

# Methodological Issues Surrounding the Use of Baseline Health-Related Quality of Life Data to Inform Trial-Based Economic Evaluations of Interventions Within Emergency and Critical Care Settings: A Systematic Literature Review

Melina Dritsaki<sup>1</sup> · Felix Achana<sup>2</sup> · James Mason<sup>2</sup> · Stavros Petrou<sup>2</sup>

Published online: 6 January 2017  
© Springer International Publishing Switzerland 2017

## Abstract

**Background** Trial-based cost-utility analyses require health-related quality of life data that generate utility values in order to express health outcomes in terms of quality-adjusted life years (QALYs). Assessments of baseline health-related quality of life are problematic where trial participants are incapacitated or critically ill at the time of randomisation. This review aims to identify and critique methods for handling non-availability of baseline health-related quality of life data in trial-based cost-utility analyses within emergency and critical illness settings.

**Methods** A systematic literature review was conducted, following PRISMA guidelines, to identify trial-based cost-utility analyses of interventions within emergency and critical care settings. Databases searched included the National Institute for Health Research (NIHR) Journals Library (1991–July 2016), Cochrane Library (all years); National Health Service (NHS) Economic Evaluation Database (all years) and Ovid MEDLINE/Embase (without time restriction). Strategies employed to handle non-availability of baseline health-related quality of life data in final QALY estimations were identified and critiqued.

**Results** A total of 4224 published reports were screened, 19 of which met the study inclusion criteria (mean trial size 1670): 14 (74 %) from the UK, four (21%) from other European countries and one (5%) from India. Twelve studies (63%) were based in emergency departments and seven (37%) in intensive care units. Only one study was able to elicit patient-reported health-related quality of life at baseline. To overcome the lack of baseline data when estimating QALYs, eight studies (42%) assigned a fixed utility weight corresponding to either death, an unconscious health state or a country-specific norm to patients at baseline, four (21%) ignored baseline utilities, three (16%) applied values from another study, one (5%) generated utility values via retrospective recall and one (5%) elicited utilities from experts. A preliminary exploration of these methods shows that incremental QALY estimation is unlikely to be biased if balanced trial allocation is achieved and subsequent collection of health-related quality of life data occurs at the earliest possible opportunity following commencement of treatment, followed by an adequate number of follow-up assessments.

**Conclusion** Trial-based cost-utility analyses within emergency and critical illness settings have applied different methods for QALY estimation, employing disparate assumptions about the health-related quality of life of patients at baseline. Where baseline measurement is not practical, measurement at the earliest opportunity following commencement of treatment should minimise bias in QALY estimation.

**Electronic supplementary material** The online version of this article (doi:10.1007/s40273-016-0485-x) contains supplementary material, which is available to authorized users.

✉ Melina Dritsaki  
melina.dritsaki@ndorms.ox.ac.uk

<sup>1</sup> Oxford Clinical Trials Research Unit, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford OX3 7LD, UK

<sup>2</sup> Clinical Trials Unit, Warwick Medical School, University of Warwick, Coventry CV8 7AL, UK

### Key Points for Decision Makers

Employ an appropriate randomisation strategy to ensure baseline comparability across treatment groups.

Conduct the initial health-related quality of life assessment at the earliest time possible post randomisation.

Include a constant or imputed baseline value rather than ignoring it.

## 1 Introduction

Economic evaluations are increasingly being conducted alongside phase III and phase IV randomised controlled trials of various interventions such as surgical procedures, drug treatments, diagnostic tests and behavioural interventions [1]. In the UK, government agencies such as the National Institute for Health and Care Excellence (NICE) for England and Wales, the All Wales Medicines Strategy Group (AWMSG) for Wales, and the Scottish Medicines Consortium (SMC) for Scotland have established decision-making processes that draw heavily upon economic evidence collected within the context of randomised controlled trials, whilst research funding bodies such as the National Institute for Health Research (NIHR) routinely request the inclusion of economic assessment methods within large-scale clinical trials [1, 2]. Similarly, economic evidence collected within the context of randomised trials is increasingly being used to inform the regulatory and reimbursement decisions of government agencies in other nations [3, 4]. Depending on the research question, trial-based economic evaluations can take the form of cost-consequence analyses, cost-effectiveness analyses or cost-utility analyses. Cost-utility analyses are particularly appealing to decision makers as they permit cost-effectiveness comparisons to be made using the quality-adjusted life year (QALY) metric for different health care interventions across disparate health conditions. It is unsurprising therefore that cost-utility analysis using the QALY outcome measure (which combines length of life and health-related quality of life in a single measure of health consequence) remains the preferred evaluative method in the technology appraisal guidance of many government agencies [1, 2].

To generate the QALYs needed to inform trial-based cost-utility analyses, data on survival and data on health-related quality of life measured at baseline and subsequent follow-up time points are required for trial participants.

Randomised controlled trials conducted in emergency and critical care settings have used various multi-attribute utility instruments, including the EQ-5D [5], SF-6D [6] and Health Utilities Index-3 (HUI-3) [7], to reflect preferences for patient health states; these are normally converted into health utility values using established algorithms [8–10]. In critical care settings, the psychometric properties of the EQ-5D and the SF-12 (from which the SF-6D can be derived) were recently examined in two patient populations, namely patients diagnosed with acute respiratory distress syndrome [11] and survivors of out of hospital cardiac arrest [12]. The authors of both studies reported satisfactory performance of these instruments in the respective patient populations. However, asking patients to complete these measures around the time of recruitment (commonly taken as the baseline measurement) into randomised controlled trials conducted within emergency and critical care settings can be problematic; patients are commonly incapacitated and unable to provide a self-assessment of their health status at or around the time of randomisation [13, 14]. Problems also arise because the event of interest is often acute in nature rather than pre-planned; the unknown timing makes it difficult to collect baseline data from participants during the occurrence of the event and at the point of recruitment into the trial. Alternative strategies used to collect patient-reported outcome data in clinical trials more broadly, such as researcher-administered interviews (face-to-face or, in the context of longer term follow-up, by telephone), can also be problematic to implement when patients are critically ill for the same reasons. Even where trial participants are conscious, the nature of some health conditions (for example, cardiac arrest and serious traumatic injury) around the time of randomisation can often raise ethical objections to collecting patient-reported outcome data. Solutions adopted by previous trial-based economic evaluations within emergency and critical illness settings have included delaying the time at which health-related quality of life is assessed until patients are well enough to complete questionnaires [15–17], asking patients to retrospectively recall [18] their pre-randomisation health state and use of proxies such as patients' next of kin or health professionals [19]. The impact of this heterogeneity of method upon the findings of economic analyses is unclear.

An early systematic review of the critical care literature conducted in 2002 provided evidence of the difficulty of conducting within-trial cost-utility analyses involving critically ill or injured patient populations [20]. Of the 29 economic analyses identified in that review, none was a cost-utility analysis. This systematic review therefore aimed to identify and critique approaches to collection of health-related quality of life data and subsequent estimation of QALYs in the absence of directly and

contemporaneously measured baseline values in trial-based cost-utility analyses of interventions within emergency and critical illness settings. To our knowledge this problem has not been addressed before, and there is scope to develop recommendations for future best practice to inform health economics researchers in emergency or critical care. The paper is structured as follows: Sect. 2 outlines the systematic review methods, and is followed by presentation of the results in Sect. 3. A critical appraisal of the methods identified in the review for dealing with a lack of baseline health-related quality of life data when estimating QALYs is presented in Sect. 4. The aim is to understand the implications of the assumptions underlying each method and the likely impact on cost-effectiveness results. The discussion and conclusions are presented in Sect. 5.

## 2 Methods

The NIHR Journals Library (1991–July 2016), Cochrane Library (all years); National Health Service (NHS) Economic Evaluation Database (all years) and Ovid MEDLINE/Embase (without time restriction) were searched. Search item groupings included terms and derivatives for “intensive care”, “economic evaluation” and “randomised controlled trial”. Full details of the search strategy are provided in Appendix A in the supplementary electronic material (online resource 1).

Included studies were cost-utility analyses of interventions based in emergency or critical care settings, for example, accident and emergency departments or intensive care units, which were conducted alongside randomised controlled trials. The randomisation (allocation to trial arm) had to be made whilst patients were in an emergency or critical care setting and consequently were incapacitated/unable to provide self-assessment of their health status. Eligible studies had to have collected preference-based health-related quality of life data from trial participants themselves, by proxy (e.g. relatives, health care professionals) or from an external source (for example, another study or expert opinion) to support the subsequent economic analysis. Studies were excluded if they did not include a cost-utility analysis or if the condition was non-acute, e.g. management of influenza, where patients could normally give written consent. In addition, because of the way disability-adjusted life years (DALYs) are calculated (i.e. the disutility weights are calculated for specific disease conditions and not based on patient preferences [21]), within-trial cost-utility analyses that reported outcomes in terms of DALYs were excluded. Mental health care-related studies were excluded because of the particular methodological challenges presented. Non-English studies and the grey literature were also excluded.

Literature searches and reviews were performed in two stages in accordance with Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [22] and included studies published up to July 2016. First, titles and abstracts were screened to identify and retrieve potentially relevant reports. Second, retrieved full reports were assessed for eligibility. Both stages were completed independently by two reviewers (MD, FA) checking against pre-specified inclusion and exclusion criteria, with disagreements resolved through consensus. For eligible studies, data were extracted on the clinical setting, clinical condition, study perspective, time horizon, sample size, participant demographics, preference-based health-related quality of life instrument(s), timing of data collection, source of (and where applicable, accompanying assumptions around) baseline health-related quality of life data and methods used to estimate QALYs. The quality of included studies was assessed using the strategy reported by Kendrick et al. [23] and included the quality of the randomisation process, blinding of outcome assessment and completeness of follow-up (see Appendix B in the supplementary electronic material for further details). The review was registered on the PROSPERO register of systematic reviews (registration number CRD42016046174).

Finally, the conduct and reporting of each trial-based economic evaluation was assessed against selected items on the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) checklist [24] for reporting single study-based economic evaluations of interventions and expanded to include the following: (1) study methods, including description of target population, clinical settings, perspective of the analysis, study time horizon and whether or not cost and effects were discounted and if so by what amount; (2) method of data collection (including the type of preference-based health-related quality of life instrument used and the follow-up time points at which data collection was conducted); (3) method used to calculate QALYs, including, where applicable, how the non-availability of baseline health-related quality of life data was handled when estimating QALYs; (4) characterisation of uncertainty; and (5) a critical and thematic appraisal of the reported methods used to handle non-availability of baseline health-related quality of life data in subsequent QALY estimation. Characterisation of uncertainty was assessed by examining whether studies reported parameter estimates together with associated measures of uncertainty (e.g. standard errors, confidence intervals, etc.) and investigated the impact of known methodological assumptions on final estimates of cost-effectiveness. Simple descriptive statistics were used to summarise characteristics of included studies and the methods used to handle non-availability of baseline health-related quality of life data in the cost-utility analyses. Results are presented narratively in textual

format, providing numbers and corresponding percentages in brackets where appropriate.

### 3 Results

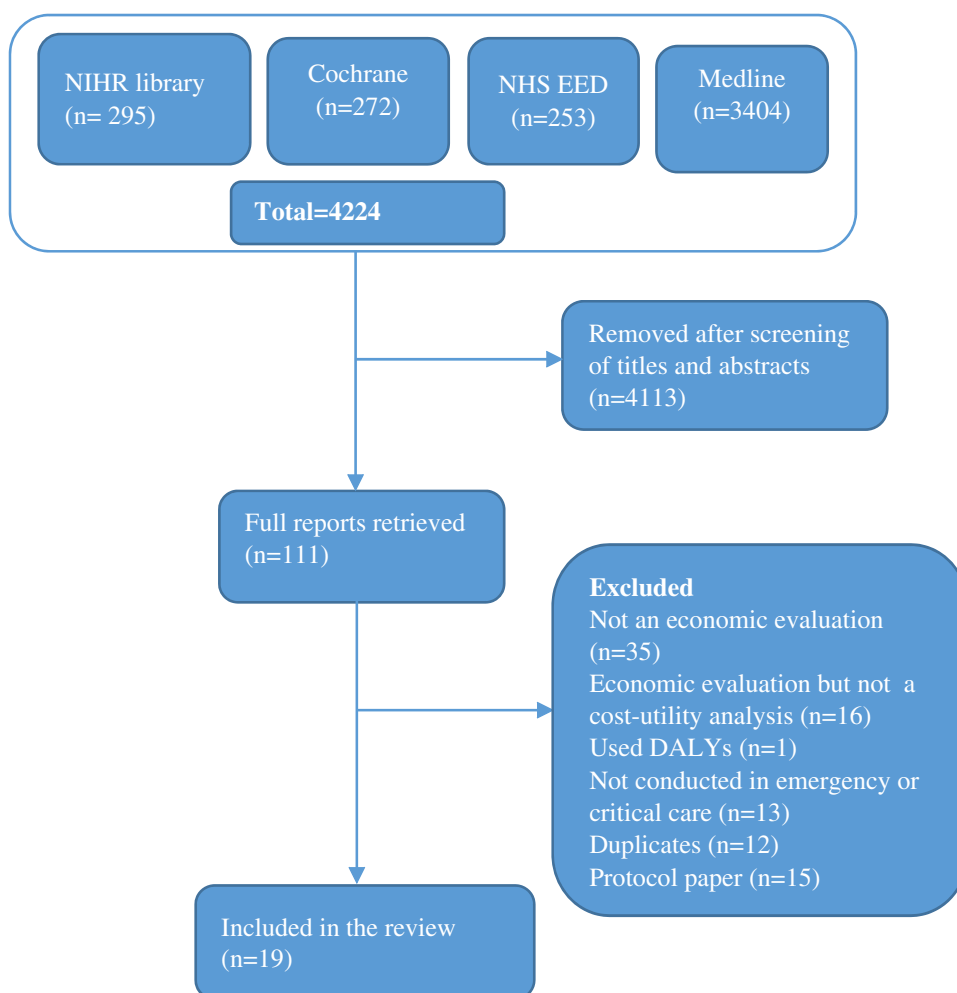
#### 3.1 Summary of Randomised Controlled Trials Included in the Review

A total of 4224 published reports (e.g. published papers, book chapters, monographs, etc.) were screened (Fig. 1), of which 4113 were excluded after initial review of titles and abstracts. Of the remaining 111 full reports retrieved, 92 (83%) were excluded: 35 reported no economic evaluation outcomes at all (i.e. did not report cost-effectiveness, cost-consequence or cost-utility outcomes), 16 reported economic evaluation outcomes that were not cost-utility based (i.e. were either cost-consequence or cost-effectiveness analyses in which the measures of health outcomes were not synthesised into preference-based metrics), 13 were conducted in non-emergency or

critical care settings, 12 were duplicate reports, 15 were protocol papers and one study expressed the health outcomes of the economic evaluation in DALY terms. A list of all 111 reports that reached the second stage of the review process is provided in Appendix C in the supplementary electronic material.

Table 1 summarises the baseline characteristics of the 19 trial-based cost-utility analyses included in the review, published between 2004 and 2016. The majority of studies, 14 (74%), were conducted in the UK, four (21%) in other European countries (namely Denmark, Norway, Germany and the Netherlands/Switzerland) and one (5%) in India. The mean number of patients in the underpinning randomised controlled trials was 1760 (range 180–6182), mean age was 53 years (range of means 0.51–78) and mean percentage of males was 57% (range 30–76%). In terms of clinical setting, 12 studies (63%) were based in emergency departments and included conditions such as emergency resuscitation for out-of-hospital cardiac arrest or acute asthma in adults and children, and seven (37%) were in intensive care units.

**Fig. 1** Flow chart of study identification and selection. *DALYs* disability-adjusted life years, *NHS EED* National Health Service Economic Evaluation Database, *NIHR* National Institute for Health Research



**Table 1** Summary of trial-based cost-utility analyses included in the review

First author (year), acronym	Country	Clinical setting	Disease/condition	Perspective	Time horizon (months)	Discounting costs and effects <sup>d</sup>	Sample size	Mean age (years)	% male
Goodacre (2004) [41]	UK	ED	Acute, undifferentiated chest pain	NHS	6	No	972	50	64
Harvey (2006) [29]	UK	ICU	Critically ill patients in adult general intensive care	NHS/PSS	12 <sup>a</sup>	Yes, 3.5%	1041	65	58
Dixon (2009) [43]	UK	ED	Care of older people following a call for an emergency ambulance	NHS/PSS	12	No	3018	78	43
Peek (2010), CESAR [33]	UK	ICU	Severe adult respiratory failure	NHS, Societal	6 <sup>b</sup>	Yes, 3.5%	180	40	58
Bohmer (2011) [18]	Norway	ED	ST-elevation myocardial infarction	Societal	12	No	266	61	76
Goodacre (2011), 3CPO [17]	UK	ICU	Acute cardiogenic pulmonary oedema	NHS/PSS	6 <sup>a</sup>	Yes, 3.5%	1069	78	43
Goodacre (2011), RATRAC [16]	UK	ED	Acute myocardial infarction	NHS	3	No	2243	55	58
Lamb (2012), MINT [42]	UK	ED	Acute whiplash injury	NHS	12	No	3851	37	43
Gates (2013), BALTI-2 [13]	UK	ICU	Acute respiratory distress syndrome	NHS/PSS	12	No	326	55	65
Powell (2013), MAGNETIC [19]	UK	ED	Acute severe asthma in children	NHS	1	No	508	4	58
Goodacre (2014), 3Mg [40]	UK	ED	Acute asthma in adults	NHS/PSS	1	No	1084	36	30
Macrae (2014), CHIP [28]	UK	ICU	Blood glucose control in paediatric intensive care units	NHS/PSS	Life-time	Yes, 3.5%	1369	0.51 <sup>c</sup>	55
Gyrd-Hansen (2015), PHANTOM-S [25]	Denmark	ED	Acute stroke	Third-party	3 <sup>b</sup>	Yes, 3.0%	6182	74	44
Lall (2015), OSCAR [26]	UK	ICU	Acute respiratory distress syndrome	NHS/PSS	12	No	795	55	62
Mouncey (2015) [15]	UK	ED	Early septic shock	NHS/PSS	12 <sup>a</sup>	No	1260	65	56
Perkins (2015), PARAMEDIC [35]	UK	ED	Out of hospital cardiac arrest	NHS/PSS	3	Unknown	4471	71	63
Rosenthal (2015) [27]	India	ICU	Nosocomial infections	Health care payer	Not stated	Not reported	1096	60	67
Schuster (2015), IABP-SHOCK II [31]	Germany	ED	Acute myocardial infarction	Health care (Germany)	12	No	600	70 <sup>c</sup>	69
Sierink (2016) REACT-2 [32]	Netherlands, Switzerland	ED	Trauma patients requiring advanced life support	Health care	12	Unknown	1403	44	76

3CPO noninvasive ventilation in patients with acute severe cardiogenic pulmonary oedema, 3Mg intravenous or nebulised magnesium sulphate versus placebo in adults with acute severe asthma, ARDS acute respiratory distress syndrome, BALTI-2 Beta-Agonist Lung injury Trial-2, CESAR Conventional ventilatory support versus extracorporeal membrane oxygenation for severe adult respiratory failure, CHIP Control of Hyperglycaemia in Paediatric Intensive Care, ED emergency department, IABP-SHOCK II Intra-aortic Balloon Pump in Cardiogenic Shock II, ICU intensive care unit, MAGNETIC MAGNESium Trial in Children, MINT Managing Injuries of the Neck Trial, NHS National Health Service, OSCAR High-frequency oscillatory ventilation against conventional artificial ventilation for adults with acute respiratory distress syndrome, PARAMEDIC Mechanic versus manual chest compression for out-of-hospital cardiac arrest, PHANTOM-S Prehospital thrombolysis, PSS Personal Social Services, RATPAC Randomised Assessment of Treatment using Panel Assay of Cardiac markers, REACT-2 Immediate total-body CT scanning versus conventional imaging and selective CT scanning in patients with severe trauma

<sup>a</sup> Extrapolated to lifetime

<sup>b</sup> Extrapolated to 60 months

<sup>c</sup> Median age reported in original trial paper

<sup>d</sup> Discount rate per annum



One study [27] did not report a time horizon for the economic evaluation. The mean time horizon for the within-trial component of the economic evaluations in the remaining 18 studies was 8 months (median 9 months and range 1–12 months). Some studies extrapolated outcomes beyond the trial follow-up period using decision analytic modelling methods [25, 26, 28], with up to 60 months [25] and lifetime [15, 16, 28, 29] extrapolations beyond the study follow-up periods. Most economic evaluations were conducted from the perspective of the UK NHS ( $n = 4$ ) or the NHS/Personal Social Services ( $n = 9$ ) in accordance with NICE guidance for appraising health technologies [30]. One further UK study adopted a societal as well as an NHS perspective [30]. Of the five non-UK studies [18, 25, 27, 31, 32] included in the review, four studies [18, 27, 31, 32] adopted a health services perspective, whilst one study adopted a third-payer perspective described as excluding costs to sectors other than the health sector and out-of-pocket expenses [25]. Most of the economic evaluations did not discount costs and effects in line with the relatively short time horizons of the within-trial analyses. Where studies had extrapolated cost and effects beyond the trial follow-up, discount rates of between 3.0% [25] and 3.5% [16, 28, 29, 33] per annum were applied to both costs and effects. In terms of study quality, all 19 studies (100%) reported using a randomisation process that was assessed as adequate according to the criteria described in Appendix B in the supplementary electronic material: 16 (84%) were un-blinded and nine (47%) reported  $\geq 80\%$  completion rates for the primary outcome at end of follow-up.

### 3.2 Measurement of Health-Related Quality of Life

The most widely used generic preference-based health-related quality of life instruments include the EuroQoL EQ-5D [5] and the SF-6D [6], but other generic preference-based instruments such as the HUI-3 [7] and the 15D instrument [34] are also available for use.

Table 2 presents a summary of the health-related quality of life data collection methods applied in the included studies and how non-availability of health-related quality of life data at baseline was handled in the QALY estimation. For the purpose of this review, we measured the baseline (or first) time point for describing health-related quality of life as reported by individual studies; conventionally, in trial-based economic evaluations, this is taken as the time of randomisation. Nine (47%) of the 19 studies used the EQ-5D to measure health-related quality of life of patients, five (26%) used the EQ-5D in combination with another instrument (primarily the SF-12/36 [26, 33, 35], HUI-3 [32] and the paediatric PedsQL [36]), one (5%) [18] used the 15D instrument [34] and another one (5%) used

the HUI-3 [28]. The remaining three studies (17%) [25, 27, 29] did not report a primary health-related quality of life data collection process. Rather, the economic evaluations in these three studies were informed by utility data extracted from external sources. In the study by Harvey et al. [29], age- and gender-specific utility values for the UK adult population were combined with survival estimates from the trial in order to estimate QALYs in the cost-utility analysis. Specifically, Harvey et al. [29] “estimated the quality-adjusted life expectancy for each survivor at hospital discharge based upon the Office of National Statistics age and sex-specific life expectancy tables and the EQ-5D age- and sex-specific quality of life weights”. Gyrd-Hansen et al. [25] used 5-year QALY estimates for patients living with stroke, obtained from the Oxford Stroke study [37] and stratified by stroke severity, to inform the cost-utility analysis. Rosenthal et al. [27] applied a QALY reduction of 0.37 per day for patients in an intensive care unit, which they obtained from a secondary study [38]. In addition, “an extra decrement of 0.2 QALYs was assumed for patients at age 65 years, and an annual decrement of 0.005 [39] each year over 65 was considered as well”.

### 3.3 Methods Used to Handle Non-Availability of Baseline Health-Related Quality of Life in QALY Estimation

Only one [40] of the 16 studies that prospectively collected health-related quality of life data was able to do so at baseline [using data from 932 (86%) of the 1084 study participants]. This study recruited patients with acute severe asthma from emergency departments. Patients had to be able to at least provide verbal consent, and those with life-threatening illness were excluded. EQ-5D data were collected at baseline by the recruiting physician. In the remaining 15 studies, the earliest time point recorded for data collection directly from study participants varied from 2 days post randomisation [41] to 12 months post randomisation. The reported reasons for not assessing health-related quality of life at randomisation mostly reflected the condition of trial participants at this time point, concerns around utility measurement in these clinical settings, and reluctance to prioritise health-related quality of life assessment in studies with substantial data collection burden.

The reported strategies used to handle the non-availability of health-related quality of life data at baseline in subsequent cost-utility analyses were available from 14 of the 15 studies that failed to collect baseline data, and can be classified into four broad categories (note, the economic evaluation based on REACT-2 trial [32] data had not yet been published at the time this review was undertaken):

**Table 2** Summary of health-related quality of life measurement and assumptions around how baseline health-related quality of life information was incorporated into QALY estimation

First author (year)	Instrument	Data collection time points (months)	Source of baseline health-related quality of life data	Methods used to estimate overall QALYs	Was the impact of baseline utility values used on cost-effectiveness results investigated and/or reported through sensitivity analyses?
Goodacre (2004) [41]	EQ-5D	0.07 <sup>a</sup> , 1, 6	First time point data	EQ-5D data over 6 months were combined with survival data to estimate QALYs using the area-under-the-curve methods	No
Harvey (2006) [29]	None	None	UK population norms	EQ-5D age- and sex-specific health-related quality of life weights for UK adult population combined with survival estimates from the trial	No
Dixon (2009) [43]	EQ-5D	1	Identical baseline scores assumed	Incremental QALYs estimated assuming a linear change in EQ-5D scores and that the two groups have identical scores at baseline	No
Peek (2010) [33]	EQ-5D; SF-36	6	Zero score assumed	Linear interpolation	No
Bohmer (2011) [18]	15D instrument	0.5, 3, 12	14-day retrospective recall	Linear interpolation	No
Goodacre (2011), [17]	EQ-5D	1, 3, 6	Zero score assumed	Area under the curve	No
Goodacre (2011), [16]	EQ-5D	1, 3	Zero score assumed	“Data are reported assuming that EQ-5D was zero at baseline, although means for any baseline score between 0 and 1 can be estimated by adding a constant $k/24$ , where $k$ is the baseline EQ-5D score of interest. As 3 months is approximately one-quarter of a year, the maximum possible number of QALYs accrued is 0.25 (assuming EQ-5D was 1 at baseline)”	No
Lamb (2012) [42]	EQ-5D	0.5, 4, 8, 12	UK pop norm assumed	Baseline utility were taken from the Health Survey for England and matched on age	No
Gates (2013) [13]	EQ-5D	6, 12	Unconscious (−0.402)	An unconscious patient (−0.402) for the baseline utility	No

Table 2 continued

First author (year)	Instrument	Data collection time points (months)	Source of baseline health-related quality of life data	Methods used to estimate overall QALYs	Was the impact of baseline utility values used on cost-effectiveness results investigated and/or reported through sensitivity analyses?
Powell (2013) [19]	Proxy EQ-5D, PedsQL	1	Delphi exercise followed by mapping	“A physician panel made up of two respiratory nurses and a consultant mapped the Yung Asthma Severity Score (ASS) scores on to EQ-5D health states from which baseline utility scores were estimated. In the base-case analysis, ASS scores of 1–3 were mapped on to an EQ-5D health state of 11111; ASS scores of 4–6 were mapped on to an EQ-5D health state of 22222; and ASS scores of 7–9 were mapped on to an EQ-5D health state of 33333”	Yes. This study reported varying the mapping algorithms used to estimate baseline utility as part sensitivity analyses. The results showed different assumptions about baseline utility values had little impact on the cost-effectiveness results (see page 55 of the HTA report)
Goedaere (2014) [40]	EQ-5D	0, 1	Baseline data collected	Not reported	Not applicable as study collected self-reported health-related quality of life data at baseline
Macrae (2014) [28]	HUI-3	12	First time point	Health-related quality of life was collected from a subset of patients who had traumatic brain injury at 12 months from ICU admission. The health-related quality of life of non-traumatic brain injury patients alive at 12 months post ICU admission was not assessed. Instead patients in this subgroup were assigned utility weights from another study. QALYs were then derived by applying the utility weights of each life year lived predicted for trial participants from 12 months post ICU admission. Patients who died before the 12 months period were assigned zero QALYs	No
Gyrd-Hansen (2015) [25]	None	None	External source	QALY estimates from external source (Oxford Vascular study) were stratified by stroke severity	
Lall (2015) [26]	EQ-5D, SF-12	6, 12	Unconscious state (–0.40)	Area under the curve assuming unconscious state (utility of –0.40) at baseline	No



Table 2 continued

First author (year)	Instrument	Data collection time points (months)	Source of baseline health-related quality of life data	Methods used to estimate overall QALYs	Was the impact of baseline utility values used on cost-effectiveness results investigated and or reported through sensitivity analyses?
Mouncey (2015) [15]	EQ-5D	3, 12	First time point data	QALYs estimated by combining survival data with quality-of-life scores at 90 days	No
Perkins (2015) [35]	EQ-5D, SF-12	3, 12	Unconscious (−0.402)	Cost-effectiveness analysis not yet published. Personal communication with the trial health economics team indicated that the analysis will assume an unconscious state at baseline by applying a utility score of −0.402	Cost-effectiveness analyses and results not yet reported
Rosenthal (2015) [27]	None	None	External source	QALYs and QALY decrements were obtained from another study and assigned to individual patients stratified by age	
Schuster (2015), [31]	EQ-5D	1, 6, 12	First time point	Area under the curve, starting from the first time point at which data were collected	No
Sierink (2016) [32]	EQ-5D, HUI-3	3, 6, 12	Cost-effectiveness analyses and results not yet published	Cost-effectiveness analyses and results not yet published	Cost-effectiveness analyses and results not yet published

HTA health technology assessment, HUI-3 Health Utilities Index-3, ICU intensive care unit, pop population, QALY quality-adjusted life year

<sup>a</sup> Retrospective recall at 14 days post randomisation

- (i) Eight studies (57%) assigned a fixed health utility value to all participants at baseline. This included assuming a zero value or a health state equivalent to death [16, 17, 33] or a utility value of  $-0.40$ , reflecting an unconscious health state for the EQ-5D-3L [13, 26, 35]. One study [42] obtained baseline utility from an external source (i.e. the Health Survey of England) stratified by age and sex, and another study [43] assumed equivalent baseline utility values across trial arms without reporting the actual values used.
- (ii) Four studies (29%) [15, 28, 31, 41] estimated QALYs using only the available data {i.e. from the first time point at which health-related quality of life was measured: 2 days post randomisation in the study by Goodacre et al. [41]; 1 month post randomisation in the study by Schuster et al. [31]; 3 months post randomisation in the study by Mouncey et al. [15]; and 1 year post randomisation in the subset of CHiP trial participants with traumatic brain injury [28]}. This effectively ignored the impact of interventions on participants' health-related quality of life prior to these time points on final QALY calculations.
- (iii) One study (7%) [18] asked patients to retrospectively recall at 14 days post randomisation their pre-randomisation health state.
- (iv) Finally, one other study (7%) [19] elicited external evidence on utility values associated with specific baseline health states using Delphi methods. The health status of participants measured at baseline was then translated/mapped onto EQ-5D-3L health states using evidence elicited from experts for incorporation in the final QALY calculations.

### 3.4 Assessment of Uncertainty Around Assumptions Used to Incorporate Baseline Utility in Final QALY Estimation

Extensive sensitivity analyses were used by the included studies to investigate the impact that methodological assumptions (mostly around the inclusion of different cost variables reflecting alternative perspectives for the economic evaluation) had on incremental cost-utility estimates. However, only one [19] of the 15 studies that did not collect baseline health-related quality of life data directly from patients (excluding the yet to be published analysis based on the REACT-2 trial [32]) specifically assessed the impact of the method and assumptions used to estimate baseline utilities on the cost-effectiveness results (Table 2). In that particular study [19], varying the assumptions used to estimate baseline utilities had little impact on the final cost-effectiveness results.

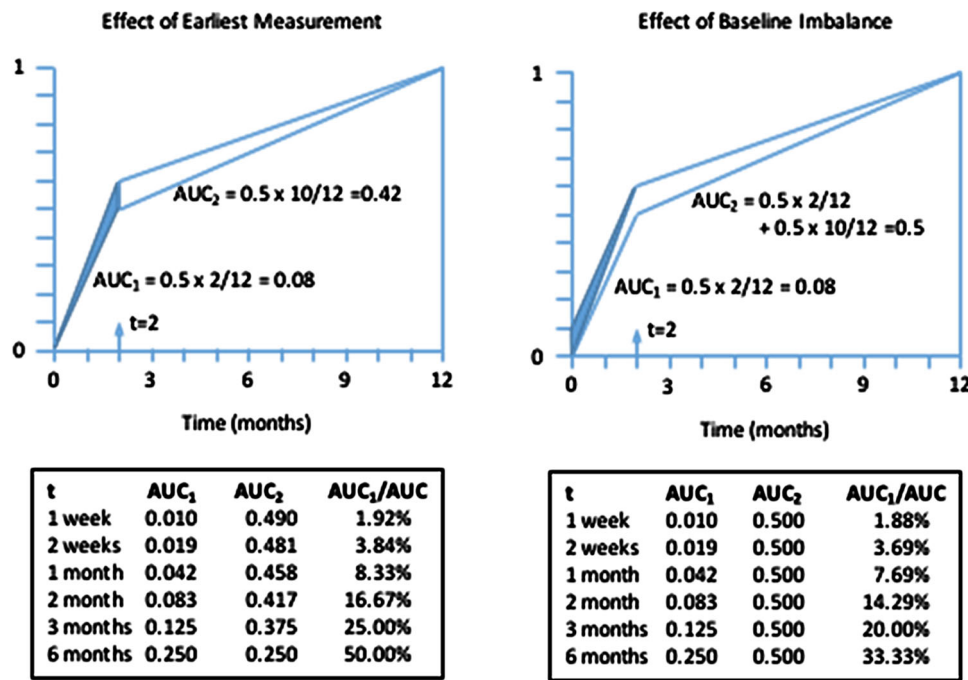
## 4 Implications of Methods for Handling Non-Availability of Baseline Quality of Life Data

### 4.1 Ignoring or Assuming a Fixed Baseline Utility Value

The selection of baseline health-related quality of life data in trial-based cost-utility analyses is significant in two ways: first, as an adjustment covariate within regression to estimate incremental costs and QALYs [44]; and second, as the first point in an area-under-the-curve (AUC) estimation of individual patient QALYs.

The importance of the method of baseline health-related quality of life measurement is driven by the success of the trial randomisation in achieving a balanced allocation of individuals (in terms of patient characteristics) between treatment arms. Baseline adjustment of health-related quality of life as a covariate within regression has become normal practice because of the need to manage the effect of baseline imbalances [44]. In the presence of baseline imbalances, different approaches to adjust for missing baseline health-related quality of life are likely to yield different answers. In this circumstance, exploring alternative baseline proxy covariates may provide the best approach, although trial stratification variables may adequately achieve this. As a general point, when estimating cost-effectiveness ratios with incremental QALYs close to zero, findings are likely to appear (perhaps artificially) sensitive to assumptions about baseline adjustment, since the incremental cost-effectiveness ratio (ICER) denominator may switch sign according to the approach taken.

In terms of AUC calculations, in the presence of balanced allocation, incremental QALY estimation may be robustly estimated. If it is assumed there is a true and variable unobserved baseline health-related quality of life, then assuming a fixed value should not systematically bias the incremental QALY gain or cost-effectiveness estimation. This is shown algebraically in Appendix D in the supplementary electronic material. In fact, in the presence of imbalances, imposing a fixed baseline in the presence of an adequate number of multiple follow-up valuations may only introduce limited bias, as only the first of a series of measurements contributing to the AUC is affected. For example, if QALY estimation is captured over a 1-year follow-up period and the first measurement is at 2 weeks, then any baseline assumption will have a small effect on the overall incremental QALY gain. Ignoring the true baseline and starting from a delayed first measurement may introduce more significant bias, since the area between the baseline and first measurement is lost. Conversely, this bias would be exacerbated in the absence of an adequate number of multiple follow-up valuations. Algebraically,



**Fig. 2** Effect of early measurement and baseline imbalance in health-related quality of life on incremental quality-adjusted life year (QALY) estimation using area-under-the-curve (AUC) approaches. Health-related quality of life (utility) weight is displayed on the vertical axis and follow-up time on the horizontal axis. The left plot shows the effect of early measurement and the right plot the effect of baseline imbalance on incremental QALY estimation. For example, assume the maximum follow-up is 12 months, there are no long-term differences between treatment groups and there is a difference of 0.1

at the earliest follow-up, time *t*. Assume further that *t* is a trial design choice and can be varied. Then the error of taking the AUC from the earliest follow-up and not attempting a baseline estimate is minimal at 1 week and considerable at 6 months, as shown in the box under the left plot, where  $AUC = AUC_1 + AUC_2$ . Similarly, having a baseline assessment is increasingly important the more delayed the first measurement. The right plot shows that when there is imbalance at baseline, ignoring it and choosing a common baseline value will have a minimal effect for an early first measurement

the degree of bias is proportional to the magnitude of the time interval between randomisation and the first data collection point and the magnitude of the QALY gain between the two time intervals (Appendix E in the supplementary electronic material). Similarly, Fig. 2 shows an example of a trial with a 12-month follow-up period where we assume no long-term differences between treatment groups and a difference of 0.1 at the earliest follow-up, time *t*. Assume time *t* is a trial design choice and can be varied. The error of taking the AUC from the earliest follow-up and not attempting a baseline estimate is minimal at 1 week and considerable at 6 months. If an analyst does include a baseline assessment, then it does not matter what baseline value is chosen between 0 and 1 (note, utility values can take negative values in practice),  $AUC_1$  is the same regardless. Having a baseline assessment is increasingly important the more delayed the first measurement. When there is imbalance at baseline, ignoring it and choosing a common baseline value will have a minimal effect for an early first measurement. Suppose the baseline imbalance in health utility was 0.1 (the same as the treatment effect at time *t*). Then the bias of missing the

imbalance in the baseline model is similar to the error of not adjusting for baseline in a baseline balanced model.

Consequently the eight studies that assigned a fixed baseline value should have produced unbiased estimates of incremental benefit in the absence of baseline imbalance; the three studies starting estimation from post treatment would similarly be adequate if the duration between randomisation and the timing of the first measurement is a small proportion of the overall follow-up period. In all circumstances, the frequency of follow-up time points needs to be adequate to characterise the treatment effect, but has been simplified in Fig. 2 for illustration.

#### 4.2 Retrospective Recall of the Baseline Health-Related Quality of Life Data

In the study by Bohmer et al. [18], “patients were carefully instructed to report health-related quality of life as it was experienced 14 days before the infarction (baseline value)”. The main appeal of retrospective recall is that baseline health-related quality of life data can be obtained from trial participants themselves. QALY estimates can

also be adjusted to account for potential imbalances between groups [44]. The most obvious limitation is that it is not possible to obtain direct estimates from deceased or permanently incapacitated patients (who would not be missing at random). For example, in the PARAMEDIC 1 trial [35], only 6.6% of 1471 individuals experiencing out-of-hospital cardiac arrest recruited into the trial survived to 3 months post arrest, the first time point at which health-related quality of life data were collected. Asking patients to retrospectively recall their baseline health-related quality of life would not be an option for the majority of this trial's participants.

Another limitation of retrospective recall is the possibility of introducing recall bias in the final QALY estimation. The extent of any recall bias may depend on the clinical and demographic characteristics of patients and the length of the recall period [45]; the longer this is, the more difficult it will be for patients to accurately recall and report on their baseline health-related quality of life. Wilson et al. [46] compared the use of retrospective recall of baseline health status versus population norms (New Zealand) in estimating change in health state valuations following acute-onset illness or injury. Their findings indicate a small but significant difference between pre- and post-injury health-related quality of life for people who had fully recovered, with recalled pre-injury health-related quality of life being higher than reported post-injury health. The reported health-related quality of life of the fully recovered patients was also higher than adult population norms. The authors concluded that "retrospective evaluation of health status is more appropriate than the application of population norms to estimate health status prior to acute-onset injury or illness, although there may be a small upward bias in such measurements." Finally, provided that recall bias is similar in magnitude and direction across treatment arms, it is unlikely to lead to a large effect on incremental QALYs.

### 4.3 Eliciting External Evidence and Using Mapping Techniques to Derive Baseline Health-Related Quality of Life Data

Powell et al. [19] employed a more sophisticated technique to estimate baseline health-related quality of life based on a clinical outcome [Yung Asthma Severity Score (ASS) in the context of acute severe asthma in children] measured at baseline. A physician panel comprising two respiratory nurses and a consultant were asked to translate/map (not on the basis of a pre-existing association but rather their clinical opinion) ASS scores measured at baseline onto EQ-5D-3L health states, from which baseline utility scores were estimated. More generally, 'mapping' techniques can be used to derive baseline health-related quality of life data

if a condition-specific outcome measure that is better able to reflect changes to individuals' health statuses can be collected at baseline and at subsequent follow-up time points when individuals are able to provide information on their health-related quality of life. The relationship between the condition-specific outcome measure and the preference-based health-related quality of life measure can be derived using the data at follow-up time points when both outcome measures are collected. The mapping coefficients can then be used to derive baseline utility values based on responses to the condition-specific outcome measure at baseline.

As with the management of non-availability of baseline health-related quality of life data, the choice of external utility values elicited from experts is unlikely to have a significant effect on incremental QALY estimation if balanced trial allocation is achieved, and may not introduce significant bias where baseline differences are small, or subsequent health-related quality of life measurement is adequately informative.

## 5 Discussion

### 5.1 Summary

This review describes the conduct of cost-utility analyses alongside randomised controlled trials in emergency and critical care settings. In this context, the estimation of QALYs is problematic because of difficulties in collecting health-related quality of life data from acutely ill or injured patients around the time of recruitment into trials. Four approaches for handling the lack of baseline health-related quality of life data in QALY calculations were identified among the 19 studies included in the review: (1) assigning a fixed health state utility value (typically assumed to be zero, a utility value for an unconscious health state, or stratified by important predictors of health-related quality of life) to all patients at baseline; (2) ignoring baseline health-related quality of life and estimating QALYs on the basis of available data at later time points; (3) retrospective recall of baseline health-related quality of life; and (4) mapping from disease-specific outcomes measured at baseline onto generic preference-based health-related quality of life outcomes. The results suggest that there is no uniformity in approach amongst researchers conducting trial-based economic evaluations regarding the most appropriate strategy for dealing with the problem of non-availability of primary baseline health-related quality of life data when estimating QALYs to inform trial-based cost-utility analyses within emergency and critical care settings.

Some implications of the methods used for dealing with the lack of baseline health-related quality of life data have been explored. To permit robust trial-based cost-utility analysis, a critical factor is whether the treatment arms are balanced with respect to the health-related quality of life of patients at baseline. By definition this is not observed directly, but may be implied by measured baseline differences in trial covariates. In this circumstance, proxy covariate adjustment of health-related quality of life estimates should be explored. In terms of AUC QALY estimation, provided randomisation has resulted in balanced treatment groups at baseline and the first health-related quality of life assessment occurs early in the overall period of assessment (e.g. 2 weeks into a 52-week follow-up), then all methods are likely to give fairly similar answers, and estimated incremental QALYs will not vary significantly. However, use of a fixed baseline may reconstruct treatment arm QALY gains more faithfully and may be preferable to simply starting QALY estimation from the first post-randomisation measurement; the fixed baseline health state utility value cancels out in the calculation of incremental QALYs (Appendix D in the supplementary electronic material). In general, ignoring the baseline health-related quality of life measurement may increase potential biases, particularly if there is substantial delay in the first trial measurement point (Appendix E in the supplementary electronic material and Fig. 2).

A potential limitation of the review is the omission of studies that meet the inclusion criteria in the review process (either at the search or screening and selection stages of the review). This can occur where eligible studies fail to report sufficient details in titles and abstracts to enable them to be identified as trial-based economic evaluations. For example, where trials fail to find difference in effectiveness between comparator interventions, there may be insufficient interest to include health economic outcomes within the main trial report or publish separate cost-effectiveness findings. The goal of the review was to characterise the types of approach used to compensate for unobtainable baseline utilities within the literature. Although some further eligible studies might have been obtained by more sensitive search methods, we have not identified any further approaches not already captured within this review.

## 5.2 Recommendations for Design of Future Trial-Based Economic Evaluations in Emergency and Critical Care Settings

- It is evident from the discussions above that an appropriate randomisation strategy should be employed to promote treatment groups that are similar in observable and unobservable patient characteristics. This in turn makes it likely that an unbiased estimate of

incremental QALYs is produced irrespective of the strategy for dealing with the lack of baseline health-related quality of life data. It also limits the need for more complicated adjusted analyses to correct for the imbalances in baseline health utilities. It is acknowledged that, as the outcome of randomisation is probabilistic, the best that randomisation can achieve is groups that are 'similar'. It is thus possible that a perfectly valid strategy can still end up with a chance imbalance because there is a limit to the number of stratifying variables within randomisation, and blocking breaks down with low recruiting centres where these randomisation strategies are employed.

- It is also evident that when possible, the initial assessment of patient health-related quality of life should be conducted at the earliest time possible post randomisation. This might mean the initial assessment of health-related quality of life is conducted at different time points as and when each patient is able to complete the health-related quality of life questionnaire. The differential times for the initial assessment would then be taken into account in the subsequent analyses. This is unlikely to cause problems if variation in first measurement time is random or small relative to the total follow-up. If, however, different treatments lead to substantially different durations to first measurement, this might be an issue for incremental QALY estimation. Further research should be considered to investigate whether or not collecting data at early time points offers advantages over data collection at a fixed time point for all patients.
- Ignoring baseline utilities altogether in final QALY estimation is generally not preferable as this approach may result in biased estimates of incremental QALYs, as demonstrated in Sect. 4.1 and Fig. 2.
- Identification and, where possible, collection of data on clinical variables (outcomes) that are strongly correlated with health-related quality of life and hence can be used to predict baseline health-related quality of life using mapping algorithms offer a route for further research enquiry.
- In the context of data collection, if the incorporation of health-related quality of life data is considered burdensome by trialists, it is important that health economists provide clear methodological guidance on best methods that balance the need to minimise respondent burden against the requirement for minimising analytical biases.

## 5.3 Concluding Remarks

Baseline health-related quality of life measurement is problematic in trial-based economic evaluations conducted in emergency and critical care settings. Consequently, trial-



based cost-utility analyses have used different methods that make different assumptions about baseline health utilities. Key messages that come out of this study include the need to employ appropriate randomisation strategies to ensure baseline comparability across treatment groups, initial assessment of health-related quality of life of patients at the earliest time possible post randomisation and, where appropriate, inclusion of a constant or imputed baseline utility value rather than ignoring it. Further research is needed in order to determine the impact of different assumptions upon cost-effectiveness results, and to identify best methodological practice in this area.

**Author contributions** The concept of this manuscript was jointly conceived by Melina Dritsaki, Felix Achana, and Stavros Petrou. Melina Dritsaki and Felix Achana gathered and reviewed the literature and data and drafted the initial manuscript. James Mason drafted the initial work presented in Sect. 4. All authors participated in the interpretation of data and preparation of the final manuscript.

#### Compliance with Ethical Standards

Melina Dritsaki, Felix Achana, James Mason, and Stavros Petrou have no conflicts of interest to disclose. There was no research funding support for this study.

**Data availability statement** The authors declare that the data supporting the findings of this study are available within the article and its supplementary information files.

#### References

- Glick HA, Doshi JA, Sonnad SS, Polsky D. Economic evaluation in clinical trials. Oxford: Oxford University Press; 2007.
- Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. *BMJ*. 2011;342:d1548.
- CDR Directorate at Canadian Coordinating Office for Health Technology Assessment. Canada common drug review submission guidelines for manufacturers. CCOHTA, 2013.
- Guidelines for the Pharmaceutical Industry on Preparation of Submissions to the Pharmaceutical Benefits Advisory Committee. Canberra: PBAC; 2002.
- Dolan P, Roberts J. Modelling valuations for Eq-5d health states: an alternative model using differences in valuations. *Med Care*. 2002;40(5):442–6.
- Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ*. 2002;21(2):271–92.
- Furlong WJ, et al. The Health Utilities Index (HUI) system for assessing health-related quality of life in clinical studies. *Ann Med*. 2001;33(5):375–84.
- Dolan P. Modeling valuations for EuroQol health states. *Med Care*. 1997;35:1095–108.
- Devlin N, Shah K, Feng Y, Mulhern B, Van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. Research Paper 16/01, Office of Health Economics; 2016.
- Brazier J, Roberts JR. The estimation of a preference-based index from the SF-12. *Med Care*. 2004;42(9):851–9.
- Shah HA, Dritsaki M, Pink J, Petrou S. Psychometric properties of Patient Reported Outcome Measures (PROMs) in patients diagnosed with Acute Respiratory Distress Syndrome (ARDS). *Health Qual Life Outcomes*. 2016;14:15.
- Smith K, Andrew E, Lijovic M, Nehme Z, Bernard S. Quality of life and functional outcomes 12 months after out-of-hospital cardiac arrest. *Circulation*. 2015;131:174–181.
- Gates S, et al. Beta-Agonist Lung injury TRial-2 (BALTI-2): a multicentre, randomised, double-blind, placebo-controlled trial and economic evaluation of intravenous infusion of salbutamol versus placebo in patients with acute respiratory distress syndrome. *Health Technol Assess*. 2013;17(38):v–vi, 1–87.
- Young D, et al. High-frequency oscillation for acute respiratory distress syndrome. *N Engl J Med*. 2013;368(9):806–13.
- Mouncey PR, et al. Trial of early, goal-directed resuscitation for septic shock. *N Engl J Med*. 2015;372(14):1301–11.
- Goodacre S, et al. The RATPAC (Randomised Assessment of Treatment using Panel Assay of Cardiac markers) trial: a randomised controlled trial of point-of-care cardiac markers in the emergency department. *Health Technol Assess*. 2011;15(23):iii–xi, 1–102.
- Goodacre S, et al. Health utility and survival after hospital admission with acute cardiogenic pulmonary oedema. *Emerg Med J*. 2011;28(6):477–82.
- Bohmer E, et al. Health and cost consequences of early versus late invasive strategy after thrombolysis for acute myocardial infarction. *Eur J Cardiovasc Prev Rehabil*. 2011;18(5):717–23.
- Powell CV, et al. MAGNESium Trial In Children (MAGNETIC): a randomised, placebo-controlled trial and economic evaluation of nebulised magnesium sulphate in acute severe asthma in children. *Health Technol Assess*. 2013;17(45):v–vi, 1–216.
- Pines JM, Fager SS, Milzman DP. A review of costing methodologies in critical care studies. *J Crit Care*. 2002;17(3):181–6.
- Lopez AD, Mathers CD, Ezzati M, Jamison D, Murray CJL. The global and regional burden of disease and risk factors, 2001: systematic analysis of population health data. *Lancet*. 2006;367(9524):1747–57.
- Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. The PRISMA Group (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine* (OPEN ACCESS). 2009.
- Kendrick D, Young B, Mason-Jones AJ, Ilyas N, Achana FA, Cooper NJ, Hubbard SJ, Sutton AJ, Smith S, Wynn P, Mulvaney C, Watson MC, Coupland C. Home safety education and provision of safety equipment for injury prevention (Review). *Evid Based Child Health*. 2013;8(3):761–939.
- Husereau D, Drummond M, Petrou S, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS)—explanation and elaboration: a report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value Health*. 2013;16:231–50.
- Gyrd-Hansen D, et al. Cost-effectiveness estimate of prehospital thrombolysis: results of the PHANTOM-S study. *Neurology*. 2015;84(11):1090–7.
- Lall R, et al. A randomised controlled trial and cost-effectiveness analysis of high-frequency oscillatory ventilation against conventional artificial ventilation for adults with acute respiratory distress syndrome. The OSCAR (OSCillation in ARDS) study. *Health Technol Assess*. 2015;19(23):1–177, vii.
- Rosenthal VD, Udwardia FE, Kumar S, Poojary A, Sankar R, Orellano PW, Durgad S, Thulasiraman M, Bahirune S, Kumbhar S, Patil P. Clinical impact and cost-effectiveness of split-septum and single-use prefilled flushing device vs 3-way stopcock on central line-associated bloodstream infection rates in India: a randomized clinical trial conducted by the International



- Nosocomial Infection Control Consortium (INICC). *Am J Infect Control*. 2015;43:1040–1045.
28. Macrae D, Grieve R, Allen E, Sadique Z, Betts H, Morris K, et al. A clinical and economic evaluation of Control of Hyperglycaemia in Paediatric intensive care (CHiP): a randomised controlled trial. *Health Technol Assess*. 2014;18(26):1–210. doi:10.3310/hta18260
  29. Harvey S, et al. An evaluation of the clinical and cost-effectiveness of pulmonary artery catheters in patient management in intensive care: a systematic review and a randomised controlled trial. *Health Technol Assess*. 2006;10(29):iii–iv, ix–xi, 1–133.
  30. National Institute for Health and Clinical Excellence. NICE guide to the methods of technology appraisal. NICE, 2013.
  31. Schuster A, Faulkner M, Zeymer U, Ouarrak T, Eitel I, Desch S, Hasenfuß G, Thiele H. Economic implications of intra-aortic balloon support for myocardial infarction with cardiogenic shock: an analysis from the IABP-SHOCK II-trial. *Clin Res Cardiol*. 2015;104:566–73.
  32. Sierink JC, Treskes K, Edwards MJR, Beuker BJA, den Hartog D, Hohmann J, Dijkgraaf MGW, Luitse JSK, Beenen LFM, Hollmann MW, Goslings JC, For the REACT-2 study group\*. Immediate total-body CT scanning versus conventional imaging and selective CT scanning in patients with severe trauma (REACT-2): a randomised controlled trial. *Lancet*. 2016;388:673–83.
  33. Peek GJ, et al. Randomised controlled trial and parallel economic evaluation of conventional ventilatory support versus extracorporeal membrane oxygenation for severe adult respiratory failure (CESAR). *Health Technol Assess*. 2010;14(35):1–46.
  34. Sintonen H. Comparing properties of the 15D and the EQ-5D in measuring health-related quality of life. *Arch Hellen Med*. 2001;18(2):156–160.
  35. Perkins GD, et al. Mechanical versus manual chest compression for out-of-hospital cardiac arrest (PARAMEDIC): a pragmatic, cluster randomised controlled trial. *Lancet*. 2015;385(9972):947–55.
  36. Varni JW, Limbers CA. The PedsQL 4.0 Generic Core Scales Young Adult Version: feasibility, reliability and validity in a university student population. *J Health Psychol*. 2009;14(4):611–22.
  37. Luengo-Fernandez R, et al. Quality of life after TIA and stroke: ten-year results of the Oxford Vascular Study. *Neurology*. 2013;81(18):1588–95.
  38. Sznajder M, Aegerter P, Launois R, Merliere Y, Guidet B, CubRea V. A cost-effectiveness analysis of stays in intensive care units. *Intensive Care Med*. 2001;27:146–53.
  39. Dick AW, Perencevich EN, Pogorzelska-Maziarz M, Zwanziger J, Larson EL, Stone PW. A decade of investment in infection prevention: a cost-effectiveness analysis. *Am J Infect Control*. 2015;43:4–9.
  40. Goodacre S, Cohen J, Bradburn M, Stevens J, Gray A, Benger J, et al. 3Mg trial: a randomised controlled trial of intravenous or nebulised magnesium sulphate versus placebo in adults with acute severe asthma. *Health Technol Assess*. 2014;18(22):1–168.
  41. Goodacre S, et al. Randomised controlled trial and economic evaluation of a chest pain observation unit compared with routine care. *BMJ*. 2004;328(7434):254.
  42. Lamb SE, et al. Managing Injuries of the Neck Trial (MINT): a randomised controlled trial of treatments for whiplash injuries. *Health Technol Assess*. 2012;16(49):iii–iv, 1–141.
  43. Dixon S, et al. Is it cost effective to introduce paramedic practitioners for older people to the ambulance service? Results of a cluster randomised controlled trial. *Emerg Med J*. 2009;26(6):446–51.
  44. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Econ*. 2005;14(5):487–96.
  45. Schmier JK, Halpern MT. Patient recall and recall bias of health state and health status. *Expert Rev Pharmacoecon Outcomes Res*. 2004;4(2):159–63.
  46. Wilson R, et al. Retrospective evaluation versus population norms for the measurement of baseline health status. *Health Qual Life Outcomes*. 2012;10:68.