



OPEN ACCESS

Evaluation of a predevelopment service delivery intervention: an application to improve clinical handovers

Guiqing Lily Yao,¹ Nicola Novielli,¹ Semira Manaseki-Holland,¹ Yen-Fu Chen,¹ Marcel van der Klink,² Paul Barach,^{3,4,5} Peter J Chilton,¹ Richard J Lilford,¹ on behalf of the European HANDOVER Research Collaborative*

► Additional supplementary files are published online only. To view these files please visit the journal online (<http://dx.doi.org/10.1136/bmjqs-2012-001210>).

For numbered affiliations see end of article.

Correspondence to

Dr Richard J Lilford, Department of Public Health, Epidemiology and Biostatistics, University of Birmingham, Edgbaston, West Midlands, Birmingham B15 2TT, UK; r.j.lilford@bham.ac.uk

Received 22 May 2012
Accepted 7 August 2012
Published Online First
13 September 2012

ABSTRACT

Background: We developed a method to estimate the expected cost-effectiveness of a service intervention at the design stage and ‘road-tested’ the method on an intervention to improve patient handover of care between hospital and community.

Method: The development of a nine-step evaluation framework:

1. Identification of multiple endpoints and arranging them into manageable groups;
2. Estimation of baseline overall and *preventable* risk;
3. Bayesian elicitation of expected effectiveness of the planned intervention;
4. Assigning utilities to groups of endpoints;
5. Costing the intervention;
6. Estimating health service costs associated with preventable adverse events;
7. Calculating health benefits;
8. Cost-effectiveness calculation;
9. Sensitivity and headroom analysis.

Results: Literature review suggested that adverse events follow 19% of patient discharges, and that one-third are preventable by improved handover (ie, 6.3% of all discharges). The intervention to improve handover would reduce the incidence of adverse events by 21% (ie, from 6.3% to 4.7%) according to the elicitation exercise. Potentially preventable adverse events were classified by severity and duration. Utilities were assigned to each category of adverse event. The costs associated with each category of event were obtained from the literature. The unit cost of the intervention was €16.6, which would yield a Quality Adjusted Life Year (QALY) gain per discharge of 0.010. The resulting cost saving was €14.3 per discharge. The intervention is cost-effective at approximately €214 per QALY under the base case, and remains cost-effective while the effectiveness is greater than 1.6%.

Conclusions: We offer a usable framework to assist in ex ante health economic evaluations of health service interventions.

INTRODUCTION

The decision to *adopt* a service level intervention turns on the evidence of its effectiveness. However, decisions also have to be made to develop an intervention in the first place, and on whether to continue development once started.^{1 2} This paper is concerned with estimating the potential cost-effectiveness of service interventions at the design and development stages, that is, at the formative stage before the intervention is rolled out in practice. This type of early economic evaluation would take place behind company doors in the case of pharmaceuticals. However, it can take place in the health services in the case of most service delivery interventions, which are typically developed ‘in the service, by the service, for the service’.³

Service delivery interventions have been classified as targeted (near patient) and generic (far patient).⁴ Targeted interventions (eg, a guideline to use thromboprophylaxis before surgery, or a forced function to prevent misconnecting the oxygen supply) are relatively easy to evaluate for effectiveness and cost-effectiveness.⁵ However, generic interventions (such as an improved nurse-to-patient ratio, or changing safety culture) have diffuse effects that spread out to affect many clinical processes and outcomes, as shown in [figure 1](#). A recent systematic review⁶ demonstrates that most economic evaluations of such interventions are based on cost minimisation only. These studies do not capture the non-monetary value of improved outcomes for which cost utility/benefit analyses are necessary. However, our recent review (Bramley and Lilford—in preparation) revealed no articles discussing the methodological implications of this type of study for generic service interventions.

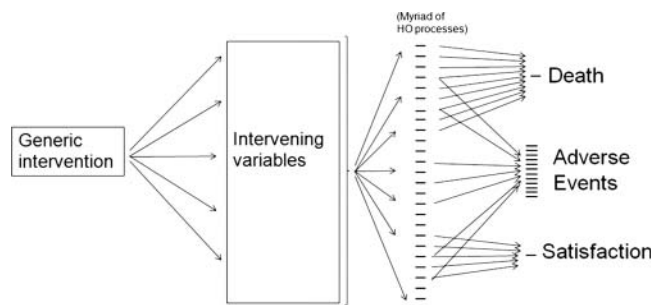


Figure 1 Representation of the widespread effects of a generic intervention. Endpoints such as mortality, or those measuring satisfaction partially measure the effect of the intervention. A sensible grouping of adverse events allows the measurements of different dimensions of effectiveness.

We found no articles dealing with preimplementation assessments of service delivery interventions. In this paper, we describe a method for prospective evaluation of generic service delivery interventions. We demonstrate how it can be applied using a particular generic intervention—the ‘HANDOVER’ project (<http://www.handover.eu>), as a case study. Briefly, this was a European Union funded project to develop an educational intervention to improve patient care at the point of discharge from hospital to the community; that is, a service delivery/educational intervention to improve clinical handover of the patient from one clinical microsystem to the next.⁷ The intervention was based on earlier (exploratory) studies conducted within the HANDOVER project itself, and also knowledge of the clinical and behavioural literature dealing, for example, with the importance of clinical skills⁸ and teamwork in improving clinical transitions.⁹ The HANDOVER project is described elsewhere in this issue of the journal, and in a report to the European Commission.¹⁰ We deal here with the most intensive form of the intervention, which consists of classroom instruction, supported by a number of internet-based educational resources known as the ‘Handover Toolbox’.

METHODS

Framework for the evaluation of a generic service delivery intervention at the design phases

A framework for evaluation of a generic service delivery intervention was developed based on nine steps. The first step was concerned with identifying and classifying the endpoints for the evaluation. Steps 2–6 were used to derive input data. Steps 7–9 were concerned with the calculations of health benefit, net costs and cost-effectiveness.

Step 1: identification and classification of suitable endpoints

We selected endpoints on the basis of expert opinion informed by an inventory of endpoints identified from

systematic reviews of the topic. As described above (and demonstrated in figure 1), the number of endpoints that might be affected by a generic service intervention is large when compared with those affected by targeted service or clinical interventions.¹¹ Ensuring that the range of salient endpoints is captured requires expert knowledge. In the case of the HANDOVER project this was acquired from the scientific management committee, informed by an inventory of endpoints used in previous reviews.^{12–15} Service interventions are often designed to improve the quality and safety of care, that is, reduce the number of adverse events. Each of these events is associated with a possible contingent cost saving. In the case of a cost utility/benefit evaluation, each of these events has to be valued (in the sense that a preference weighting/utility must be assigned). It would be difficult to assign cost savings or utilities to each endpoint because there are so many. Yet, assigning one utility to cover all endpoints seems too coarse given the very different nature of the salient endpoints. Our solution was to categorise the endpoints, on the basis of the severity and duration of the associated morbidity, in order to provide a manageable number of categories.

Step 2: estimation of the baseline risk associated with each endpoint

The identification of suitable endpoints (or clusters of endpoints) is followed by an estimation of the baseline risk related to each. There are two kinds of baseline risk to consider: overall baseline risk for that endpoint, and the reversible component of that risk. This is analogous to the concepts of risk and attributable risk in classical epidemiology.¹⁶ In the service context, we are interested in the distinction between the bad outcome in question and the proportion of that bad outcome that is amenable to improved care (ie, that is preventable). For example, if the death rate is 10%, and 20% of those deaths are preventable, then the preventable risk is 2%.¹⁷ This establishes a plausible upper bound for the overall effectiveness of the proposed intervention.

Step 3: elicitation of expected effectiveness

Summative (demand-side) evaluation of interventions is based on direct empirical evidence, such as that derived from head-to-head comparisons. At the design and early development stages, however, such empirical evidence is absent or scant. Expert opinion must therefore suffice for formative (supply-side) evaluations. Such expert opinion must be quantified if it is to serve as the basis for a cost-effectiveness analysis. The necessary estimates of effectiveness should be elicited from domain experts who do not have a psychological or material personal stake in the outcome.^{18 19} Before they make their

judgements, the experts should be party to a detailed description of the proposed intervention, and their knowledge should be brought up to date by drawing their attention to recent relevant reviews.¹⁹

The elicitation is typically used to construct a Bayesian 'prior' in the form of a 'density' of prior probabilities over a range of parameter values. The methods we used for eliciting a Bayesian prior have been described elsewhere.¹⁹ The centre of the probability distribution represents an expert's best guess for the effectiveness of an intervention. The priors from multiple experts can be amalgamated mathematically to yield a consolidated prior, the centre of which is a collective best guess on an effectiveness parameter. The statistical aspects of Bayesian inference are described elsewhere.^{19–22} Ideally, the elicitation should be repeated for each endpoint or endpoint type. Arguably, this requirement can be relaxed if the endpoints are all highly correlated and thought to belong to the same 'latent class'.²³

Step 4: estimating the utility values

The utility measures of disability in the case of clinical interventions are required for only one (or a small number of) health states. For example, patients with multiple sclerosis can be evaluated with respect to a disease-specific quality-of-life measure that has been mapped onto a utility scale. However, in the case of a generic service delivery intervention, such as a clinical handover intervention, a very large number of adverse events may be affected. Each event is associated with a loss of utility, but only a few of these have been calibrated. To circumvent the logistic problem of eliciting utilities for each and every adverse event, we recommended (above) that the adverse events be grouped into similar categories with respect to severity and duration, thereby limiting the number of states for which utility estimates must be obtained. These utilities may be estimated in one of two ways:

1. The disability category is assigned a 'typical' score on a generic quality-of-life scale (eg, EuroQoL) that can be converted into a utility by means of a standard tariff;
2. A specific elicitation is made for the category, for example, by means of a standard gamble or a conjoint analysis.

Step 5: estimation of the costs of the intervention

Intervention costs are built up from fixed costs and variable costs in the usual manner. The fixed costs refer to the costs of the development, while the variable costs include the salary costs of the staff and materials deployed in delivering the intervention.

Step 6: healthcare costs associated with adverse events

Adverse patient events lead to increased health expenditure. The intervention costs are standard, but the costs generated through adverse events depend on the category of event; the more serious the event, the greater the costs associated with it on average. The estimates of costs associated with various types of endpoints may be available in the literature, and we were fortunate to find such a source of data for categories similar to those used in the HANDOVER study. In the absence of such data, we would have had to model cost-savings contingent on effectiveness (step 3) using national reference costs.

Step 7: expected health benefits and net costs

The expected health benefit (EHB) is a function of the expected effectiveness of an intervention. Resource allocation decisions require that the benefits resulting from changes in patient outcomes are captured on the same scale, for example, a Quality Adjusted Life Year (QALY) scale. Health states, each weighted by their corresponding utility value, are combined to produce the QALY. It is often convenient, for reasons we will explain below, to express the expected health gain as an expected monetary benefit (EMB). In order to convert a QALY gain to an EMB an assumption of societal willingness to pay (λ) is necessary:

$$\text{EMB} = \text{EHB}_{(\text{QALY gain})} \times \lambda$$

The net costs are calculated by subtracting the health service costs saved from the intervention costs. While this is the standard method, it was necessary to repeat it for each of the severity and duration categories identified. The results were then combined, and weighted by the relative frequencies of the various categories.

Step 8: estimation of cost-effectiveness and cost-benefit analysis

Cost-effectiveness analysis is based on combining the costs and benefits to calculate an incremental cost-effectiveness ratio (ICER) in the usual way. Suppose an intervention incurs additional costs (or saving) over usual practice (at ΔCOSTs), and the improved effectiveness measured on the QALY scale is ΔQALYs . Then, the ICER can be expressed as

$$\text{ICER} = \frac{\Delta\text{COSTs}}{\Delta\text{QALYs}}$$

When the expected benefit is expressed as a monetary value, we can estimate the expected net benefit (ENB) from an intervention as follows:

$$\text{ENB} = \text{EMB} - \Delta\text{COSTs} = (\Delta\text{QALYs} \times \lambda) - \Delta\text{COSTs}$$

Where λ is the society willingness to pay for a QALY gained.

Step 9: sensitivity analysis and headroom calculation

The above calculation can be repeated under various estimates of the effectiveness parameter drawn from the different points on the consolidated Bayesian prior. The headroom method is a way to explore the minimum health benefits that are needed to make an intervention cost-effective at a given level of societal willingness-to-pay value (λ). It can be calculated by the following equation:

$$\text{Min}(\Delta\text{QALYs}) = \frac{\Delta\text{COSTS}}{\lambda}$$

If the minimum necessary health benefit lies outside the bounds (say 95% credible interval) of the collective prior, then the intervention is unlikely to offer good value for money.

RESULTS—APPLICATION OF THE FRAMEWORK TO THE EMPIRICAL STUDY**The HANDOVER project**

The HANDOVER project described in the introduction and elsewhere in this issue was used as a case study for the nine-step framework described above.

Step 1: endpoints

The literature demonstrates that the patient readmissions and adverse events are the most frequent targets for interventions to improve patient handover.^{12 14 15} However, they are not independent, since preventable readmissions are contingent on preventable adverse events. The scientific management group for the study, therefore, selected adverse events as the salient

endpoint, both to model benefits to patients, and to estimate the healthcare savings that result from the reduced hospital admissions.

There are hundreds of possible adverse events, and it was therefore decided to categorise them into a manageable number. The best known classification was that used by Brennan and colleagues in the famous Harvard Medical Practice Study.²⁴ This system gives the relative frequencies in six categories of adverse events. Forster *et al*¹³ have produced an alternative system based on 400 patient discharges, specifically to assess adverse events following patient discharge, again with six severity levels. These two classification systems are placed side by side in **table 1**. We had to use expert judgement to ‘reconcile’ the two systems since they are different. For example, the proportion of deaths among patients who suffered adverse events, as cited by Brennan *et al*, seemed rather high at 13% for a discharge population; while no deaths, as found in the limited sample used by Forster *et al*, seemed optimistic. We therefore compromised at 5% (of the adverse events).

Step 2: estimation of the overall risk and preventable risk

We searched the literature and identified one paper by Forster *et al*¹³ in which the risk of adverse events following hospital discharge was estimated. The study found that at least one adverse event occurs in 19% of discharged patients within 1 month of discharge. These adverse events were preventable (ie, the result of a failure in handover practice) in approximately one-third of cases (ie, 6.3% of all discharges).¹³ The timeframe was considered appropriate as it has been shown that

Table 1 Comparison of two adverse event classification systems and the ‘reconciled’ version taken forward in this study

Forster <i>et al</i> ¹³ (patients experiencing adverse events after discharge)		Brennan <i>et al</i> ²⁴ (patients experiencing adverse events in hospital)		Reconciled	
Health Outcome	Proportion with the outcome	Health states	Proportion in each state	Health states	Proportion in each state
Death	0	Death	0.136	Death	0.05
Permanent disability*	0.03	Permanent disability >50%	0.026	Permanent disability >50%	0.02
		Permanent disability ≤50%	0.039	Permanent disability ≤50%	0.03
Readmission	0.21	Moderate impairment, recovery >6 months	0.028	Moderate impairment, recovery >6 months	0.10
A & E visit	0.11	Moderate impairment, recovery 1–6 months	0.137	Moderate impairment, recovery 1–6 months	0.30
Physician visit	0.14	Minimal impairment, recovery <1 month†	0.634	Minimal impairment, recovery <1 month	0.50
No extra use of health service	0.51				

*These patients were also assumed to be readmitted to hospital (the observed readmission in the study was 24%).

†Brennan *et al* included an ‘unclassified’ category for 6.5% of adverse events—we have subsumed this into the ‘minimal impairment’ category.

the risk of preventable patient readmissions following hospital discharge is much reduced after 1 month.²⁵

Step 3: the expected effectiveness of the intervention

The plausible estimates for expected effectiveness were elicited from ‘experts’. The elicitation process is described elsewhere.²⁶ Elicitation involved experts from Europe and the USA who attended a HANOVER project conference. The intervention (see Introduction) was described by researchers leading the workstream responsible for development of the intervention. Usable estimates of the effectiveness of the intervention were given by 24 experts. The individual probability densities were consolidated statistically.²⁷ The midpoint of this pooled prior indicates a 21% reduction in incidence of adverse events (figure 2). This represented a reduction from 6.3% to 4.7% for the risk of a preventable adverse event (an absolute reduction in risk of 1.3%). However, as shown in figure 2, there was considerable variation over the central estimate between experts. The experts

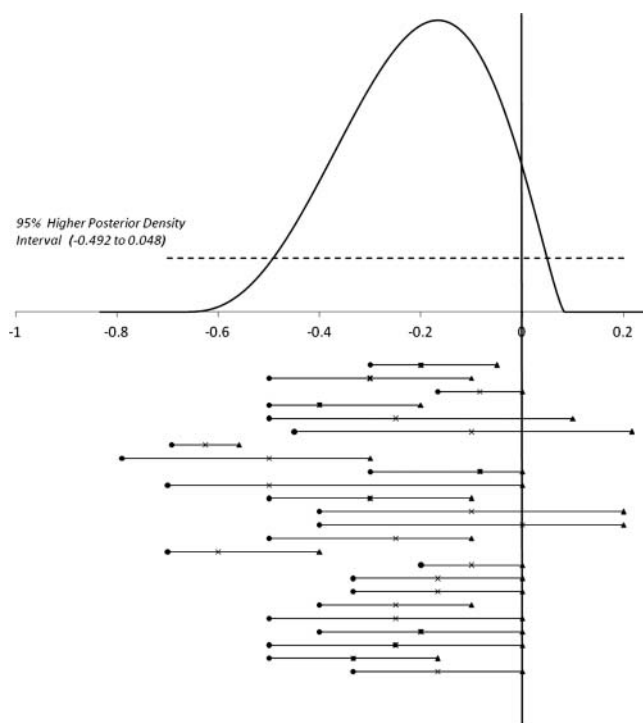


Figure 2 Results of the exercise to elicit experts' estimates on expected effectiveness. Above the x-axis: pooled expert opinion on the relative reduction of the attributable risk of adverse events as a result of the implementation of a proposed intervention to improve the quality of handover (mean -0.21 (21% reduction) and 95% Higher Posterior Density Interval (-0.492 to 0.048)). Below the x-axis: Best estimate (x), lowest (•), and upper (▲) plausible value of the relative reduction of the attributable risk of adverse events as a result of the implementation of a proposed intervention to improve the quality of handover, as elicited from each of the 24 experts.

also expressed considerable individual uncertainties. We made the assumption that the proportional risk reduction would be the same over various categories of adverse events.

Step 4: utility loss due to adverse events

The next step was to assign a utility function to the adverse events. First, each category of event was allocated to a ‘typical’ (indicative) state on a generic quality-of-life tool, the EQ5D. The result is shown in table 2. An additional health state, ‘minimal impairment,’ is not included, as generic quality-of-life instruments are not sufficiently sensitive to detect such a change.²⁸ We therefore assumed a disutility of 0.05 for this health state (in our warfarin example below, this might correspond to a patient whose international normalised ratio (INR) is out of range, requiring an additional visit to the physician).

The utility associated with each health state identified in this way was obtained from the EQ5D manual, and then multiplied by the mean duration of the state in years (or fractions of a year) to provide an estimate of the QALY loss associated with each adverse event. It was assumed that a person can have only one adverse event.

Step 5: unit costs of the intervention

In this study, we considered a HANOVER programme based on a facilitated classroom instruction supported by a suite of internet materials, as described in this issue and elsewhere.¹⁰ The unit cost (ie, cost averaged over discharges) for this resource-intensive form of the HANOVER intervention has been explicated elsewhere for five European countries.²⁹ The Netherlands study sample was selected as it was the country where the costs resulting from patient admissions due to adverse events were derived—see the next step. The unit cost for The Netherlands was €16.6.

Step 6: costs associated with each adverse event

We must also consider the possible savings contingent on reducing adverse events in each category. A recent European study by Hoonhout *et al*³⁰ reported the mean cost of admissions associated with preventable adverse events by disability categories in 2003 to be €2979 in cases of no disability; €5973 in cases of minor disability; €6649 in cases of permanent disability; and €3831 in cases of death. Note that the study of Hoonhout and colleagues is relevant to our model since, like our model, it is based on *preventable* (not overall) adverse events. However, they used four categories, as opposed to Brennan *et al* who subdivided the permanent category into two. Moreover, Hoonhout *et al* deals only with cases requiring hospital admission, whereas Forster *et al*

Table 2 'Typical' EQ5D states associated with adverse events in each severity class, as defined by Brennan *et al*²⁴

Health state	Example: Warfarin complication	Mobility	Self-care	Usual activities	Pain/ discomfort	Anxiety/ depression	Index	Utility	Disutility
Moderate Impairment, recovery 1– 6 months	Readmission for minor bleeding	1	1	2	1	1	11211	0.88	0.12
Moderate Impairment, recovery >6 months	Major Gastrointestinal bleeding	2	1	2	1	1	21211	0.81	0.19
Permanent Impairment	Stroke, leaving residual moderate hemiplegia	2	2	2	1	2	22212	0.64	0.36
Permanent— Severe	Stroke, leaving dense hemiplegia	2	2	3	1	2	22312	0.31	0.69

For each dimension 1 represents no handicap, 2 represents a moderate handicap, and #3 represents a severe handicap.

(see above), found that only one quarter of adverse events associated with handover resulted in patient readmission. The costs of attendance at the emergency department or doctor's office were provided by Tan *et al*³¹. We used our judgement to map all the above costs onto the adverse event categories identified above in step 1, as shown in [table 3](#). All costs have been expressed in euros at 2010/2011 prices. An annual inflation rate of 3.5% was applied to convert the unit costs of adverse event categories from 2003 to 2011 values, thus corresponding to the year when the intervention was costed.

Step 7: calculation of health benefit and cost savings

In [table 3](#), we calculate the average QALY loss and the healthcare-associated costs for each patient experiencing an adverse event. The QALY loss is calculated over each category of adverse event, weighted by its probability among people with an adverse event. Patients who die or become permanently disabled are assumed to have a mean life expectancy of 10 years (see above) given that the incidence of adverse events associated with permanent disability and deaths is higher among patients aged >65 years.³² The estimated QALY loss per person that

Table 3 Calculation of QALY loss and healthcare costs associated with each patient who experiences an adverse event

Health states	Proportion of patients with adverse events in each state (a)	Average duration in state (years) (b)	Disutility of state* (c)	Cost per adverse event (€) (d)	QALY loss (a×b×c)	Cost (€) (a×d)
Death	0.05	10	1	5045	0.5	252
Permanent >50%	0.02	10	0.69	8755	0.138	175
Permanent ≤50%	0.03	10	0.36	7865	0.108	236
Moderate impairment, recovery >6 months	0.10	1	0.19	3923	0.019	392
Moderate impairment, recovery >1 to 6 months	0.30	0.25	0.12	82†	0.009	25
Minimal impairment, recovery <1 month	0.50	0.08	0.05	0	0.002	0
Expected total per person with adverse event					0.775	1080

*From [table 2](#).

†Assuming 44% of patients incurred a visit to the emergency department (costed at €151 per visit) and in 56% of patients it involved a visit to their primary care physician (costed at €28 per visit). Based on data from Forster *et al*,¹³ and cost estimates from Tan *et al*.³¹ QALY, Quality Adjusted Life Year 2012.

Table 4 Calculation of cost-effectiveness measures (EMB and ICERs) for the intervention in a hospital with 50 000 discharges per year: the estimated total cost, QALYs, and the cost-effectiveness results

Total discharge	50000	Number of adverse events	QALY loss	Costs of adverse events (€)
Unit cost of the intervention per discharge (€)	16.6			
Cost of the intervention (€)	827900			
Willingness to pay (€) per QALY	20000			
Rate of adverse event	0.19	9500	*0.775×9500=7365.1	*1080×9500=10261714†
Attributable to handover errors	0.333	3164	0.775×3164=2452.6	1080×3164=3417151

*From table 3.

†This assumes that the costs of non-preventable adverse events is the same as that used in the calculations for preventable adverse events—there is some evidence that preventable adverse events are more costly.³³

EMB, monetary benefit; ICER, incremental cost-effectiveness ratio; QALY, Quality Adjusted Life Year.

experiences an adverse event is 0.775, and the estimated associated cost is €1080 per person.

The QALY loss across all discharges is 0.15 (0.775×0.19). The QALY loss for all patient discharges due to *preventable* adverse events is 0.049 (0.775×0.063). If the intervention effectiveness is 21% (base-case scenario), then the QALY gain per discharged patient is 0.010 (0.21×0.049)—equivalent to one healthy life-year per 100 patient discharges.

The costs associated with adverse events (step 6 above) were mapped onto the severity categories, as shown in table 3. Avoiding adverse events can defray costs apart from those associated with readmission, especially in the case of events resulting in permanent impairment and lifelong rehabilitation. Our cost estimates are liable to underestimate total savings for this reason. The healthcare costs associated with preventable adverse events is €68 (1080×0.063) when averaged over all types of patient discharges. This amounts to €14.3 (68×0.21) given the base-case assumption on the effectiveness of the intervention. Since the *unit* cost of the intervention is €16.6, the *net* cost of the HANDOVER intervention is €2.3 (16.6–14.3).

Step 8: calculation of cost-effectiveness

The estimated results were based on the assumption of a large European (ie, The Netherlands) hospital with 50 000 discharges. The calculations for costs and benefits are given in table 4. In a hospital with 50 000 discharges

each year, ineffective patient handovers result in a loss equivalent to over 7000 healthy life-years. It also results in nearly €3.5 million in health service costs that could be saved if all preventable adverse events were prevented. Under the base case, 21% of the above QALY loss can be averted (table 5). The expected cost of the intervention included the implementation cost of the HANDOVER project intervention and the cost released by the intervention. The unit cost per discharge for such a programme in The Netherlands is €16.6, and the cost for a hospital with 50 000 discharges is, therefore, €827 900 per year. This cost is subtracted from the cost savings from the reduced adverse events to give the overall net cost of the interventions. These results are given in table 5.

Last, the net costs and benefits are reconciled in table 6 to provide ICERs and ENBs. The proposed intervention under the base case (21% effectiveness) is highly cost-effective at only about €214 per QALY gain.

Stage 9: headroom and sensitivity analysis result

If the intervention were 100% effective in reducing preventable adverse events, it would ‘dominate’ (table 6)—it would be both effective and cost-saving, placing it in the southeast quadrant of the cost-effectiveness plane. Dominance is lost when effectiveness drops below 24.3% (table 5).

At the societal willingness to pay of €20 000 per QALY (a conservative assumption), the minimum effect the

Table 5 Calculation of cost-effectiveness measures (EMB and ICERs) for the intervention in a hospital with 50 000 discharges per year: Estimated costs per QALY at different levels of effectiveness

Intervention effectiveness	Proportion of discharges where AE avoided	Number of AEs avoided	QALY gained	Costs saved (€)	Net cost of intervention (€)*
100%	0.063	3164	2452.6	3417151	–2589251
21% (base case)	0.013	664	515	717602	110298
24.3%	0.015	769	596.0	830368	–2468
1.6%	0.001	51	39.2	54674	773226

*Costs of intervention, minus costs saved.

EMB, monetary benefit; ICER, incremental cost-effectiveness ratio; QALY, Quality Adjusted Life Year.

Table 6 Calculation of cost-effectiveness measures (EMB and ICERs) for the intervention in a hospital with 50 000 discharges per year: Estimated cost-effectiveness measured in ICER and EMB

Intervention effectiveness (%)	EMB	ICER	ENB*
100	49051649	-1055.72	51640900
21	10300846	214.15	10190548
24.3	11919551	-4.14	11922018
1.6	784826	19704.37	11601

*ENB=EMB-Cost.

EMB, monetary benefit; ENB, expected net benefit; ICER, incremental cost-effectiveness ratio; QALY, Quality Adjusted Life Year.

intervention must achieve to make it cost-effective is less than 1.6%. This reflects the low unit cost of the intervention (£16.6). Nevertheless, this represents a considerable outlay (£827 900) for a large (50 000 discharge) hospital. Under the base-case effectiveness estimate (21%), the intervention cost can increase to €220.36 before it becomes cost-ineffective at the given willingness-to-pay threshold.

We also subjected the life expectancy for people with permanent disability and who died, to one-way sensitivity analysis, assuming five- and 20-year time intervals. This resulted in ICERs of €412.67 and €109.14 per QALY, respectively.

DISCUSSION AND RECOMMENDATION

This paper describes the economic evaluation methodology for generic service interventions. It goes beyond most economic studies by valuing the non-monetary benefits of a complex, social intervention. It is novel in that it deals with the methodology for such studies, and focuses on the formative (preimplementation) stage of a new service. At this stage, it is necessary to estimate the effectiveness of the intervention on the basis of expert opinion. We considered it desirable to describe baseline rates for selected endpoints, and to identify the preventable (attributable) component (that is, the proportion of endpoints that result from failure of the process at which the intervention is targeted). We believe this is necessary to inject realism into the process and attenuate overoptimistic estimates by focussing respondents' minds on the component of the endpoint that is amenable to service change.

For the HANDOVER intervention, the magnitude of the expected effects of the proposed intervention is modest (a 1.7 percentage point (ie, absolute) reduction in all adverse events). However, the model demonstrates that such an effect would, nevertheless, be highly cost-effective. The analysis shows that even small benefits (less than a 2% relative risk reduction or 0.3% absolute

risk reduction in adverse events of all types) can be cost-effective for an intervention that will cost a hospital over €800 000. The small effect sizes that may be cost-effective for generic service delivery interventions have been pointed out before.⁴ However, this previous study did not take into account the health service savings that are achieved when adverse events are avoided. The result is that health service interventions will frequently be cost-effective at low magnitudes of effect, and may dominate. Although the intervention was cost-effective at an effect size of less than 2%, it had to be nearly 25% effective (ie, 1.7% absolute risk reduction) to be cost-releasing.

An intervention that might sound very expensive at the institutional level is modest when costs were hypothecated for individual patients. Under the base case, the HANDOVER intervention would be cost-effective up to approximately €220 per patient—modest when compared with clinical interventions, but a ruinous €11 million when applied to a hospital with 50 000 discharges per year. These costs may be even harder to bear when savings do not accrue to the institution that bears the cost.

A corollary of these findings is that studies to detect minimum magnitude of effects that justify their cost need to be very large to avoid type 2 error—that is, a false null result. Studies demonstrating that adverse events can be reduced from a baseline of 19% to a rate of 17.7% (ie, the 1.3% reduction posited in the base case) would require 14 060 observations, both before and after the intervention at an alpha of 0.05 and a beta of 0.8. A less biased study with contemporaneous controls would require a much larger sample size,³⁴ and would be difficult to organise given the clustered nature of the data. Likewise, detecting the less than 0.3% absolute improvement level at which the intervention becomes cost-effective would be a totally quixotic undertaking. Even after full implementation, the cost-effectiveness of inexpensive interventions may have to be modelled, rather than estimated directly from observations of improved outcome for patients. The model developed at the outset can be repopulated as development proceeds and information of various sorts gradually accumulates.³ This evidence may be surrogate evidence for patient outcomes. Such evidence is collated at the system level, upstream of the patient, as described elsewhere.⁴ While evidence on effectiveness may be very difficult or impossible to obtain, other potential inputs to economic models, such as prevalence rates of adverse events and costs, can be obtained at lower cost and/or with less difficulty.

Limitations of this study

This study is an ex ante evaluation and, therefore, by definition, encompasses many unavoidable uncertainties. This

is the first study we know of to attempt an ex ante health economic assessment of a proposed service delivery intervention. It builds on methods we have developed for the device industry.^{3 35 36} The main problem we have encountered is the large number of potential adverse events, and many of the limitations stem from this conundrum. Since there is little literature to guide us in this area, we had to break new ground, and we do not claim to have provided the definitive solution. The approach we have taken is to group adverse events by severity and duration. This approach is associated with many challenges:

1. Baseline rates of preventable adverse events have to be apportioned across the categories, and this had to be induced from indirect evidence as described.
2. Ideally, separate Bayesian densities for effectiveness estimates should be elicited for each group—we assumed a common effect across groups. Severe and minor errors are not well correlated,³⁷ and adverse event categories may be affected differently by a given intervention.²³
3. Different utilities are needed for each group, and these do not appear to be available in the literature—for example, they are not available in the Tufts database.³⁸ We imputed EQ5D scores in an ad hoc way and translated these into utilities by applying the standard conversion factor (tariff). Clearly, it would have been preferable to obtain these estimates from a more representative sample of respondents. If the idea of categorising endpoints catches on and a classification system can be agreed upon, then a formal study of utilities associated with different types of adverse events (and according to how they arise) would be needed.³⁹
4. Cost savings are very different for each category. Hoonhout *et al*'s study hypothecated costs on categories of adverse events, but the categories he used did not map precisely onto the categories we used (based on the Harvard Medical Practice Study). Furthermore, the costings in Hoonhout *et al*'s study omitted costs of long-term care, loss of earnings, and litigation.
5. The average life expectancy of people who survive with permanent disability was difficult to ascertain, and there will always be a problem in estimating life-years lost in people who die following adverse events.

At the very least, our paper demonstrates the need to develop a consensus on the intellectual framework for assessing the outcomes of complex service and safety interventions, and then to set up epidemiological studies to collect data in the requisite form.

CONCLUSION

We have argued for the utility of a predevelopment economic evaluation of service interventions, and described a method for achieving this. We applied the methodology

to a clinical handover intervention aimed at improving transfer of patients from hospital to the community. Adverse events constitute the outcome of interest, but our major challenge was to deal with the wide range of such events. We think that the method we have proposed shows promise, but it is in its infancy and much more work is necessary to develop supply-side health economic modelling for service delivery interventions. However, refinements in the method will not overturn one stark conclusion—service delivery interventions are cost-effective at small absolute effect sizes and at costs that cannot be borne under current thresholds used to ration new treatments.

Author affiliations

¹Department of Public Health, Epidemiology and Biostatistics, University of Birmingham, Birmingham, UK

²Centre for Learning Sciences and Technologies, Open Universiteit Nederland, Heerlen, The Netherlands

³Utrecht Medical Center, Utrecht, The Netherlands

⁴University of Stavanger, Stavanger, Norway

⁵University College Cork, Cork, Ireland

*The European HANDOVER Research Collaborative consists of: Venneri F, Molisso A (Azienda Sanitaria Firenze, Italy), Albolino S, Toccafondi G (Clinical Risk Management and Patient Safety Center, Tuscany Region, Italy), Barach P, Gademan P, Göbel B, Johnson J, Kalkman C, Pijnenborg L (Patient Safety Center, University Medical Center Utrecht, Utrecht, The Netherlands), Wollersheim H, Hesselink G, Schoonhoven L, Vernooij-Dassen M, Zegers M (Scientific Institute for Quality of Healthcare, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands), Boshuizen E, Drachler H, Kicken W, van der Klink M, Stoyanov S (Centre for Learning Sciences and Technologies, Open University, Heerlen, The Netherlands), Kutryba B, Dudzik-Urbaniak E, Kalinowski M, Kutaj-Wasikowska H (National Center for Quality Assessment in Health Care, Krakow, Poland), Suñol R, Groene O, Orrego C (Avedis Donabedian Institute, Universidad Autónoma de Barcelona, Barcelona, Spain), Öhlén G, Airoso F, Bergenbrant S, Flink M, Hansagi H, Olsson M (Karolinska University Hospital, Stockholm, Sweden), Lilford R, Chen Y-F, Novielli N, Manaseki-Holland S (University of Birmingham, Birmingham, United Kingdom).

Acknowledgements We thank all the members of the European HANDOVER Research Collaborative for their participation in this study. Paul Barach, Loes Pijnenborg, Julie Johnson, Beryl Göbel, Cor Kalkman, Basia Kutryba, Julie Johnson, Halina Kutaj-Wasikowska, Ewa Dudzik-Urbaniak, Marcin Kalinowski, Francesco Venneri, Giulio Toccafondi, Antonio Molisso, Sara Albolino, Hub Wollersheim, Gijs Hesselink, Lisette Schoonhoven, Myrra Vernooij, Marieke Zegers, Helen Hansagi, Mariann Olsson, Susanne Bergenbrant, Maria Flink, Gunnar Ohlen, Carola Orrego, Rosa Sunol, Oliver Groene. We would also like to thank Melita Shirley (University of Birmingham) for aiding in the preparation of this manuscript.

Contributors GLY refined and executed the economic model; NN collected and analysed Bayesian priors and developed first iteration of economic model; SMH provided editorial assistance; YFC conducted the systematic review; MvdK developed the intervention; PB provided editorial assistance and was the principal investigator of the HANDOVER intervention; PC provided editorial assistance and verified calculations; RJL provided the original concept of the study, was primary author of the manuscript, and oversaw the evaluation/economic work package of HANDOVER.

Funding This study was supported by a grant from the European Union, the Framework Programme of the European Commission (FP7-HEALTH-F2-2008-223409); by a National Institute of Health Research (NIHR) Programme Grant (RP-PG-1209-10099), Investigating the implementation, adoption and effectiveness of ePrescribing systems in English hospitals: a mixed-methods national evaluation; The EPSRC MATCH

project and by the NIHR Collaborations for Leadership in Applied Health Research and Care (CLAHRC) for Birmingham and Black Country.

Competing interests All authors have completed the Unified Competing Interest form at http://www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author), and all authors want to declare (1) Financial support for the submitted work was supported by a grant from the European Union, the Framework Programme of the European Commission (FP7-HEALTH-F2-2008-223409); by a National Institute of Health Research (NIHR) Programme Grant (RP-PG-1209-10099), Investigating the implementation, adoption and effectiveness of ePrescribing systems in English hospitals: a mixed-methods national evaluation; the West Midlands Quality Institute; and by the NIHR Collaborations for Leadership in Applied Health Research and Care (CLAHRC) for Birmingham and Black Country; (2) No financial relationships with commercial entities that might have an interest in the submitted work; (3) No spouses, partners or children with relationships with commercial entities that might have an interest in the submitted work; (4) No non-financial interests that may be relevant to the submitted work.

The study sponsors had no role in the study design; collection, analysis and interpretation of the data; or in the writing of the article and decision to submit the article for publication.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

- Brown C, Hofer T, Johal A, *et al*. An epistemology of patient safety research: a framework for study design and interpretation. Part 1. Conceptualising and developing interventions. *Qual Saf Health Care* 2008;17:158–62.
- Medical Research Council (MRC). *Developing and evaluating complex intervention: new guidance*. London, UK: MRC, 2008.
- Vallejo-Torres L, Steuten LM, Buxton MJ, *et al*. Integrating health economics modeling in the product development cycle of medical devices: a Bayesian approach. *Int J Technol Assess Health Care* 2008;24:459–64.
- Lilford RJ, Chilton PJ, Hemming K, *et al*. Evaluating policy and service interventions: framework to guide selection and interpretation of study end points. *BMJ* 2010;341:c4413.
- Mason J, Freemantle N, Nazareth I, *et al*. When is it cost-effective to change the behavior of health professionals? *JAMA* 2001;286:2988–92.
- Ovretveit J. Does improving quality save money?—Health Foundation. <http://www.health.org.uk/publications/does-improving-quality-save-money/> (accessed 10 Jul 2012).
- Mohr J, Batalden P, Barach P. Integrating patient safety into the clinical microsystem. *Qual Saf Health Care* 2004;13:34–8.
- Philibert I. Use of strategies from high-reliability organisations to the patient hand-off by resident physicians: practical implications. *Qual Saf Health Care* 2009;18:261–6.
- Apostolakis G, Barach P. Lessons learned from nuclear power. In: Hatlie M, Tavill K, eds. *Patient safety: international textbook*. Faithersburg, MD: Aspen Publications, 2003:205–25.
- van der Klink M, Kicken W, Drachsler H, *et al*. Deliverable 9—Evaluation report regarding training and use of tools. 2011.
- Brown C, Lilford R. Evaluating service delivery interventions to enhance patient safety. *BMJ* 2008;337:a2764.
- Australian Council for Safety and Quality in Health Care. Clinical Handover and Patient Safety—Literature review report. 2005. [http://www.health.gov.au/internet/safety/publishing.nsf/Content/AA1369AD4AC5FC2ACA2571BF0081CD95/\\$File/clinhovrlitrev.pdf](http://www.health.gov.au/internet/safety/publishing.nsf/Content/AA1369AD4AC5FC2ACA2571BF0081CD95/$File/clinhovrlitrev.pdf) (accessed 21 May 2012).
- Forster AJ, Murff HJ, Peterson JF, *et al*. The incidence and severity of adverse events affecting patients after discharge from the hospital. *Ann Intern Med* 2003;138:161–7.
- Mistiaen P, Francke AL, Poot E. Interventions aimed at reducing problems in adult patients discharged from hospital to home: a systematic meta-review. *BMC Health Serv Res* 2007;7:47.
- Shepperd S, McClaran J, Phillips CO, *et al*. Discharge planning from hospital to home. *Cochrane Database Syst Rev* 2010;(1). Art. No.: CD00313.
- Kirkwood BR, Sterne JA. *Essential medical statistics*. 2nd edn. Malden, MA: Blackwell Science Ltd, 2003:449–51.
- Lilford R, Pronovost P. Using hospital mortality rates to judge hospital performance: a bad idea that just won't go away. *BMJ* 2010;340:c2016.
- Khalil EL. The Bayesian fallacy: distinguishing internal motivations and religious beliefs from other beliefs. *J Econ Behav Organ* 2010;75:268–80.
- O'Hagan A, Buck CE, Daneshkhan A, *et al*. *Uncertain judgements: eliciting experts' probabilities*. London, UK: Wiley, 2006.
- Ashby D, Smith AF. Evidence-based medicine as Bayesian decision-making. *Stat Med* 2000;19:3291–305.
- Lilford R, Braunholtz DA. The statistical basis of public policy: a paradigm shift is overdue. *BMJ* 1996;313:603–60.
- Spiegelhalter DJ, Freedman LS, Parmar MK. Applying Bayesian ideas in drug development and clinical trials. *Stat Med* 1993;12:1501–11.
- Bowater R, Lilford R. Clinical effectiveness in cardiovascular trials in relation to the importance to the patient of the end points measured. *J Eval Clin Pract* 2011;17:547–53.
- Brennan TA, Leape LL, Laird NM, *et al*. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. 1991. *NEJM* 1991;324:370–6.
- Halfon P, Eggli Y, van Melle G, *et al*. Measuring potentially avoidable hospital readmissions. *J Clin Epidemiol* 2002;55:573–87.
- Novielli N, Chen Y-F, van der Klink M, *et al*. Deliverable nr D7—report quantifying the expected benefits of the planned interventions. 2010. <http://www.handover.eu/upload/library/z740hd0knqw1akbauau25.pdf> (accessed 19 Apr 2012).
- Johnson SR, Tomlinson GA, Hawker GA, *et al*. Methods to elicit beliefs for Bayesian priors: a systematic review. *J Clin Epidemiol* 2010;63:355–69.
- Whitehead SJ, Shehzad A. Health outcomes in economic evaluation: the QALY and utilities. *Br Med Bull* 2010;96:5–21.
- Novielli N, van der Klink M, Manaseki-Holland S, *et al*. Deliverable nr D11—report quantifying the resources actually consumed in each country by type of intervention. 2011. <http://www.handover.eu/upload/library/zbsrzp1g4d1f12b9gu7h8.pdf> (accessed 19 Apr 2012).
- Hoonhout LH, de Bruijne MC, Wagner C, *et al*. Direct medical costs of adverse events in Dutch hospitals. *BMC Health Serv Res* 2009;9:27.
- Tan SS, Bouwmans CA, Rutten FF, *et al*. Update of the Dutch manual for costing in economic evaluations. *Int J Technol Assess Health Care* 2012;28:152–8.
- Zegers M, De Bruijne MC, Wagner C. Adverse events and potentially preventable deaths in Dutch hospitals: results of a retrospective patient record review study. *Qual Saf Health Care* 2009;18:297–302.
- Rodríguez-Monguió R, Otero MJ, Rovira J. Assessing the economic impact of adverse drug effects. *Pharmacoeconomics* 2003;21:623–50.
- Brown C, Hofer T, Johal A, *et al*. An epistemology of patient safety research: a framework for study design and interpretation. Part 2. Study design. *Qual Saf Health Care* 2008;17:163–9.
- Girling A, Young T, Brown C, *et al*. Early-stage valuation of medical devices: the role of developmental uncertainty. *Value Health* 2010;13:585–91.
- Girling AJ, Lilford RJ, Young TP. Pricing of medical devices under coverage uncertainty—a modelling approach. *Health Econ*. Published Online First: 20 October 2011. doi: 10.1002/hec.1807.
- Coleman JJ, Hemming K, Nightingale PG, *et al*. Can an electronic prescribing system detect doctors who are more likely to make a serious prescribing error? *J R Soc Med* 2011;104:208–18.
- Tufts Medical Center. CEA Registry Website. 2012. <https://research.tufts-nemc.org/cea4/Home.aspx> (accessed 18 May 2012).
- Steuten L, Buxton M. Economic evaluation of healthcare safety: which attributes of safety do healthcare professionals consider most important in resource allocation decisions? *Qual Saf Health Care* 2010;19:e6.