Measuring health system resource use for economic evaluation: a comparison of data sources

CHRISTINE POLLICINO, ROSALIE VINEY AND MARION HAAS

Christine Pollicino is a Research Officer at the Centre for Economics Research and Evaluation (CHERE).
Rosalie Viney is a Deputy Director of CHERE and Senior Lecturer in Health Economics at the University of Technology, Sydney.
Marion Haas is a Deputy Director of CHERE and Senior Lecturer in Health Services Research at the University of Technology, Sydney.

Abstract

A key challenge for evaluators and health system planners is the identification, measurement and valuation of resource use for economic evaluation. Accurately capturing all significant resource use is particularly difficult in the Australian context where there is no comprehensive database from which researchers can draw. Evaluators and health system planners need to consider different approaches to data collection for estimating resource use for economic evaluation, and the relative merits of the different data sources available. This paper illustrates the issues that arise in using different data sources using a sub-sample of the data being collected for an economic evaluation. Specifically, it compares the use of Australia's largest administrative database on resource use, the Health Insurance Commission database, with the use of patient-supplied data. The extent of agreement and discrepancies between the two data sources is investigated. Findings from this study and recommendations as to how to deal with different data sources are presented.

Identification, measurement and valuation of resource use

Economic evaluation of health care involves the systematic comparison of alternative interventions in terms of consequences (impacts on health and other outcomes) and costs (resource use) (Drummond, O'Brien et al. 1997). It provides information about which interventions are likely to maximise outcomes from available resources. There are two components to economic evaluation: identification, measurement and valuation of consequences; and identification, measurement and valuation of resource use.

The appropriate approach to identifying, measuring and valuing resource use is to identify the opportunity costs of additional resources used. We are interested in the net impact on health system resources as a result of the intervention. This assessment of costs should be made at the margin: as a result of the intervention, what is the value of additional resources used, and what is the value of any additional resources freed up?

There is a range of different approaches to economic evaluation. Economic evaluations can be undertaken as decision tree or Markov modelling studies, where the choice of interventions is set up in a decision tree framework and the resource use and outcomes for an average patient are assessed (Drummond, O'Brien et al. 1997). Other economic evaluations may involve detailed assessment of resource use in a clinical trial setting: collecting individual data on all patients in the control and intervention groups. Health system planners and decision makers may also undertake partial economic evaluations (costing studies), involving examination of the costs of different interventions without the comparison between alternatives, and the assessment of

consequences. In all instances, resource use may be collected prospectively or retrospectively. Regardless of the purpose and nature of an economic evaluation or costing study, the same theoretical framework is appropriate.

A key challenge for evaluators and health system planners is the identification of appropriate data for economic evaluation. There are a number of different methods that can be used, and data sources available. Each has its limitations and its own implications for costing methods. The purpose of this paper is to identify, in the Australian context, the sources of data for economic evaluation, to discuss their limitations, and to draw conclusions about how to use the different data sets. In particular, the paper involves a comparison of two common methods for costing health system resource use: the use of patient-supplied and administrative data. We report on the extent of agreement between these two sources in a specific study.

The next section of the paper discusses the methods and possible data sources for costing health services in Australia, how they might be used and the issues involved in using each of them. In particular, the focus of this paper is on the collection of actual resource use data for the purposes of costing. Then the methods and results of a study involving data from two different sources are presented and discussed. The final section of the paper provides some guidelines for undertaking economic evaluations.

Methods of costing

As discussed above, there are two main approaches to costing for economic evaluation, or for other purposes (such as resource allocation, priority setting or planning):

- Construction of decision analytic models of the resource use and outcomes associated with typical treatment pathways for alternative interventions; and
- Observation and measurement of the differences in costs at the individual level for individuals undergoing alternative interventions.

In the first approach, decision tree models or other similar models (for example, a Markov model) are constructed to estimate the overall net impact on resource use of alternative interventions for a group of patients. The difference in average resource use can then be estimated. The margin is the difference between the two interventions. The identification and measurement of resource use in this type of study is often based on protocols, on estimates of outcomes and the associated resource use from other published studies and expert opinion, rather than on observation of individuals. The process of costing involves determination of what typically happens at each node of the decision model, and attaching appropriate values. For example, the appropriate values may be the Medicare Benefits Schedule fee or the Australian Refined Diagnosis Related Groups (AR-DRG) cost. This approach to costing is deterministic, and although individual level data may sometimes be used as an input to the costing in the model, the resource use estimated is that of an average pathway of treatment.

The second approach is generally a bottom-up method of estimating the resource use associated with different interventions. Data are collected on the resource use and outcomes for individual patients undergoing alternative interventions. Economic evaluation based on actual resource use is preferable for a number of reasons. First, given it involves the collection of individual level data, it provides a more accurate reflection of the variability in resource use across patients. Second, it may reveal differences in resource use between treatment groups that may not be evident from modelled costs. Lastly, collection of individual resource use data from a sample of patients allows for statistical analysis of costs.

The data collection for this approach may be prospective or retrospective. Retrospective data collection involves identifying groups of patients who have undergone the different interventions being evaluated (for example in a case-control study), and extracting information about resource use from medical records and administrative databases. The advantage of retrospective data collection is that it is often a cheap and efficient way to collect the necessary resource use information, because it is already recorded in administrative databases and medical records. The disadvantages of retrospective data collection are that not all relevant resource use may have been recorded, the samples of patients may be biased, and it is often difficult to obtain the appropriate ethics approvals to extract the required information.

Prospective data collection involves following patients who have been allocated to the different interventions (preferably in a randomised controlled trial, but it may also be in an observational study), and collecting all

relevant resource use information at the time that treatment occurs. The advantages of prospective data collection are that it is possible to identify and develop data collection methods for all resource use that is likely to be relevant, there is more control over selection of patients and study design, it is more likely that patient and clinician permission to access the relevant information can be obtained, and more options for methods of data collection are available. Particularly, it allows for both administrative and patient-supplied data to be collected. However, the costs and time delays for collection of prospective data often make it impractical.

In any one economic evaluation, both approaches may be used. A decision tree or Markov modelling approach can be enhanced by the use of observational data on individual resource use for particular aspects of treatment (for example, to obtain an accurate estimate of the average number of follow-up tests required to confirm a particular diagnosis following a screening test, or to estimate the proportion of patients requiring re-admission following a particular procedure). Alternatively, even where an economic evaluation is undertaken in the context of a randomised controlled trial with observation of individual level data, it may be necessary to model particular components of treatment, especially if estimation of total costs and outcomes for the intervention goes beyond the timeframe or scope of the trial. For example, the endpoint of the trial may be a particular event, and yet many of the relevant costs may occur as a result of this event.

Further, even when individual resource use data are able to be collected, the process of costing will require some assumptions in order to attach a cost to each component of the treatment received. A range of approaches may be used here, but it is important that they are logical and that the assumptions used are clear. For example, hospital episodes may be costed using relevant average AR-DRG costs, which themselves have been estimated based on modelling of individual level data (Hindle, 1999), or there may be within the study a much more detailed process of attribution of costs, such as using the length of time of a consultation as a basis to attribute salary costs to individual patients.

Within one economic evaluation it may also be necessary to use both prospective and retrospective data collection. For example, an intervention may involve periods of hospital stay and follow-up care in the community. It may not be appropriate to collect actual resource use and cost data for the hospital stay – this would only reflect the costs in that particular hospital. Instead, individual data may be collected at the level of AR-DRGs for the hospital episodes and then costed using published cost-weights. However, for the follow-up care, individual numbers of visits may be counted to estimate costs of this component. On this point, it is useful to make a distinction also between how the resource use data are collected and how values are attached to resource use.

Resource use data sources

The challenge that individual level costing for economic evaluation poses is the collection of data. There are two main sources of resource use data; administrative and patient-supplied.

Administrative data have been routinely collected to meet the administrative needs of service providers. For example, the Health Insurance Commission (HIC) collects data on provider service claims for payment purposes. The advantages of administrative data include its comprehensiveness, possibility of accessing much larger samples and availability of longitudinal data. Thus this source of data is extremely valuable for estimating use of health services. However, there are several limitations associated with the use of these data, as follows:

- Administrative data are not designed for research purposes. Researchers are thereby restricted to measure those resources collected in the database.
- Informed patient consent needs to be obtained before extraction of individual patient level data can be made. This consent must specify the time period over which the data can be collected. This can be difficult if the data collection is retrospective.
- Researchers often need to wait a certain period of time to ensure the majority of claims have been processed. For the HIC, complete data are not available until at least six months after the last date of service, and then it may take another month for data to be processed. This can delay the completion of the costing considerably.
- These data typically lack clinical detail.
- Administrative data may suffer from programming and coding errors, and little is known about the extent of these errors.

Data retrieved from medical records is another example of administrative data. These data are more readily available than other data sources and contain clinical detail. However, collection may be time-consuming, often requiring a person with clinical expertise to extract the required data.

An alternative to the use of administrative data is patient-supplied data. There can be collected in a number of ways, such as through the use of a diary or conduct of an interview. Irrespective of the method of collection, these data are typically collected prospectively over a period of time giving researchers an opportunity to collect variables of interest. This source of data also suffers from several problems, including the following:

- It is subject to memory decay, recency bias, and external and internal telescoping (McCullum, Lonergan et al. 1993). Memory decay is when more distant events tend to be forgotten. Recency bias is when recent events blot out the memory of more distant events. External telescoping is the incorporation of services occurring outside the defined time period into the time period. Internal telescoping is the tendency to bunch recalled events at more recent times in the time period.
- Burden rests on respondents.
- Informed patient consent is required prior to data collection.
- Data collection and processing is complex.

The next section of this paper provides an example of how the two main sources of data (HIC and patient-supplied data) compare. We report on the extent of agreement in a small study undertaken as part of a larger randomised controlled trial. For the purposes of this paper, only a small sample of the total resource use data collected for the trial will be analysed.

Method

Data on resource use were collected from two independent sources; the patient, and HIC. These data were obtained for patients with apparently resectable non-small-cell lung cancer registered in a larger randomised controlled trial. The data presented are a sub-sample of all the data collected for the trial. We obtained this sub-sample early in the overall follow-up for the trial for the purposes of comparing the two methods of data collection and ensuring that the HIC data would meet the requirements of the costing to be undertaken for the trial.

As part of the randomised controlled trial, patients were asked to complete diaries relating to their health system resource use and consent to have their Medicare data retrieved from HIC and/or the Department of Veterans' Affairs (DVA). The resource use diaries contained questions about the use of general practice, specialist care, pathology, hospitals, chemotherapy, radiotherapy, community or home based care and alternative therapies. Diaries were completed on a fortnightly basis for the first year of follow-up and then monthly the year after. Prepaid envelopes were provided to encourage response. Patients were contacted by telephone if diaries were not returned within a reasonable period of time.

Medicare data were extracted from HIC. These data contain detailed service information for the Australian population from the inception of Medicare in February 1984 to the present. DVA granted permission for HIC to release Medicare data relating to Veterans. To ensure that the majority of claims made by patients in the trial had been presented and processed by HIC and DVA, the request for data was made six months after the date of the last service being rendered. The data covered services that qualified for Medicare Benefit such as physician and specialist attendances, pathology tests, diagnostic imaging, and radio and nuclear therapy. Medical services for privately admitted patients in public and private hospitals were also included. Not included were services provided by hospital doctors to public patients in public hospitals or services as a result of a compensation or insurance claim.

Patient details (surname, first name, date of birth, sex) along with the patient's Medicare number were supplied to HIC for record matching. Matched Medicare data were compared with data reported by the patient. The type and number of services utilised were compared between the data sources. The numbers of patient-reported services were reviewed by two researchers independently. Limitations of Medicare data meant in-hospital episodes, community or home based care, and alternative therapies were excluded from this comparison. Up to four months of resource use data were compared for each patient.

Results

Medicare data were requested for 24 patients. Of these, one patient could not be matched by HIC and DVA, and was excluded from subsequent analysis. A total of 177 dairies were returned. 19 patients completed all eight diaries covering a four-month period, three patients completed six of the eight diaries required and one patient returned seven. Missing diaries (4 patients) were supplemented with information supplied by the patient's health care provider.

Table 1 presents the extent of agreement between the two data sources by broad type of service. The largest discrepancy in the number of services utilised occurred in the documentation of pathology tests (93%), followed by unreferred attendances covering vocational registered general practitioners and others (87%), diagnostic imaging (75%) and specialist attendances (59%).

Table 1: Extent of agreement between HIC/DVA data and patient diaries by broad type of service (BTOS)

	BTOS	N	HIC/DVA & patient agreed on # of services utilised	HIC/DVA & patient disagreed on # of services utilised
Unreferred attendances	A & B	23	3	20
Specialist attendances	C	22	9	13
Diagnostic imaging	G	20	5	15
Pathology tests	F & N	14	1	13

A comparison of the number of unreferred and specialist services recorded between data sources appears in Table 2. Overall, Medicare data identified a greater number of unreferred and specialist attendances being utilised. Absolute differences in unreferred and specialist attendances ranged from one to ten and one to three respectively.

Nine patients consistently under-reported the number of unreferred and specialist attendances utilised. Two patients over-reported both service attendance types. Only two patients agreed with HIC/DVA data when recording both the number of unreferred and specialist attendances utilised.

Table 2: Unreferred and specialist attendances compared between data sources

Patient #	Unreferred attendances			Specialist attendances		
	# reported by patient	# recorded by HIC/DVA	Difference	# reported by patient	# recorded by HIC/DVA	Difference
1	15	5	10	6	4	2
2	7	3	4	0	1	-1
3	8	4	4	0	3	-3
4	6	4	2	1	2	-1
5	5	4	1	3	2	1
6	7	6	1	1	1	0
7	7	6	1	3	3	0
8	5	4	1	1	1	0
9	8	8	0	1	1	0
10	10	10	0	2	2	0
11	11	11	0	2	3	-1
12	10	11	-1	5	5	0
13	11	12	-1	3	3	0
14	3	4	-1	1	4	-3
15	3	4	-1	0	3	-3
16	4	6	-2	2	2	0
17	4	6	-2	0	2	-2
18	3	5	-2	1	5	-4
19	4	6	-2	-	-	-
20	7	10	-3	2	4	-2
21	4	9	-5	0	3	-3
22	7	16	-9	0	5	-5
23	8	19	-11	1	2	-1
TOTAL	157	173	-16	35	61	-26

Summary of results

There is little agreement between the data sources. The majority of disagreement occurred in the documentation of pathology tests, unreferred attendances and/or diagnostic imaging. Patients tended to under-report resource use when compared to HIC and DVA data.

The different data sources result in different estimates of resource use. The difference in resource use, specifically unreferred and specialist attendances, between sources can be estimated at $1,619.80 (using standard consultation fees). Potentially, this could change the ranking of relative cost-effectiveness.

Discussion

Discrepancies in resource utilisation between administrative and patient-supplied data sources have been previously reported in the literature. Differences have been noted in the reporting of physician consultations, hospital admissions and specific procedures such as Pap tests (Glandon, Counte et al. 1992; McCullum, Lonergan et al. 1993; Marshall, Grayson et al. 2001; Norrish, North et al. 1994; Hancock, Sanson-Fisher et al. 1998).

Two of the published studies have been undertaken in Australia. McCallum et al. (1993) conducted a pilot study which explored whether self-reported general practitioner and specialist utilisation were adequate substitutes for Medicare data extracted from HIC. The authors found that self-reported data resulted in under-reporting and extreme outliers, and that good measures of patterns of use and accuracy of recording were only gained from HIC data.

Marshall et al. (2001) compared self-reported medical care consumption to matched HIC and DVA data for Australian veterans. They showed that administrative data confirmed only 51% of veterans self reporting medical care consumption and that there was up to 30% over-estimation of actual services used in the two-week reference period.

Several studies have attempted to explain the discrepancy with health status and socio-demographic variables (Glandon, Counte et al. 1992; McCullum, Lonergan et al. 1993). In this study, there are a number of possible explanations for the observed discrepancies.

A small proportion of the disparity in the number of pathology tests could be explained by differences in the recording of data. For example, multiple pathology tests would be recorded separately on the HIC database but they were recalled as a single event by the patients.

There is a trade-off in designing patient diaries between obtaining sufficient information to ensure there is no double counting or missed services, and ensuring the burden on patients is not excessive. In this study, patients were not explicitly asked for the date of their consultation. Had we done so, patients would have had to complete diaries more frequently increasing the burden on patients. Duplication of resource use across fortnightly diaries was evident for a small number of patients who recorded the date of their unreferred attendances. The inclusion of consultation dates to patient diaries would be valuable in understanding the extent of duplication of resource use, but inclusion is at a trade-off to patient burden.

The duplication of resource use within diaries may also suggest that some patients have had difficulty in understanding the questions asked. For example, patients could not distinguish between a visit to their local doctor's office and a doctor's visit to their home.

Differences in diagnostic imaging use could be explained by the nature of treatment follow-up. Routine follow-up of patients with non-small-cell lung cancer includes periodic chest x-rays, and bone scans. It was these periodic services which were often not recorded by the patient. Further development of the patient diary should take this into account. One corrective action might be to ask patients explicitly if they had a chest x-ray and then bone scan singularly rather than asking a general question.

Conclusion

Often study constraints and convention dictate the source of resource data used for an economic evaluation. In general, validation studies use administrative data as the gold standard to compare resource use. However, in the context of the Australian health care system there is no comprehensive source of resource use data from which researchers can draw. For example, HIC data does not include public patient hospital services. Similarly, patient-supplied data does not provide specific pathology data. Linkage of data sources is therefore needed to capture complete resource use.

Data linkage involves matching multiple data sources on the basis of user-determined criteria. Linkage can be done using name, address, birth date or sex. A number of pilot studies have been conducted to date. The Commonwealth Department of Health and Aged Care makes mention of these studies in a recent publication (Commonwealth Department of Health and Aged Care, 2000). These have found record linkage to be feasible.

However, they have also identified the challenge and limitations of such an approach especially outside the context of a prospective trial – such as obtaining signed patient consent prior to the release of data.

In the absence of record linkage, we recommend that both sources of data be collected and compared for a small sub-sample of the population. This exercise will provide researchers with an opportunity to gauge the extent of the agreement between sources and make informed decisions as to which data source is more appropriate for the population of interest. It will also assist researchers to define plausible ranges for specific variables to be used in subsequent sensitivity analysis. When the collection of both data sources is not possible, we suggest the use of HIC/DVA data if timeliness is not an issue and this database captures the majority of significant resource use. Where HIC data does not capture the resources of interest (for example, community care), diaries remain an essential research tool.

References

Commonwealth Department of Health and Aged Care 2000, *Insights into the utilisation of health services in Australia based on linked administrative data*. Canberra, Commonwealth of Australia.

Drummond MF, O'Brien B, Stoddart GL & Torrance G 1997, *Methods for the economic evaluation of health care programmes*. Oxford, Oxford Medical Publications.

Glandon GL, Counte MA & Tancredi D 1992, An analysis of physician utilisation by elderly persons: Systematic differences between self-report and archival information. *Journal of Gerontology*, 47: ppS245-253.

Hancock L, Sanson-Fisher R & Kentish L 1998, Cervical cancer screening in rural NSW: Health Insurance Commission data compared to self-report. *Australian & New Zealand Journal of Public Health*, 22: pp307-12.

Hindle D 1999, *Product costing in Australian hospitals: a background guide to national costing work*. Canberra, Commonwealth of Australia.

Marshall R, Grayson D, Jorm A & O'Toole B 2001, Are survey measures of medical care utilisation misleading? A comparison of self-reported medical care consumption with actual medical care utilisation. *Australian Health Review*, 24: pp91-99.

McCullum J, Lonergan J & Raymond C 1993, *The NCEPH record linkage pilot study: A preliminary examination of individual Health Insurance Commission records with linked data sets*. Canberra, National Centre for Epidemiology and Population Health, The Australian National University.

Norrish A, North D, Kirkman P & Jackson R 1994, Validity of self-reported hospital admission in a prospective study. *American Journal of Epidemiology*, 140: pp938-42.