12 months, >12 months)
6. Costing perspective (one or more of health service, personal social services, patient [and carer] out-of-pocket costs, educa-

tion sector, employers, societal, voluntary, social security benefits/contributions, other with option to specify)

7. Piloting (not piloted, piloted by patient group within trial, piloted by separate patient group, piloted by healthy volunteers, piloted by health care professionals, used previously, or other, with an option to specify)
8. Validation (option to specify)

Prior to distribution, five experienced health economists were invited to complete the survey and provide their comments. Responses were received from three, who gave favorable comments, and the questions remained unchanged. No follow-up was conducted to ascertain why the other two health economists did not pretest the instrument.

### Assessment

Based on the instruments provided, associated publications, and responses to the online survey, we extracted information on potentially important elements by using a good practice checklist for resource-use data capture alongside clinical trials [1]. These were divided into *characteristics* such as content, completion method, population, setting of care, disease area, resource-use area, recall period, and periods of administration; *qualities* such as reliability, validity, responsiveness, development, piloting, and cognitive preparation; and *supporting material* such as instructions. Following data extraction and review, we determined a list of fields for the database.

## Results

One hundred and forty-three electronic SurveyMonkey questionnaires were returned, of which 79 (55%) carried useful information such as publication references, instrument details, and contact e-mail address. Of these, 26 respondents came from the NHS Economic Evaluation Database category, 12 were from the category of published HTA reports, and 23 from the category of ongoing HTA studies. The remaining 18 respondents may have come from the JISCMail group but could equally have been prompted to respond by members of the other three groups. Questionnaire responses yielded data on 80 instruments. Of these, 16 were unavailable because the respondents specifically indicated an unwillingness to share. Follow-up of willing-to-share respondents and other subsequent contacts yielded a total of 54 resource-use instruments considered suitable for the database (once exclusion criteria and non-responses were taken into account). The majority (36 of 54) were from National Institute for Health Research HTA-funded trials. Research charities funded 6; the Department of Health and NHS trusts each funded 4 studies, 2 were from pharmaceutical company-funded research, and 2 were funded by the European Union. Forty-seven (87%) instruments took the form of questionnaires, with the remainder being diaries.

As well as the key identification fields such as instrument name and instrument developer, the following database fields were defined on the basis of a structured review of the 54 instruments.

### Resource-use instrument characteristics

#### Administration

Administration describes who completes the resource-use instrument. Thirty-seven instruments were designed for completion by patients, their carers, or guardians. The remainder were designed for completion by researchers or health care professionals while conducting telephone or face-to-face interviews. Some instruments had been adapted for administration by patient/carer/guardian and by researcher/health care professional during pa-

tient interviews. For example, the designers of the Annotated Cost Questionnaire [5] recognized that patient completion alone was not always appropriate, whereas the Client Service Receipt Inventory [6], while originally designed for trained interviewers and directed at client carers, has been adapted for administration in multiple contexts (e.g., for completion by patients [7], parents [8], and health care professionals/carers [9]).

We identified the most appropriate administrative descriptors for the resource-use instruments as patient-completed form (reported in 26 of 54), parent-/carer-completed form (7 of 54), patient/carers diary (7 of 54), researcher-/health care-professional-completed based on patient recall (23 of 54), face-to-face interview (22 of 54), telephone interview (9 of 54), and computer administration (1 of 54). It was clear from the sample of instruments that a combination of these categories often applied to a single questionnaire or diary (median 1; range 1–6; interquartile range [IQR] 1–2). Each resource-use instrument could therefore have a number of administrative descriptors, each defined by a binary yes/no category.

### Resource categories

For the purpose of this article, we define “resource categories” as individual or grouping of questions on discrete elements of health care, such as general practitioner (GP) visits, hospital outpatient clinic attendance, accident and emergency visits, medicines, and the like. The authors of 25 of the 54 resource-use instruments reported to have identified resource categories for costing at an early stage in their development. This was usually done by consultation with health care professionals, pilot studies, or literature searches, but details on the extent to which this was done were not always clear.

The principal resource categories identified in the instruments included GP (47 of 54), inpatient (46), outpatient (38), and community/practice nurse services (45). Use of dental, optician, dietician, and the services of other professions allied to medicine was measured in 36 instruments, while that of medications was captured in 32 instruments, physiotherapy in 30, employer costs in 29, patient-incurred costs in 27, informal carer costs in 28, social services in 27, day hospital in 27, and accident and emergency in 29. Non-NHS services included state benefits (11), criminal justice system (7), and education (6). In general, the layout of all instruments made them particularly amenable to costing with standard sources of UK unit cost [10–13]. Most instruments (50) covered four or more categories (median 8; range 1–16; IQR 6–11), which needed to be reflected in the coding for the database.

### Disease

Disease classification was modeled on the Cochrane Library listing [14]. All disease areas of the 54 instruments could be mapped to this listing, although five (under the category “other”) needed a further free-text disease category field to explain their application, typically because these instruments were designed generically for measuring resource use in a wide variety of medical conditions, an example being the Expert Patients Programme questionnaire [15].

### Setting of care

We define setting of care as the setting in which the clinical trial is principally being conducted. This was not always defined explicitly, and in such cases, it was difficult to distinguish between (for example) primary care and community care and to specify a single setting for health care interventions that spanned many. In some instances, difficulty was also experienced in ascertaining where it ended (e.g., community) and began (e.g., hospital inpatient). We identified the most common “setting of care” descriptors for the resource-use instruments as community (30 of 54), GP practice (19 of 54), inpatient (16 of 54), and outpatient (16 of 54). Other lesser-

used categories included accident and emergency, day hospital, residential care, and specialist tertiary centers. A combination of these settings of care was often captured within questionnaires or diaries (median 2; range 1–9; IQR 1–2), and therefore each care settings is defined within the database by the binary yes/no category.

### Population

Categories of population ages were initially considered as mutually exclusive groups: pediatric, adult, and elderly; however, it became clear that these were not practicable because mutual exclusion could not always be inferred or assumed. We therefore recorded the upper and lower age limits of the populations in which the instruments were administered.

### Specified recall period

Within the study sample, 53 (of 54) instruments clearly stated recall periods and of these, 46 were less than 6 months. On a continuous scale, median recall period was 3 months (IQR 0.5–6 months). For the purposes of the database, a specified recall period would constitute a categorical ordinal variable and utilize descriptors of 0 to 1 week, 1 to 4 weeks, 1 to 3 months, 4 to 6 months, and so on. The likelihood is that many were determined by study protocols and follow-up for clinical assessments rather than based on empirical evidence on optimal (or maximum) recall periods.

### Frequency of data collection

Data collection periods were specified in 42 of the 47 instruments (excluding diaries) and ranged from a few weeks to 10 years, with up to 12 resource data collecting episodes per instrument. The median number of data collection episodes was 3 (IQR 1–4), with a median data collection interval of 4 months (IQR 3–6 months). Data collection intervals did not necessarily equate to recall periods but, as with recall periods, may be determined by the clinical protocol. The frequency of data collection would constitute another categorical ordinal variable within the database and utilize a similar range of ordered descriptors (0–3 months, 4–6 months, etc.).

### Resource-use instrument qualities

#### Reliability

Reliability refers to the consistency with which a resource-use instrument measures what it is supposed to be measuring [16]. On review of the instruments, however, there was very little evidence of any reliability testing (4 of 54 instruments) and, therefore, very little opportunity to assign descriptors for a reliability field within the database.

#### Validity

Validity refers to how well a resource-use instrument measures what it is intended to measure. While not necessarily reported as such, varying degrees of three specific validative descriptors were observed in a number of resource-use instruments (29 of 54). For example, expert panel review and literature comparisons were evident in 12 cases, representing a degree of content validation; acceptance of scale by patients or health care professionals was evident in 18 instances, showing a degree of face validation; and comparisons with GP or hospital notes (16 of 54) demonstrated an acknowledgment of criterion validation (if based on an assumption that health care professionals' notes provide a gold standard) or convergent validation (if exploring different methods to gain the same information).



### Cognitive interview development

Evidence of cognitive questionnaire development, present in 5 (of 54) instruments within the sample, demonstrated that steps had been taken in a few studies to ensure that patients understood the questionnaires. No evidence existed, however, for the presence of readability measures, such as the Gunning Fog Index, in either instructions or instruments. Database fields to reflect this information would be appropriate because patient comprehension is a prior requisite to recall.

### Pilot testing

Within the sample, a number of instruments (25 of 54) were considered piloted, although as observed in our previous review [1], authors used a variety of methods ranging from direct application of an annotated system [5,17], inclusion in specifically designed pilot programs [18], and reliance on previously used questionnaires from similar [19] and dissimilar [15] studies. Because of the importance of pilot testing, it would be necessary to include details of the methods employed within the database.

### Questionnaire completion

Attempts were made to assess patients' acceptability of instruments including the number of returned resource-use questionnaires, the number of unanswered questions, and completion times. There was, however, little evidence of any widespread analysis (4 of 54 instruments) in these areas—especially because some studies were still ongoing—and therefore no opportunity to assign descriptors within the database. In addition, a resource-use questionnaire may form only a part of the overall patient-reporting package and may typically appear alongside health-related quality-of-life measures. This makes it difficult to assess whether the completion rates are due to the resource-use questionnaire *per se* or due to other questionnaires/tasks within the package, or indeed, the overall length of the package itself. Nevertheless, because patient acceptability is an important criterion for judging the effectiveness of a resource-use questionnaire, it would be useful for the database to include relevant details where these are available.

### Resource-use instrument supporting materials

Instructions for completion were available for 24 of the 54 instruments. These ranged from a short paragraph at the top of a postal questionnaire to multipage manuals of instructions for interviewers.

## Discussion

There is general agreement that in the absence of cognitive and/or physical impairments, patients are an important source of information on the use of health care and other resources [20,21]. Diaries and questionnaires offer a structured means of gathering information on the use of a wide range of health, out-of-pocket, informal, and social care resources that may not be possible by other methods. When used alone, or to complement other methods—such as routine medical records and case report form entries—they provide valuable data to inform economic evaluations of health care interventions.

Comparing data based on patient recall with data from hospital notes or other secondary source abstraction allows for correlative corroboration to improve validity [21–23]. There are, however, clear limitations to estimating resource use based on patient recall. These are related to the length of the recall period, questionnaire design, mode of data collection, impact on patient, frequency of administration, level of social desirability, and patient cognition and demographics, which affect the va-

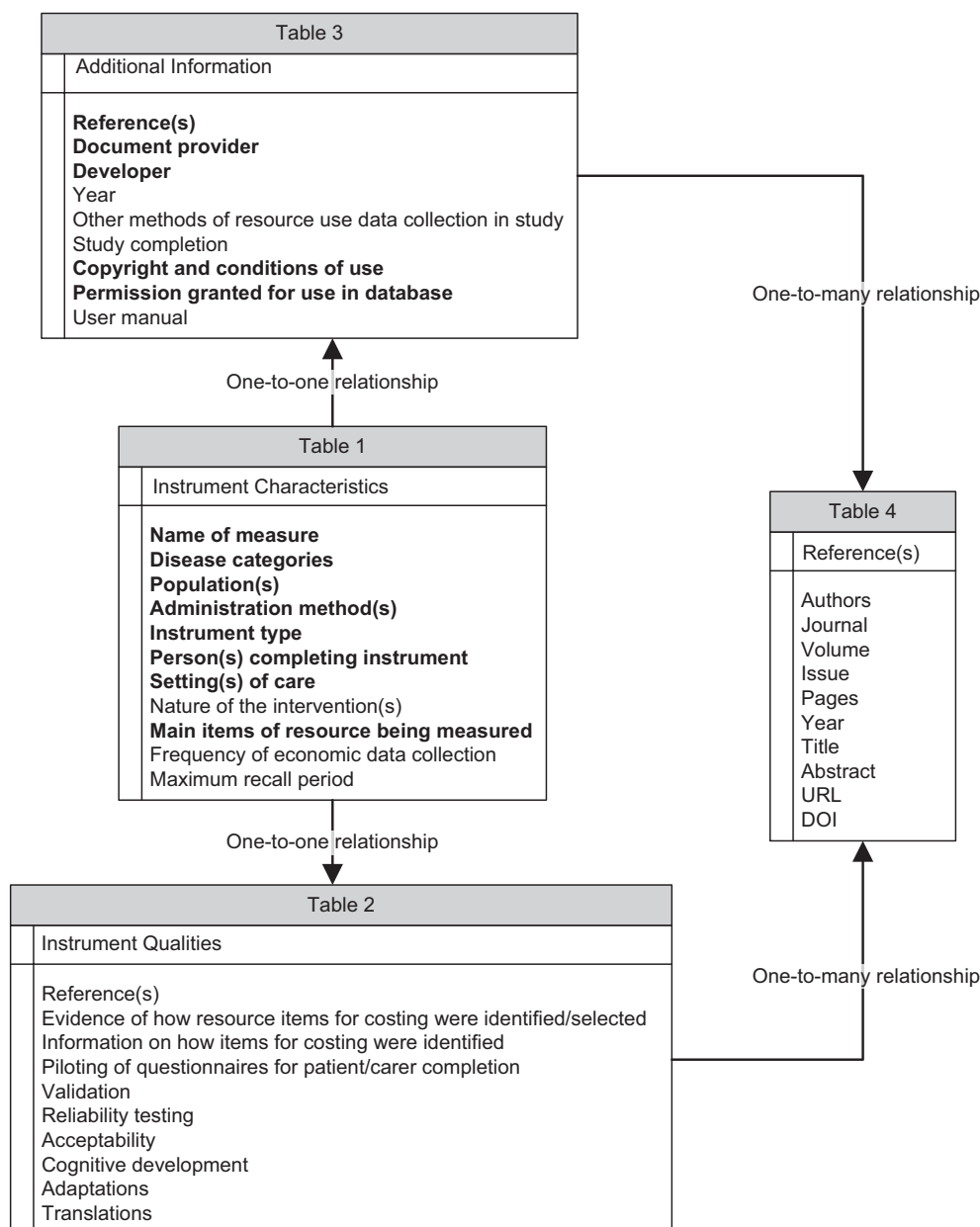
lidity of reported results [20,24]. Nevertheless, because clinical records are often fragmented across different health system sectors and established for clinical and/or payment purposes rather than research, and because sufficient information for accurate costing may not always be readily available from these sources, patient-reported resource-use instruments are unlikely to be superseded in the short term. The balance between data accuracy and the costs of alternative methods should always be considered during the design of a trial-based economic evaluation. Key steps in the process of questionnaire development involve the identification of key cost-driving elements (e.g., based on expert opinion, use of focus groups, and literature reviews), piloting, validation, and reliability testing. However, methods vary from study to study, and there is no acceptable gold standard for the development of resource-use instruments. Researchers relying on patient recall are normally left with a choice of either designing their own instruments or adapting instruments from previous studies. Reliance on instruments developed by others may facilitate this process, but access to such instruments is often limited, and reliability and validity may not be generalizable to different populations or contexts. One example of this is the way instruments are administered. Some instruments identified in the present review were used in multiple contexts (e.g., patient completed, face-to-face interview, and telephone interview) within the same trial, without evidence of testing for appropriateness across the different methods of administration. A second example relates to the age-appropriateness of instruments. The use of an instrument intended for adults in a pediatric population would need—as a minimum—some adaptation reflecting the cognitive distribution of particular age groups.

There are limitations to our analysis. First, we focused exclusively on UK studies and the fields contained within the Database of Instruments for Resource-Use Measurement (DIRUM) may not be appropriate to resource-use instruments used in other countries. Furthermore, the methods of resource-use data collection invariably differ between health care systems, with less reliance on patient recall in private or third-party payer health systems where bills are itemized and charges may be substituted for costs. Second, our sample was limited to instruments used within clinical trials. Other formats of data collection (e.g., national surveys) were not considered. There is no reason, however, to exclude these, or instruments from other countries, and to modify the data fields within DIRUM, accordingly.

Our systematic approach of defining the scope, pertinent fields for data extraction, and feasibility of establishing a repository of resource-use instruments based on patient recall has resulted in the DIRUM (accessible via [www.DIRUM.org](http://www.DIRUM.org)). It supports data navigation, sorting, searching, advanced filtering, record addition, modification, deletion, and file uploads. A conceptual structure of the database and field content is presented in Figure 1. A key feature is the author submission form that allows questionnaire developers to submit their instruments for inclusion on the database.

An important intended application of DIRUM is to facilitate further research in the area of resource-use data collection. There are several questions that are hitherto unanswered, and it is beyond the scope of this article to provide specific guidance. There is, however, a clear need to establish a research agenda that encompasses questions such as the following:

- Effect of question sequencing. Should resource items that are the principal cost drivers be asked first? What is the effect of the positioning of the resource utilization questions among other sections of case report forms or diaries?
- Free-text entries. Are there potential benefits of free-text entry, when balanced against the burden of data management?



**Fig. 1 – Conceptual structure of database and field content. Notes:** Items in bold are essential criteria needed for each resource-use instrument. Without these, an instrument will not be included in the database. Tables 1 to 3, respectively, relate to instruments' characteristics, qualities, and additional information. Each instrument is characterized by at least one (but sometimes more) reference relating to its use or development (Table 4).

- Optimal recall length [25]. The trade-off between recall period—often defined by trial protocol and/or patient burden—and the accuracy of cost estimates warrants further empirical analysis.
- Accuracy in recall of medical terms. Are patients able to provide reliable responses in relation to their medicines or distinguish (for instance) between a clinical psychologist and a psychiatrist?
- Ability of instrument (and respondents) to distinguish between resource use related/unrelated to the context of use (specificity of the instrument).
- Comprehensiveness of the items recorded by the instruments (sensitivity of the instrument) and the related issue of length of questionnaire and potential burden on respondent.
- Missing data. Assessment of the underlying causes of missing data in resource-use questionnaires based on patient recall and analysis of missing data to minimize biased cost estimates [26].
- Reliability testing. Defining applicability and appropriateness of different methods to assess for and to ensure the reliability of instruments.
- Validity testing. Defining applicability and appropriateness of different methods to assess for and to ensure the validity of instruments.
- Appropriateness (or otherwise) of generic resource-use instruments or defining a core set of resource-use measures.
- Transferability. Appropriateness of using an instrument in different populations, settings, or context.

- Cognition and readability. To what extent do respondents comprehend the questions?
- To what extent might large national health surveys, such as the Health Survey for England, the Welsh Health Survey, Australia's National Health Survey, or the Medical Expenditure and National Health Interview Surveys in the United States, inform or be informed by trial-based data collection instruments?
- Challenges associated with the development of resource-use instruments for multinational trials.

The ability to access resource-use instruments via DIRUM, and to choose among instruments or parts of instruments that are applicable to specified areas of disease, population, or settings of care, will provide a valuable resource to health economists and help encourage more focus on future instrument development.

Source of financial support: This study was funded by the Medical Research Council, as part of the North West Hub for Trial Methodology Research.

## REFERENCES

- [1] Ridyard CH, Hughes DA. Methods for the collection of resource use data within clinical trials: a systematic review of trials funded by the UK Health Technology Assessment programme. *Value Health* 2010;13: 867–72.
- [2] Emery M, Perrier L, Acquadro C. Patient-Reported Outcome and Quality of Life Instruments Database (PROQOLID): frequently asked questions. *Health Qual Life Outcomes* 2005;3:12.
- [3] JiscMAIL. HEALTHECON-ALL website. Available from: <http://www.jiscmail.ac.uk/HEALTHECON-ALL>. [Accessed March 1, 2011].
- [4] SurveyMonkey. Available from <http://www.surveymonkey.com>. [Accessed March 1, 2011].
- [5] Wordsworth S, Thompson S. An annotated cost questionnaire for patients: results of piloting. *HERU Discussion Paper* 03/01, 2001:1–91.
- [6] Knapp M, Beecham J. Costing mental health services. *Psychol Med* 1990;20:893–908.
- [7] Patel A, Rendu A, Moran S, et al. A comparison of two methods of collecting economic data in primary care. *Fam Pract* 2005;22:323–7.
- [8] Knapp M, Scott S, Davies J. The cost of antisocial behaviour in younger children. *Clin Child Psychol Psychiatry* 1999;4:457–73.
- [9] Beecham J, Knapp M. Costing psychiatric interventions. *Discussion Paper* 1536, 1999:22.
- [10] Curtis L. Unit costs of health and social care 2010. Available from: <http://www.pssru.ac.uk/pdf/uc/uc2010/uc2010.pdf>. [Accessed March 1, 2011].
- [11] Street A, Dawson D. Costing hospital activity: the experience with healthcare resource groups in England. *Eur J Health Econ* 2002; 3:3–9.
- [12] Department of Health. NHS reference costs 2009–2010. Available from: [http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH\\_123459](http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_123459). [Accessed March 1, 2011].
- [13] Joint Formulary Committee. *British National Formulary* (61st ed.). London, UK: British Medical Association and Royal Pharmaceutical Society of Great Britain, 2011. Available from: <http://www.bnf.org/bnf>. [Accessed March 7, 2011].
- [14] The Cochrane Library. Available from: <http://www.thecochranelibrary.com/view/0/index.html>. [Accessed April 6, 2011].
- [15] Richardson G, Kennedy A, Reeves D, et al. Cost effectiveness of the Expert Patients Programme (EPP) for patients with chronic conditions. *J Epidemiol Community Health* 2008;62:361–7.
- [16] Bannigan K, Watson R. Reliability and validity in a nutshell. *J Clin Nurs* 2009;18:3237–43.
- [17] Thalanany MM, Mugford M, Mitchell-Inwang C. Visiting adult patients in intensive care: the importance of relatives' travel and time costs. *Intens Crit Care Nurs* 2006;22:40–8.
- [18] Fox M, Voordouw J, Mugford M, et al. Economic costs of food allergies in Europe: development of a questionnaire to measure costs and health utility. *Health Serv Res* 2009;44:1662–78.
- [19] Lewis SW, Davies L, Jones PB, et al. Randomised controlled trials of conventional antipsychotic versus new atypical drugs, and new atypical drugs versus clozapine, in people with schizophrenia responding poorly to, or intolerant of, current drug treatment. *Health Technol Assess* 2006;10:1–165.
- [20] Evans CJ, Crawford B. Data collection methods in prospective economic evaluations: how accurate are the results? *Value Health* 2000;3:277–86.
- [21] Cooper NJ, Mugford M, Symmons DPM, et al. Development of resource-use and expenditure questionnaires for use in rheumatology research. *J Rheumatol* 2003;30:2485–91.
- [22] Jolly K, Taylor R, Lip GY, et al. Birmingham Rehabilitation Uptake Maximisation Study (BRUM). Home-based versus hospital-based cardiac rehabilitation in a multi-ethnic population: cost-effectiveness and patient adherence. *Health Technol Assess* 2007;11:1–98.
- [23] Chisolm D, Knapp M, Knudsen H, et al. Client Socio-Demographic and Service Receipt Inventory-European Version: development of an instrument for international research. *Br J Psychiatry* 2000;177(Suppl.): s28–33.
- [24] Bhandari A, Wagner T. Self-reported utilization of health care services: improving measurement and accuracy. *Med Care Res Rev* 2006;63:217–35.
- [25] Clarke PM, Fiebig DG, Gerdtham UG. Optimal recall length in survey design. *J Health Econ* 2008;27:1275–84.
- [26] Noble SM, Hollingworth W, Tilling K. Missing data in trial-based cost-effectiveness analysis: the current state of play. *Health Econ* December 15, 2010; DOI: 10.1002/hec.1693. Epub ahead of print.