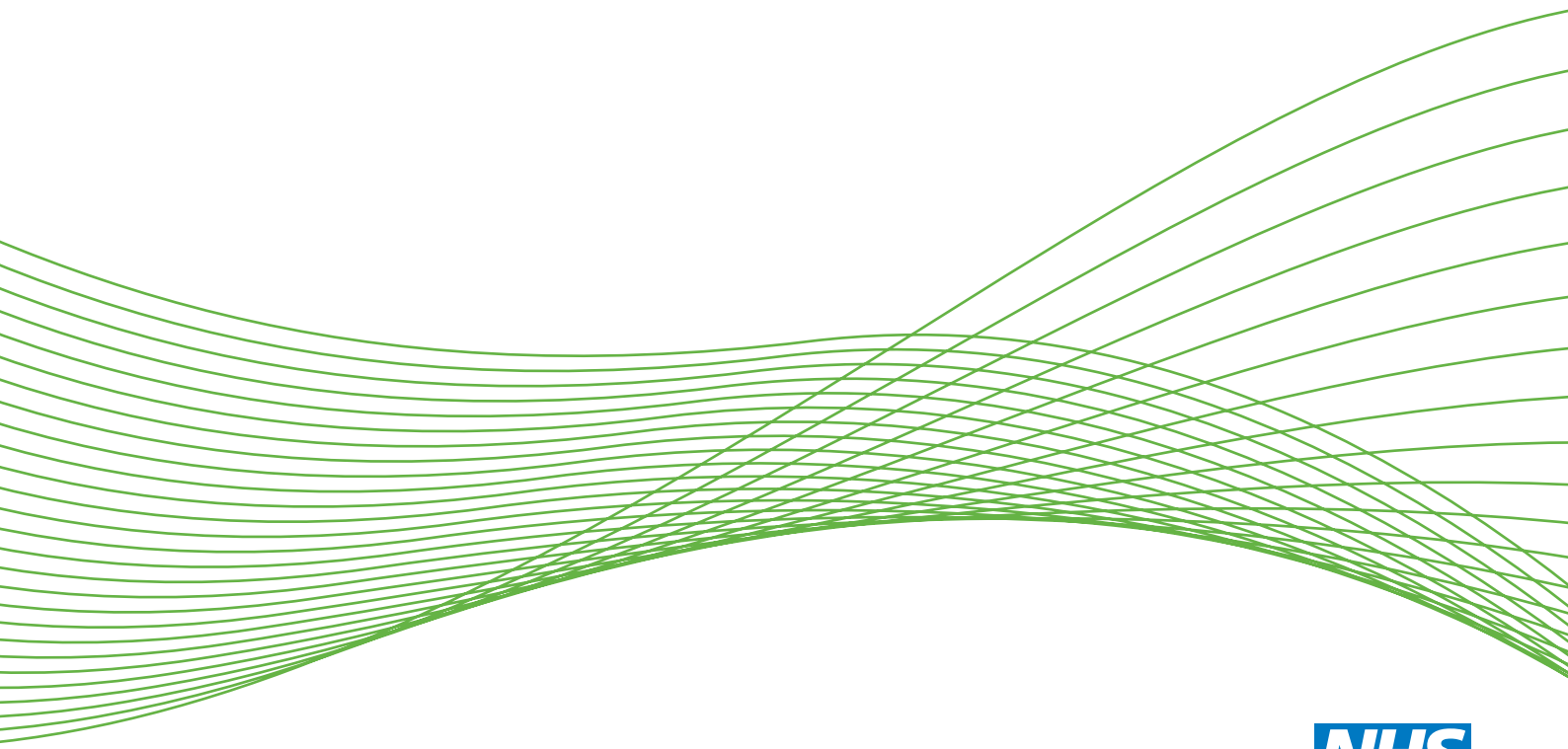
HEALTH TECHNOLOGY ASSESSMENT

VOLUME 17 ISSUE 34 AUGUST 2013
ISSN 1366-5278

Clinical effectiveness and cost-effectiveness of depth of anaesthesia monitoring (E-Entropy, Bispectral Index and Narcotrend): a systematic review and economic evaluation

J Shepherd, J Jones, GK Frampton, J Bryant, L Baxter and K Cooper



NHS
National Institute for Health Research

DOI 10.3310/hta17340

Clinical effectiveness and cost-effectiveness of depth of anaesthesia monitoring (E-Entropy, Bispectral Index and Narcotrend): a systematic review and economic evaluation

J Shepherd,* J Jones, GK Frampton, J Bryant, L Baxter and K Cooper

Southampton Health Technology Assessments Centre (SHTAC), University of Southampton, Southampton, UK

*Corresponding author

Declared competing interests of the authors: none

Published August 2013
DOI: 10.3310/hta17340

This report should be referenced as follows:

Shepherd J, Jones J, Frampton GK, Bryant J, Baxter L, Cooper K. Clinical effectiveness and cost-effectiveness of depth of anaesthesia monitoring (E-Entropy, Bispectral Index and Narcotrend): a systematic review and economic evaluation. *Health Technol Assess* 2013;**17**(34).

Health Technology Assessment is indexed and abstracted in *Index Medicus*/MEDLINE, *Excerpta Medica*/EMBASE, *Science Citation Index Expanded* (*SciSearch*®) and *Current Contents*®/Clinical Medicine.

Health Technology Assessment

ISSN 1366-5278 (Print)

ISSN 2046-4924 (Online)

Five-year impact factor: 5.804

Health Technology Assessment is indexed in MEDLINE, CINAHL, EMBASE, The Cochrane Library and the ISI Science Citation Index and is assessed for inclusion in the Database of Abstracts of Reviews of Effects.

This journal is a member of and subscribes to the principles of the Committee on Publication Ethics (COPE) (www.publicationethics.org/).

Editorial contact: nihredit@southampton.ac.uk

The full HTA archive is freely available to view online at www.journalslibrary.nihr.ac.uk/hta. Print-on-demand copies can be purchased from the report pages of the NIHR Journals Library website: www.journalslibrary.nihr.ac.uk

Criteria for inclusion in the *Health Technology Assessment* journal

Reports are published in *Health Technology Assessment* (HTA) if (1) they have resulted from work for the HTA programme, and (2) they are of a sufficiently high scientific quality as assessed by the reviewers and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

HTA programme

The HTA programme, part of the National Institute for Health Research (NIHR), was set up in 1993. It produces high-quality research information on the effectiveness, costs and broader impact of health technologies for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined as all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care.

The journal is indexed in NHS Evidence via its abstracts included in MEDLINE and its Technology Assessment Reports inform National Institute for Health and Care Excellence (NICE) guidance. HTA research is also an important source of evidence for National Screening Committee (NSC) policy decisions.

For more information about the HTA programme please visit the website: www.hta.ac.uk/

This report

The research reported in this issue of the journal was commissioned and funded by the HTA programme on behalf of NICE as project number 11/57/01. The protocol was agreed in November 2011. The assessment report began editorial review in April 2012 and was accepted for publication in September 2012. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the reviewers for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

This report presents independent research funded by the National Institute for Health Research (NIHR). The views and opinions expressed by authors in this publication are those of the authors and do not necessarily reflect those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health. If there are verbatim quotations included in this publication the views and opinions expressed by the interviewees are those of the interviewees and do not necessarily reflect those of the authors, those of the NHS, the NIHR, NETSCC, the HTA programme or the Department of Health.

Published by the NIHR Journals Library (www.journalslibrary.nihr.ac.uk), produced by Prepress Projects Ltd, Perth, Scotland (www.prepress-projects.co.uk).

© Queen's Printer and Controller of HMSO 2013. This work was produced by Shepherd *et al*. under the terms of a commissioning contract issued by the Secretary of State for Health. This issue may be freely reproduced for the purposes of private research and study and extracts (or indeed, the full report) may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising. Applications for commercial reproduction should be addressed to: NIHR Journals Library, National Institute for Health Research, Evaluation, Trials and Studies Coordinating Centre, Alpha House, University of Southampton Science Park, Southampton SO16 7NS, UK.

Abstract

Clinical effectiveness and cost-effectiveness of depth of anaesthesia monitoring (E-Entropy, Bispectral Index and Narcotrend): a systematic review and economic evaluation

J Shepherd,* J Jones, GK Frampton, J Bryant, L Baxter and K Cooper

Southampton Health Technology Assessments Centre (SHTAC), University of Southampton, Southampton, UK

*Corresponding author

Background: It is important that the level of general anaesthesia (GA) is appropriate for the individual patient undergoing surgery. If anaesthesia is deeper than required to keep a patient unconscious, there might be increased risk of anaesthetic-related morbidity, such as postoperative nausea, vomiting and cognitive dysfunction. This may also prolong recovery times, potentially increasing health-care costs. If anaesthesia is too light, patients may not be fully unconscious and could be at risk of intraoperative awareness.

Objective: The objective of this report is to assess the clinical effectiveness and cost-effectiveness of Bispectral Index (BIS), E-Entropy and Narcotrend technologies, each compared with standard clinical monitoring, to monitor the depth of anaesthesia in surgical patients undergoing GA.

Data sources: A search strategy was developed and run on a number of bibliographic electronic databases including MEDLINE, EMBASE, The Cochrane Library and the Health Technology Assessment (HTA) database. For the systematic review of patient outcomes, databases were searched from the beginning of 2009 to November 2011 for studies of BIS (and then updated in February 2012), and from 1995 to November 2011 (and then updated in February 2012) for studies of E-Entropy and Narcotrend. For the systematic review of cost-effectiveness, searches were from database inception to November 2011 (an update search was performed in February 2012).

Review methods: The systematic review of patient outcomes followed standard methodology for evidence synthesis. A decision-analytic model was developed to assess the cost-effectiveness of depth of anaesthesia monitoring compared with standard clinical observation. A simple decision tree was developed, which accounted for patients' risk of experiencing short-term anaesthetic-related complications in addition to risk of experiencing intraoperative awareness.

Results: Twenty-two randomised controlled trials comparing BIS, E-Entropy and Narcotrend with standard clinical monitoring were included in the systematic review of patient outcomes, alongside evidence from a recent Cochrane review. Six trials of patients classified with risk factors for intraoperative awareness were combined in a fixed-effect meta-analysis. The overall pooled Peto's odds ratio was 0.45 (95% confidence interval 0.25 to 0.81) in favour of BIS. However, there was statistically significant heterogeneity. The base-case cost per quality-adjusted life-year (QALY) for BIS compared with standard clinical monitoring ranged from £22,339 to £44,198 depending on patient subgroups (type of GA received; level of risk for awareness). For E-Entropy, base-case estimates ranged from £14,421 to £31,430. For Narcotrend, estimates varied from a cost per QALY of £8033 to Narcotrend dominating standard clinical monitoring.

Limitations: The analysis was limited by lack of clinical effectiveness data, particularly for E-Entropy and Narcotrend.

Conclusions: The available evidence on the impact of the technologies on reducing the likelihood of intraoperative awareness is limited. However, there were reductions in general anaesthetic consumption and anaesthetic recovery times. The cost-effectiveness of depth of anaesthesia monitoring appears to be highly dependent on a number of factors, including probability of awareness.

Study registration: PROSPERO registration number CRD42011001834.

Funding: The National Institute for Health Research Health Technology Assessment programme.

Contents

List of abbreviations

Technical terms and abbreviations are used throughout this report. The meaning is usually clear from the context, but a glossary is provided for the non-specialist reader.

ASA	American Society of Anesthesiologists
BAG-RECALL	BIS or Anaesthetic Gas to Reduce Explicit Recall
BIS	Bispectral Index
BMI	body mass index
BNF	*British National Formulary*
CHEOPS	Children's Hospital of Eastern Ontario Pain Score
CI	confidence interval
ECG	electrocardiography
EEG	electroencephalography
EQ-5D	European Quality of Life-5 Dimensions
ETAC	end-tidal anaesthetic concentration
FGF	fresh gas flow
GA	general anaesthesia
HRQoL	health-related quality of life
HTA	health technology assessment
ICER	incremental cost-effectiveness ratio
ITT	intention to treat
i.v.	intravenous
IV	inverse variance
LPS	late psychological symptoms
MAC	minimum alveolar concentration
MMSE	Mini Mental State Examination
NICE	National Institute for Health and Care Excellence
PACU	postanaesthesia care unit
PAED	Paediatric Anaesthetic Emergence Delirium
POCD	postoperative cognitive dysfunction
PONV	postoperative nausea and vomiting
PTSD	post-traumatic stress disorder
PWHS	preference weighted health score
QALY	quality-adjusted life-year
QoL	quality of life
RCT	randomised controlled trial
SD	standard deviation
SF-36	Short Form questionnaire-36 items
SF-6D	Short Form questionnaire-6 Dimensions
SG	standard gamble
STAI	State–Trait Anxiety Inventory

TIVA	total intravenous anaesthesia	VAS	visual analogue scale
TTO	time trade-off		

All abbreviations that have been used in this report are listed here unless the abbreviation is well known (e.g. NHS), or it has been used only once, or it is a non-standard abbreviation used only in figures/tables/appendices, in which case the abbreviation is defined in the figure legend or in the notes at the end of the table.

Note

This monograph is based on the Diagnostic Assessment Report produced for NICE. The full report contained a considerable number of data that were deemed commercial-in-confidence. The full report was used by the Appraisal Committee at NICE in their deliberations. The full report with each piece of commercial-in-confidence data removed and replaced by the statement 'commercial-in-confidence data removed' is available on the NICE website: www.nice.org.uk.

The present monograph presents as full a version of the report as is possible while retaining readability, but some sections, sentences, tables and figures have been removed. Readers should bear in mind that the discussion, conclusions and implications for practice and research are based on all the data considered in the original full NICE report.

Scientific summary

Background

It is important that the level of general anaesthesia (GA) is appropriate for the individual patient undergoing surgery. If anaesthesia is deeper than required to keep a patient unconscious, there might be increased risk of anaesthetic-related morbidity, such as postoperative nausea, vomiting and cognitive dysfunction. If anaesthesia is too light, patients may not be fully unconscious and could be at risk of intraoperative awareness. Intraoperative awareness is a relatively rare event with an incidence typically of around one to two patients per 1000. However, over time, awareness may cause depression, anxiety and post-traumatic stress disorder (PTSD).

During GA, patients are routinely monitored for signs of potential intraoperative awareness, including tachycardia (rapid heart rate), hypertension, sweating, lacrimation (tear production), movement/grimacing and tachypnoea (rapid breathing). In patients receiving inhaled GA, end-tidal (exhaled) anaesthetic gas concentrations may be assessed to gauge depth of anaesthesia. However, clinical observation alone may not be a reliable surrogate marker of depth of anaesthesia. Technologies have been developed using electroencephalography (EEG) to measure and interpret electrical activity in the brain to provide a measure of unconsciousness. Most devices comprise a module that collects raw EEG data via sensors placed on the patient's forehead and then processes and analyses these using a mathematical algorithm. The output is then displayed numerically on a monitor for use by the anaesthetist to judge depth of unconsciousness, and to alter anaesthetic dose accordingly. Three such devices prioritised for this report are Bispectral Index (BIS), E-Entropy and Narcotrend.

Objectives

The objective of this report is to assess the clinical effectiveness and cost-effectiveness of BIS, E-Entropy and Narcotrend technologies to monitor the depth of anaesthesia in surgical patients undergoing GA.

Methods

Systematic review of patient outcomes

A systematic review of patient outcomes associated with depth of anaesthesia monitoring was conducted. A search strategy was developed and run on eight bibliographic electronic databases. Reference lists supplied by the device manufacturers were checked to identify potentially relevant studies. Eligibility criteria were applied to titles and abstracts and to full papers by two reviewers independently. Because of the relatively large volume of evidence for BIS, we included only trials that were supplemental to a recent Cochrane systematic review of BIS. Included studies were data extracted using a standard template. Risk of bias and markers of quality were assessed. The studies were synthesised narratively, with meta-analyses from the Cochrane review of BIS updated with supplemental studies where feasible and appropriate.

Systematic review of cost-effectiveness

A systematic review of the literature on the cost-effectiveness of depth of anaesthesia monitoring compared with standard clinical monitoring was undertaken. Included studies were evaluated for their quality and for generalisability to the UK. Eligibility criteria were applied to titles and abstracts and to full papers by two reviewers independently, and the studies were synthesised narratively.

Economic evaluation

A decision-analytic model was developed to assess the cost-effectiveness of depth of anaesthesia monitoring compared with standard clinical observation. A simple decision tree was developed, which accounted for patients' risk of experiencing short-term anaesthetic-related complications in addition to a risk of experiencing intraoperative awareness.

It was assumed that a proportion of patients who experience awareness will suffer psychological symptoms and that a proportion of those will develop PTSD and may seek treatment. A systematic review of health-related quality of life (HRQoL) in PTSD was undertaken in order to estimate the quality-of-life decrement to be applied as the result of any psychological symptoms arising from an awareness episode. The costs of depth of anaesthesia monitoring consist of the capital costs associated with acquisition of the monitor and recurring costs associated with sensors that are attached to the patient. Equivalent annual costs for each monitor were calculated for an effective equipment life of 5 years. Unit costs of anaesthetic drugs were derived from the *British National Formulary* (BNF) and supplied from an NHS Trust. The baseline incidence of awareness in high-risk patients was calculated from the control arms of randomised controlled trials (RCTs) in this group of patients. The summary values of the effectiveness of depth of anaesthesia monitoring were taken from our systematic review of patient outcomes.

The model evaluates costs [UK sterling (pounds) using a 2011 price base] from the perspective of the NHS and Personal Social Services. Outcomes in the model are expressed as quality-adjusted life-years (QALYs). Both costs and outcomes are discounted using a 3.5% annual discount rate, in line with current guidance.

Results

Systematic review of patient outcomes

From a total of 776 bibliographic records, 22 RCTs comparing BIS, E-Entropy and Narcotrend with standard clinical monitoring were included in the systematic review of patient outcomes. Fifteen trials of BIS, seven trials of E-Entropy and four trials of Narcotrend all compared with standard clinical monitoring. (Note that some trials compared more than one of the three devices to standard clinical monitoring.) Some of the trials reported that in the EEG arm anaesthesia doses were titrated according to device values in conjunction with clinical signs. In other trials the use of clinical signs alongside EEG monitoring was not explicit. The Cochrane review of BIS included 31 RCTs. The trials included in both reviews span the period between 1997 and 2011 in terms of publication date.

In many cases, the risk of bias in the trials was unclear because of limitations in reporting of methodological details. The trials varied in terms of their sample sizes, from as low as 20 to over 6000 patients, but, in general, sample sizes were relatively small (e.g. fewer than 200). Fifteen of the trials in this systematic review and all of the trials in the Cochrane BIS review were conducted in adult patients, of varying mean ages. Seven of the trials in this review were conducted with children. The trials were generally single-centre studies conducted in a range of locations including Europe, North America and Asia.

Six trials were conducted with patients classified as having one or more risk factors for intraoperative awareness (e.g. planned cardiac surgery, pulmonary hypertension, end-stage lung disease), all of which evaluated BIS monitoring. The trials tended to exclude patients with significant ill health or factors that may interfere with EEG recordings.

Explicit intraoperative awareness was assessed in 16 of the trials, but in most of these no episodes were recorded. However, awareness is a relatively rare event and the trials were not statistically powered to detect it. The six trials of patients classified with risk factors for intraoperative awareness, all of which evaluated BIS, were combined in a fixed-effect meta-analysis. The overall pooled Peto's odds ratio (OR) was 0.45 [95% confidence interval (CI) 0.25 to 0.81] in favour of BIS.

Caution is advised in the interpretation of this result as, overall, there was statistically significant heterogeneity ($p = 0.009$; $I^2 = 79\%$). Both the subgroup of trials, which included a trial of mixed inhaled and intravenous anaesthesia, and the subgroup, which included trials of total intravenous anaesthesia (TIVA), statistically favoured BIS monitoring. However, in the subgroup of trials that used only inhaled anaesthesia, the Peto's OR was 1.79 (95% CI 0.63 to 5.11), favouring standard clinical monitoring, although not statistically significant.

Systematic review of cost-effectiveness

A total of 134 potentially relevant references were identified by the cost-effectiveness searches. Of these, one study comparing BIS with standard clinical monitoring met all of the inclusion criteria. The study reported cost per avoided intraoperative recall, with the incidence of recall with BIS reported as 0.04% compared with 0.18% for standard monitoring, resulting in a cost per avoided recall of US$4410. The authors of the study concluded that BIS monitoring did not appear cost-effective. However, the results and conclusions should be viewed with caution because of poor methodological and reporting quality.

Economic evaluation

For each technology we presented a base-case analysis for two modes of anaesthetic administration {TIVA and mixed anaesthesia [induction with intravenous (i.v.) anaesthesia and maintenance with inhaled anaesthesia or a combination of inhaled and i.v. anaesthetic]} and for two patient populations (those considered at high risk of intraoperative awareness and a general surgical population, at average risk of intraoperative awareness).

Bispectral Index compared with standard clinical monitoring

In cohorts of 10,000 patients at high risk of intraoperative awareness undergoing GA with TIVA, the incremental cost-effectiveness ratio (ICER) for BIS compared with standard clinical monitoring in this population was £22,339.

For the population of general surgical patients undergoing GA with TIVA, BIS monitoring was modelled as being associated with 3.8 cases (per 10,000 patients) of awareness, compared with 16 in patients receiving standard clinical monitoring. Given the lower baseline risk of awareness in this population, the QALY gain with BIS monitoring was lower (0.0003) than for high-risk patients. This resulted in a higher ICER (£34,565).

Deterministic sensitivity analyses indicated that the ICER was sensitive to the same input parameters as for the population at high risk of awareness.

The baseline estimates of awareness, late psychological symptoms (LPS) and PTSD for high-risk patients undergoing mixed GA were the same as for high-risk patients undergoing TIVA. However, given that the OR of awareness with BIS monitoring was higher in this analysis, the estimated reduction in LPS and PTSD was lower. The ICER for BIS compared with standard clinical monitoring in this population was £29,634.

The baseline estimates of awareness, LPS and PTSD in the population of general surgical patients undergoing mixed GA were the same as for TIVA. Although a proportion of the higher cost associated with BIS monitoring was offset by reduction in anaesthetic consumption, the cost-saving for inhaled anaesthesia was lower than for TIVA. As a result the incremental cost was greater. Given the lower baseline risk of awareness in this population, the QALY gain with BIS monitoring was lower (0.0003) than for high-risk patients, resulting in a higher ICER (£49,198).

Deterministic sensitivity analyses indicated that the ICER was sensitive to a number of parameters, including the baseline incidence of awareness and the effectiveness of BIS in reducing awareness.

E-Entropy compared with standard clinical monitoring

In patients at high risk of awareness undergoing GA with TIVA, the modelled cost per patient with E-Entropy monitoring was higher than with standard clinical monitoring, although some of the additional cost was offset by reduced cost associated with psychological sequelae of awareness. The ICER for E-Entropy compared with standard clinical monitoring in this population was £14,421.

In the population of general surgical patients undergoing GA with TIVA, E-Entropy monitoring had a higher cost per patient than standard clinical monitoring. There was no reduction in anaesthetic drug costs to offset the additional costs of E-Entropy monitoring. Given the lower baseline risk of awareness in this population, the QALY gain was lower than for high-risk patients, which resulted in a higher ICER (£31,131–31,430).

In patients considered at high risk of awareness undergoing mixed GA, E-Entropy monitoring had higher costs and improved outcomes compared with standard clinical monitoring. However, the QALY gain was lower than for patients undergoing TIVA. The ICER for E-Entropy compared with standard clinical monitoring in this population was £19,367.

In the population of general surgical patients undergoing mixed GA, E-Entropy monitoring had higher costs than standard clinical monitoring. In contrast with the analysis for TIVA, the clinical trial used to estimate inhaled anaesthetic drug consumption reported a substantial decrease (29%), which resulted in approximately half of the additional cost of E-Entropy monitoring being offset by a reduction in anaesthetic drug costs. Despite the lower baseline risk of awareness, which resulted in a lower QALY gain with E-Entropy monitoring than for high-risk patients, the lower incremental cost resulted in an equivalent ICER (£19,000).

Deterministic sensitivity analyses indicated that the ICER was sensitive to a number of parameters, including the baseline incidence of awareness and the effectiveness of E-Entropy in reducing awareness.

Narcotrend compared with standard clinical monitoring

In patients at high risk of awareness undergoing GA with TIVA, the modelled cost per patient with Narcotrend monitoring was higher than with standard clinical monitoring, although some of the additional cost was offset by reduced cost associated with psychological sequelae of awareness. The ICER for Narcotrend compared with standard clinical monitoring in this population was £5681. Deterministic sensitivity analyses indicated that the ICER was sensitive to a number of parameters, including the baseline incidence of awareness and the effectiveness in reducing awareness.

In the general surgical population undergoing GA with TIVA, and also undergoing mixed GA, Narcotrend monitoring had a lower cost per patient than standard clinical monitoring. The additional cost of monitoring was more than offset by reduction in anaesthetic drug consumption. Given the lower baseline risk of awareness in this population, the QALY gain was lower than for high-risk patients. Narcotrend dominated standard clinical monitoring. Narcotrend remained dominant in the majority of deterministic sensitivity analyses.

In patients at high risk of awareness undergoing mixed GA, Narcotrend monitoring had higher costs and improved outcomes than standard clinical monitoring, although the QALY gain (0.0005) was lower than for patients undergoing TIVA. The ICER for Narcotrend compared with standard clinical monitoring in this population was £8033. Deterministic sensitivity analyses indicated that the ICER was sensitive to the same parameters as for high-risk patients undergoing TIVA.

Conclusions

In general, BIS, E-Entropy and Narcotrend technologies for monitoring the depth of anaesthesia are associated with reductions in general anaesthetic consumption, and decreased anaesthetic recovery times, compared with monitoring of clinical signs alone. However, these reductions may be considered clinically modest. The available evidence on the impact of the technologies on reducing the likelihood of intraoperative awareness is limited. Overall, BIS was associated with a statistically significant reduction in intraoperative awareness in patients classified as at higher risk, although there is uncertainty in effect estimates because of significant heterogeneity. Caution is advised because of uncertainties about the risk of bias of many of the included trials, and because many outcome measures were not statistically powered.

The cost-effectiveness of depth of anaesthesia monitoring appears to be highly dependent on the incidence of awareness, the HRQoL impact of psychological sequelae of awareness and the probability of developing psychological illness following awareness, as well as the effectiveness of depth of anaesthesia monitoring in reducing awareness. Cost-savings resulting from reduced use of anaesthetic drugs may offset some of the additional cost of depth of anaesthesia monitoring. The cost of sensors attached to the patient appears to be a key factor in the additional cost of depth of anaesthesia monitoring.

This report makes the following research recommendations (in priority order):

1. RCTs of E-Entropy- and Narcotrend-guided anaesthesia monitoring are needed, in high-risk patients, with adequate statistical power to detect explicit intraoperative awareness, and of sufficient length of follow-up to detect delayed cases of awareness.
2. RCTs of all three technologies should also evaluate the effects of anaesthesia overdosing, including short-term effects, such as nausea and vomiting, as well as longer-term impact on cognitive function.
3. RCTs of E-Entropy- and Narcotrend-guided anaesthesia monitoring are also needed in children.

Study registration

This study is registered as PROSPERO CR042011001834.

Funding

Funding for this study was provided by the Health Technology Assessment programme of the National Institute for Health Research.

Chapter 1 Background and definition of the decision problem

Condition and aetiology

Background

When patients undergo surgical procedures under general anaesthesia (GA) it is important that the depth of anaesthesia provided by the anaesthetist is neither too light nor too deep. If the depth is too light, patients may not be fully unconscious and may be at risk of intraoperative awareness, which may lead to longer-term postoperative sequelae such as post-traumatic stress disorder (PTSD). If the depth of anaesthesia is deeper than the minimum needed to keep a patient unconscious, the patient may be at risk of anaesthetic-related morbidity, which can include postoperative nausea, vomiting and varying degrees of cognitive dysfunction. Provision of lighter anaesthesia is more likely to facilitate prompt recovery, and therefore potential health-care savings, but has to be balanced against the risks of inadequate analgesia and intraoperative awareness. A challenge facing the anaesthetist is to avoid under- or overdosing the anaesthetic, as the response to anaesthetic agents varies among individuals.

A primary concern with inadequate depth of anaesthesia is that a patient may experience intraoperative awareness, which the patient may recall postoperatively (explicit awareness) or may not subsequently recall (implicit awareness).[1] Although implicit awareness can exist without conscious recall, it may (or may not) influence patients' experience and behaviours after anaesthesia. Conscious recall may underestimate instances of awareness, as people are generally aware of more things intraoperatively than they remember.[2,3] Some authors have used 'wakefulness' as a term to describe the ability of a patient to respond to a command during GA without recollection of this in the postoperative period.[4] Examples of intraoperative events that have been classed as awareness by researchers but which were not recalled by patients when questioned after their surgery, include eye opening and gross motor responses during anaesthesia.[2,3]

Awareness symptoms and sequelae

Intraoperative awareness is commonly reported by patients as hearing noises or voices, a sensation of paralysis, anxiety, helplessness, panic and/or pain during their operation.[5,6] Some patients may report intraoperative awareness when interviewed in the recovery room, but many patients do not recall intraoperative awareness until several weeks after surgery.[7,8] Patients who experience intraoperative awareness may go on to experience problems including sleep disturbances, nightmares, flashbacks, anxiety during the day and/or fear about future anaesthetics,[5,7,9] and may be diagnosed with PTSD.[5,6,8,10] Some patients who have experienced symptoms following awareness will not seek treatment because the episode was so traumatic that they do not wish to discuss it, particularly if they have subsequently developed a phobia of medical personnel.

Studies that have followed up patients with intraoperative awareness for 2 years[11] or 5 years[8] estimated that around half of the patients with intraoperative awareness experienced PTSD. In these patients, the PTSD was not detectable immediately after surgery, but commenced several weeks afterwards, and then persisted throughout the follow-up period. The findings from these studies highlight the importance of conducting long-term follow-up of patients who might be at risk of intraoperative awareness, and emphasise that interviews to detect intraoperative awareness within the first few days of surgery may not detect either intraoperative awareness or sequelae including PTSD.

Incidence of intraoperative awareness

Intraoperative awareness is a rare event, so large studies are needed in order to accurately estimate the incidence. Large studies (with sample size of at least 10,000 patients) have not been conducted in the UK. Large studies in other countries, which have all been based on adult populations, suggest that the incidence rate for intraoperative awareness and recall is typically one to two patients per 1000, although a considerably lower incidence of 0.07 per 1000 patients was found in the largest study[14] which included over 87,000 patients, whereas a higher incidence of 4.1 per 1000 patients was found in a Chinese study[15] (*Table 1*).

Differences in incidence estimates between these studies might be explained by variations in data collection methods, the frequency and timing of interviews, or the characteristics of the patient populations and surgical procedures included.[14] The notably high incidence of intraoperative awareness in the Chinese study was considered by the authors to be possibly attributable to differences between Chinese and Western medical practices, including inappropriately light anaesthesia in the Chinese population.[15]

Risk factors for intraoperative awareness

Some groups of patients undergoing GA are at increased risk of intraoperative awareness because they cannot tolerate adequate doses of anaesthetic or because signs of inadequate anaesthesia are masked or because, owing to the nature of the patient's condition and the surgery, higher doses of anaesthetic were considered to be risky.[7,16] For example, patients undergoing procedures such as caesarean section were often given lower anaesthetic doses because of concerns over adverse fetal effects. However, most caesarean sections are now performed under regional anaesthesia (epidural or spinal) rather than under GA. Similarly, patients undergoing cardiac surgery were given lower doses because of concerns over adverse effects on their circulation. However, modern anaesthetic agents and improved treatment of haemodynamic effects have lessened the risks.[17]

Use of muscle relaxant drugs (e.g. to facilitate tracheal extubation) is an important risk factor for intraoperative awareness because it permits the use of less anaesthetic while at the same time preventing patients' movement responses that could signal inadequacy of anaesthesia to the anaesthetist, potentially allowing anaesthetic insufficiency to remain uncorrected. Some patients who have received muscle relaxants (and are therefore paralysed) have reported feelings of impending doom and death while experiencing intraoperative awareness, and have suffered long-term psychological ill health. Around half of all operations under GA involve the use of muscle relaxants.

Other risk factors for intraoperative awareness that have been identified include a high American Society of Anesthesiologists (ASA) physical status classification (indicating worse illness);[13,14] use of total intravenous anaesthesia (TIVA);[18] history of depression;[6] lack of benzodiazepine premedication;[18] and emergency surgery performed at night.[18]

TABLE 1 Estimates of the incidence of intraoperative awareness from studies with large sample sizes

Study	Country	Sample size (number of patients)	Awareness assessment method	Estimated incidence of intraoperative awareness per 1000 patients
Myles *et al*. (2000)[12]	Australia	10,811	NR	1.1
Sandin *et al*. (2000)[9]	Sweden	11,785	Modified Brice interview	1.0 without neuromuscular block, 1.8 with neuromuscular block
Sebel *et al*. (2004)[13]	USA	19,575	Modified Brice interview	1.3 overall (one or two per site)
Pollard *et al*. (2007)[14]	USA	87,361	Modified Brice interview	0.07
Xu *et al*. (2009)[15]	China	11,101	Modified Brice interview	4.1 (all patients had neuromuscular block)

NR, not reported.

Impact of intraoperative awareness

Patients who experienced severe long-term psychological or psychiatric symptoms following intraoperative awareness have reported that the symptoms caused a definite impairment of their lives.[11] For example, it may limit their ability to work, and have an adverse effect on relationships with family and friends. Patients with less severe symptoms of intraoperative awareness frequently experience a sense of dissatisfaction with their anaesthetic experience.[12] Such patients may be at risk of avoiding certain health-care procedures if they feel anxious or if they mistrust health professionals as a result of their previous experience.

Aside from the cost of managing the sequelae of intraoperative awareness, the NHS could be at risk of professional liability claims from those who have experienced intraoperative awareness.[19] However, the psychological trauma experienced by some people may be so great that they may be discouraged from reporting intraoperative awareness because they do not want to discuss it. The incidence of explicit awareness may therefore be underestimated. High-profile cases of intraoperative awareness in the media may influence public perceptions of the safety of anaesthetic procedures, which could influence how patients perceive information and services provided to them by the NHS. Some patients who have experienced intraoperative awareness have developed a fear of anaesthesia, which, in the event that further anaesthesia is required, could have implications for their acceptance or tolerance of subsequent care.

Measurement of intraoperative awareness

Basic signs of intraoperative awareness during anaesthesia include tachycardia (rapid heart rate), hypertension, sweating, lacrimation (tear production), movement/grimacing and tachypnoea (rapid breathing). Intermittent checking of these clinical signs has low sensitivity and specificity for detecting awareness.[20,21] Cases of intraoperative awareness do not always involve changes in haemodynamic parameters.[22]

Tests of intraoperative awareness may seek to identify awareness in situ, often using verbal, tactile or noxious stimulation,[1,2] and/or by interviewing the patient after surgery to establish whether or not they recall having been aware during the period of anaesthesia. During surgery the isolated forearm technique is one of the methods of detecting possible awareness in patients who have received neuromuscular blockade. A tourniquet is applied to the patient's upper arm, and inflated above systolic blood pressure to isolate the patient's forearm from the effects of the block. Movement of the arm, either spontaneously or to command, indicates wakefulness, although not necessarily explicit awareness. The isolated forearm technique has not been widely used in practice, though it has been used as a research tool in a number of studies.[21,23]

The most popular approach for postoperative assessment of awareness (as illustrated in *Table 1*) is to question patients using a version of the Brice interview.[24] The Brice interview poses five questions: (1) What was the last thing you remembered happening before you went to sleep? (2) What was the first thing you remember happening on waking? (3) Did you dream or have any other experiences while you were asleep? (4) What was the worst thing about your operation? (5) What was the next worst? In addition to an interview to detect intraoperative awareness, some studies have used a second interview (sometimes referred to as a follow-up questionnaire) to characterise the awareness episodes in more detail.[25,26] In some studies, independent expert verification of interview responses has been used to determine definite cases of awareness.[27]

Studies that report using modified versions of the Brice interview have to be interpreted with caution, as there may be considerable variation in the number of questions, their content and extent of overlap with the original Brice interview. None of the studies has looked into the psychometric properties of the interview questionnaires that it used, so their reliability and validity could be questionable. As noted above, not all cases of awareness would be detected if interviews are conducted immediately after surgery with a single interview,[9] as recall of intraoperative awareness has been reported up to 19 years after the event.[5] Other issues to consider when interpreting postoperative interviews are: repeated questioning may induce

false memories,[3,27] and three of the five Brice questions are about pre or post surgery or dreaming, which would not specifically reveal remembrance of an intraoperative awareness event.[28] The interview approach to assessing awareness with recall has also been criticised because it cannot assess awareness without recall, even though this may include implicit memory (i.e. still impact on postoperative patient experience or behaviour).

As noted above, awareness without explicit recall can be assessed using specialist interview approaches,[29] but these appear to be rarely used and have been restricted to experimental research settings. It is not known whether or not changes in behaviour as a result of implicit awareness are associated with longer-term morbidity.

Consequences of anaesthesia overdose

It is suggested that anaesthetists tend to provide higher doses of anaesthetic than may be necessary, in order to reduce the risk of intraoperative awareness.[23] Potential consequences of anaesthesia overdose include prolonged recovery time (which in severe cases may lead to potentially life-threatening cardiovascular and respiratory collapse), vomiting, headaches, dizziness and, less commonly, short- or long-term cognitive dysfunction, particularly in elderly patients.[30]

Outcomes relevant to assessing the consequences of anaesthesia overdose include postoperative nausea and vomiting (PONV) assessed using patient questionnaires or rating scales; assessments of time to recovery from anaesthesia using various measures (e.g. the time to extubation, eye opening, purposeful movement, discharge from the operating theatre or the recovery room or time to attain a specified recovery score); consumption of general anaesthetic or other drugs (such as analgesics and antinausea agents); and assessment of cognitive or neurological function.

Description of technologies under assessment

The depth of anaesthesia and likelihood of awareness may be monitored using a number of different approaches. As mentioned, potential awareness may be identified by monitoring of basic clinical signs such as blood pressure and heart rate (for more information see *Comparators*). Other techniques which have been used, but are considered historical, include spontaneous and provoked lower oesophageal sphincter contractility, forehead galvanometry and saccadic eye movements.

Electroencephalography (EEG) is the study of patient electrical brain activity to assess unconsciousness. During the last 15–20 years a number of EEG-based technologies have become commercially available for measuring depth of anaesthesia and for use in guiding anaesthetic management during surgery. Most comprise a module that collects raw EEG data via sensors placed on the patient's forehead and then processes and analyses these using a mathematical algorithm. Raw EEG signals can be difficult to interpret; therefore, many modules convert the signal to a number displayed on a monitor to indicate to the anaesthetist the depth of unconsciousness (e.g. from 0 to 99). EEG can be distinguished as spontaneous or derived from middle latency evoked potentials (auditory and visual). Evoked potentials measure the EEG responses to repetitive auditory or visual stimuli, and measure the integrity of the neural pathways that bring information from the periphery to the cortex.[21] A number of EEG-derived indexes have been devised based on different algorithms,[23] including the Bispectral Index (BIS), E-Entropy, Narcotrend, Cerebral State Index, the Patient State Index and NeuroSENSE.

In practice, EEG devices can be used in conjunction with observation of clinical signs to titrate anaesthetic dose (see the section *Comparators*). Expert opinion suggests that anaesthetists primarily use clinical signs with EEG values as an additional source of information. If there is a difference between them then the anaesthetist will usually favour the clinical signs and their judgement.

After consultation by the National Institute for Health and Care Excellence (NICE) with relevant stakeholders, three of the technologies currently available were prioritised for the current assessment: the BIS, E-Entropy and Narcotrend.

Bispectral Index (Covidien, Mansfield, MA, USA)

The BIS system, introduced in 1994, uses a sensor on the patient's forehead to measure electrical activity in the brain before using proprietary algorithmic analysis to process the EEG data and calculate a number between 0 (absence of brain electrical activity) and 100 (wide awake). This provides a measure of cerebral electrical response to increasing doses of anaesthetic drugs. The target range of BIS values during GA is 40–60, which indicates a low probability of consciousness.

Bispectral Index technology is compatible with a wide range of patient monitoring platforms through an interface for 'BIS Ready' systems [such as those manufactured by Mennen Medical Corporation, Feasterville-Trejose, PA, USA (e.g. VitaLogik series monitors); Philips Healthcare, Da Best, the Netherlands (e.g. IntelliVue series monitors); and Dräger Medical Inc., Telford, PA, USA (e.g. Infinity series monitors). This works via the BISx or BISx4 plug-in connector, which allows integration with existing anaesthesia systems.

E-Entropy module (GE Healthcare, Medical Diagnostics, Amersham, UK)

Entropy monitoring in anaesthesia has been studied over the last 10 years. E-Entropy (previously known as M-Entropy) is designed to aid the management of GA in patients by measuring the level of order or disorder in spontaneous brain and frontalis muscular activity. It uses a proprietary algorithm to process EEG and frontal electromyography data to produce two values that indicate the depth of anaesthesia. The first value, response entropy, is based on both EEG and frontal electromyography signals and provides an indication of the patient's responses to external stimuli and may signal early awakening. The second value, state entropy, is a stable parameter based on EEG and may be used to assess the hypnotic effect of anaesthetic agents on the brain. Response entropy is always higher than or equal to the state entropy value. The response entropy–state entropy difference may be used as a secondary target value when monitoring depth of anaesthesia.

More ordered signals, with less variation in the wavelength and amplitude, over time, produce high values of entropy and may indicate that the patient is awake. Regular signals, with a constant wavelength and amplitude over time, produce low or zero entropy values, indicating a low probability of recall and suppression of brain electrical activity. The response entropy scale ranges from 0 (no brain activity) to 100 (fully awake) and the state entropy scale ranges from 0 (no brain activity) to 91 (fully awake). The clinically relevant target range for entropy values is 40–60. Response entropy and state entropy values near 40 indicate a low probability of consciousness.

E-Entropy is a plug-in module that is compatible with the Ohmeda S/5 Anaesthesia monitor and S/5 Compact Anaesthesia monitor using software L-ANE03(A) and L-CANE03(A), and all subsequent software releases since 2003. The module will not work with software levels that are older than indicated. It is also compatible with GE Healthcare's latest monitoring product range (CARESCAPE Monitors B850 and B650), but is incompatible with monitors made by other manufacturers.

Narcotrend monitor (MonitorTechnik, Bad Bramstedt, Germany)

The Narcotrend monitor automatically analyses the raw EEG using spectral analysis to produce a number of parameters. Multivariate statistical methods using proprietary pattern recognition algorithms are then applied to these parameters to provide an automatically classified EEG. The basis for the development of the automatic classification functions were visually classified EEG. The EEG visual classification scale is from stage A (awake) to stage F (very deep hypnosis), with stage E indicating the appropriate depth of anaesthesia for surgery. As a refinement to the A–F scale, an EEG index (100 = awake, 0 = very deep hypnosis) is also calculated.

The Narcotrend-Compact M is a stand-alone monitor that stores recorded EEG data on its hard disk and can send raw and processed EEG data in real time to other anaesthesia monitors. Data can also be saved to a USB flash drive for processing and evaluation of Narcotrend EEG recordings on a remote PC using the software NarcoWin. The Narcotrend algorithms are revised continually.

Subgroups of patients

Unsuitable patient populations include those undergoing specific surgical procedures in which the sensors would impede access to the surgical site, and therefore certain ENT, ophthalmic and neurosurgical procedures may be unsuitable for EEG monitoring. In neonates the immature EEG has resulted in inconsistent linkages between anaesthetic dosing and displayed BIS values, and an inability to demonstrate a titration potential for BIS-guided anaesthesia care. The manufacturer of BIS recommends that BIS values should be interpreted cautiously in patients with known neurological disorders and patients taking psychoactive medications. E-Entropy is validated only for patients over the age of 2 years; it is not for patients undergoing procedural or conscious sedation, and seizure activity may cause interference. In addition, E-Entropy readings may be inconsistent when monitoring patients with neurological disorders or patients on psychoactive medication. Limited information is available for subgroups of patients for whom Narcotrend may not be suitable, although Narcotrend values should be interpreted cautiously in patients with a history of central nervous system diseases.

Artefacts

All EEG monitoring is subject to contamination by artefacts generated either by the patient (e.g. by eye movements, muscle activity) or from external sources (poor skin contact, mains or power line interference, electrocautery). With the BIS system most artefacts present as elevated BIS values and the recommended strategy from the manufacturer for an unexpected elevated BIS value is prompt patient assessment, confirmation of anaesthetic dosing and delivery, and consideration of artefacts. Narcotrend is equipped with artefact detection algorithms to exclude segments contaminated with artefact from further analysis. If too many artefacts are detected, no classification result will be output and only raw EEG will be visible onscreen.

Current usage in the UK

Expert opinion suggests that there is low use of EEG in practice to monitor depth of anaesthesia. Current penetration of BIS technology in UK operating theatres is still relatively low but, as most anaesthetic monitors used in the UK could be compatible with the BIS module, BIS technology could be available in the majority of UK operating theatres. The manufacturers of E-Entropy in their submission to NICE estimate that nearly 45% of UK theatres would be ready and compatible with E-Entropy and ‘believe our theatre installed base to be around 60 to 65% of UK theatres’. No data are available on the provision or diffusion of Narcotrend in the UK. (Commercial-in-confidence information removed.)

Training

It appears that little additional training in the use of these technologies is needed. The manufacturer states that no specific additional training is required to use the BIS monitoring system (although expert clinical opinion disputes this). Instructions for use are provided with both the BIS device (stand-alone or module) as well as the BIS sensors and are regarded as sufficient guidance by the manufacturer for safe and effective use. Additional educational resources are provided by the manufacturer if necessary, such as simulation devices and online multimedia courses. For E-Entropy, 30 minutes of introductory training is suggested for health-care staff before use, with particular attention being paid to sensor application. A 1-day visit from staff to give a lecture and to demonstrate the use of Narcotrend in the operating theatre is judged sufficient training by the manufacturer for the majority of Narcotrend users.

Comparators

A number of clinical signs that are routinely monitored during anaesthesia can be used to assess potential awareness. Prior to induction of anaesthesia a variety of monitoring devices may be attached to the patient, including a pulse oximeter (to measure oxygen levels); a non-invasive blood pressure monitor; an electrocardiograph (to measure heart rate); and a capnograph (to measure inhaled and exhaled carbon dioxide concentration). Devices are also used to measure airway pressure and the patient's temperature. Other markers of awareness that are monitored include movement, lacrimation and sweating.

End-tidal anaesthetic gas concentrations (ETACs) may be used to assess the concentration of volatile (inhaled) anaesthetic in a patient, expressed as a percentage. ETAC can be used to calculate the minimum alveolar concentration (MAC), which is the minimum concentration of anaesthetic agent in the lungs at one atmosphere pressure that is required to prevent movement in 50% of individuals when exposed to a standard painful stimulus. MAC provides a measure of the potency for comparison between different inhaled general anaesthetics (see *Care pathways*), and anaesthesia can be titrated to keep within a certain MAC range.

Of all the signs and variables, the key things to observe are ETAC (where inhaled anaesthetics have been used), blood pressure and heart rate. However, in practice, the combination of signs that are used is likely to vary.[31]

Care pathways

In UK health-care settings, GA is usually administered in an anaesthetic room[32] (sometimes referred to as the induction room), following which the patient is transferred to the operating theatre. Monitoring of clinical signs always commences prior to administration of GA, and continues until surgery is complete and the patient is moved from the theatre to the recovery room (also referred to as the postanaesthesia care unit, PACU), or to intensive care or a high-dependency unit if applicable. Supplementary monitoring devices such as EEG-based technologies may also be attached during anaesthesia induction and continued until surgery is complete, anaesthesia has ceased and the patient has entered the recovery phase.

General anaesthetics are generally classified as intravenous (i.v.) or inhalational. Propofol is a commonly used i.v. anaesthetic and can be used for induction and/or maintenance of anaesthesia. Use of an i.v. anaesthetic for induction and maintenance is sometimes referred to as TIVA. Ketamine is also available for induction and maintenance of anaesthesia, but is rarely used. Inhaled anaesthetics are classified as volatile agents or nitrous oxide. The latter is used for maintenance of anaesthesia in combination with i.v. or volatile agents, in a concentration of 50–66% in oxygen[33] (it can also be used for analgesia). Volatile anaesthetics can be used for induction and maintenance of anaesthesia, and also following induction with an i.v. anaesthetic. Volatile agents include isoflurane, desflurane and sevoflurane. Isoflurane is the preferred inhalational anaesthetic for use in obstetrics.[33] Desflurane is rapid acting and has about one-fifth of the potency of isoflurane. It is not recommended for induction of GA. Sevoflurane is also rapid acting, is more potent than desflurane and can be used for induction of anaesthesia. The MACs of desflurane, sevoflurane and isoflurane are 6.0, 1.8 and 1.2 for people of ages 30–60 years, and 5.2, 1.5 and 1.0 for people older than 65 years respectively.[34] MAC would be higher in children and young adults.

Summary of the decision problem

As has been described, the purpose of anaesthesia monitoring is to ensure adequate sedation of the patient under GA. If anaesthesia is too deep the patient may be at risk of adverse effects, such as a prolonged recovery time. However, if anaesthesia is not deep enough patients may be more likely to experience awareness of their surroundings, and this may have short- and long-term psychological

effects, including depression and anxiety. Optimum anaesthetic dosing may also potentially lead to drug cost-savings.

Currently, anaesthetists generally use clinical observation of vital signs and other markers to assess unconsciousness and the possibility of awareness. However, clinical observation alone may not be a reliable surrogate marker of anaesthetic depth. As an alternative, technologies have been developed using EEG to measure and interpret patient electrical brain activity to provide a measure of unconsciousness. Three such technologies, prioritised for assessment, are BIS, E-Entropy and Narcotrend.

The aim of this report, therefore, is to assess the clinical effectiveness and cost-effectiveness of BIS, E-Entropy and Narcotrend to monitor the depth of anaesthesia in surgical patients undergoing GA.

Chapter 2 Assessment methods

Systematic review of patient outcomes

The purpose of this section is to describe the methods used in the systematic review of patient outcomes associated with depth of anaesthesia monitoring. These methods were stated a priori in the published research protocol. An extract of the protocol outlining the methods is given in *Appendix 1*.

Identification of studies

A search strategy was developed for MEDLINE and pilot tested by an experienced information scientist. The MEDLINE strategy (see *Appendix 2*) was adapted where necessary to the specific vocabulary and rules of other electronic bibliographic databases. Searches were run in the following databases: Ovid MEDLINE; Ovid EMBASE; Centre for Reviews and Dissemination (CRD); The Cochrane Library (Cochrane Database of Systematic Reviews (CDSR); Cochrane Central Register of Controlled Trials (CENTRAL); Database of Abstracts of Reviews of Effects (DARE); and Health Technology Assessment (HTA) database. For E-Entropy and Narcotrend the electronic searches were conducted from 1995 (around the time of the introduction of EEG technologies) to November 2011 (with an update search performed in February 2012).

Scoping searches indicated that the volume of evidence for BIS was relatively larger than for Narcotrend and E-Entropy and it would be beyond the resources available to include all of the BIS studies in the systematic review. During preliminary scoping searches we identified a recent Cochrane systematic review of BIS[34] that had similar study eligibility criteria to our review (with the exception that it did not include studies of children). We therefore based our review of BIS on a Cochrane systematic review,[34] which contained 31 randomised controlled trials (RCTs) of BIS. The most recent date of literature searching in the Cochrane review was May 2009. We therefore searched from the beginning of 2009 to November 2011 for studies of BIS (and then updated in February 2012) (see *Method of data synthesis* for further information about how results from the Cochrane review are integrated into the current review).

In addition to the searches of electronic bibliographic databases, the following sources were searched to identify potentially relevant studies:

- contact with experts in the field (identified by NICE as part of the consultation process)
- bibliographic lists of potentially relevant studies on BIS, E-Entropy and Narcotrend as supplied by the device manufacturers (via NICE)
- reference lists of included studies
- databases of research in progress, searched on 7 December 2011: UK Clinical Research Network (UKCRN); controlled-trials.com; ClinicalTrials.gov; NIHR-Clinical Research Network Portfolio; WHO ICTRP (International Clinical Trials Registry Platform).

The titles and abstracts of studies identified from these searches were imported into a Reference Manager bibliographic database. All titles and abstracts in this database were assessed against the inclusion/exclusion criteria (see *Inclusion/exclusion criteria*). Bibliographic records that clearly did not meet any of the inclusion criteria, or met at least one of the exclusion criteria, were excluded from further consideration. For each bibliographic record that met all of the inclusion criteria, or was of unclear relevance, a full-text version was obtained and assessed against the inclusion/exclusion criteria. Full-text records that clearly did not meet all of the inclusion criteria were excluded from further consideration, and the reasons for their exclusion were noted.

Both the title and abstract selection step and the full-text selection step were conducted independently by two reviewers. After screening the bibliographic records, the reviewers compared their selection results. All initial differences in opinion were resolved through discussion, without needing to involve a third reviewer.

Inclusion/exclusion criteria

The inclusion/exclusion criteria for this report were based on the scope of the appraisal set by NICE. Only articles published in the English language were included. Abstracts that had no corresponding full-text record (e.g. conference abstracts) were excluded unless they met two criteria: they were published in 2010 or later; and they provided sufficient details to allow appraisal of the methodology and the assessment of results to be undertaken.

The inclusion/exclusion criteria were provided to each reviewer as a standard list against which each title/abstract or full-text record could be readily assessed (see *Appendix 3*). In addition to the language and publication type restrictions, the following selection criteria were applied:

Population

Studies were included if they included patients who received GA for surgery, including adults and children (over the age of 2 years) in whom the technology is licensed. Studies involving patients receiving sedation in intensive care or high-dependency units, studies carried out in healthy volunteers and studies of non-surgical anaesthesia were excluded.

Diagnostic technologies

The diagnostic technologies included were E-Entropy, BIS and Narcotrend.

Comparators

Comparators included standard clinical monitoring for monitoring delivery of anaesthesia, including one or more of the following clinical markers: end-tidal anaesthetic gas concentrations (for inhaled anaesthesia); pulse measurement; heart rhythm; blood pressure; lacrimation; and sweating.

Outcomes

Studies were included if at least one of the following outcomes was reported:

- probability of intraoperative awareness
- patient distress and other sequelae resulting from intraoperative awareness
- recovery status (e.g. Aldrete scoring system)
- time to emergence from anaesthesia
- time to extubation
- time to discharge from the recovery room
- consumption of anaesthetic agents
- morbidity and mortality including postoperative cognitive dysfunction (POCD) from anaesthetic agents, pain-relieving drugs, antibiotics, antisickness drugs and muscle relaxants.

Study design

The review was limited to prospective controlled trials (once studies had been included in the systematic review, priority was given to RCTs unless no RCT evidence for relevant parameters was available in which case non-RCT data would be considered). Systematic reviews that met the inclusion criteria were retrieved in order to check their reference lists for potentially relevant studies but were not themselves evaluated (except for the Cochrane systematic review of BIS technologies,[34] which was considered in more detail when conducting data synthesis: see *Data extraction and critical appraisal methods*).

Data extraction and critical appraisal methods

A standardised data extraction and quality appraisal template (see *Appendix 5*) was used to extract information on the relevant study characteristics for assessing the impact of the interventions on the

outcomes listed above (see *Inclusion/exclusion criteria*) and for assessing study quality. Study quality assessment criteria included: Cochrane Collaboration Risk of Bias criteria,[35] as specified in the review protocol; methods of data analysis, including the statistical tests used and whether or not studies were powered statistically to detect differences in outcomes between intervention and comparator groups; participant attrition; generalisability of the studies; and conflict of interests. Criteria for the critical appraisal of non-randomised and observational studies were specified in the protocol but were not required, as all the included studies were RCTs (see *Results of systematic review of patient outcomes*).

The data extraction and critical appraisal template was completed for each study included in the systematic review by one reviewer and was checked by a second reviewer. All initial discrepancies between the reviewers were resolved by discussion, without needing to involve a third reviewer.

Method of data synthesis

Analyses of the three monitoring devices are presented in respective separate subsections of this report (see *Results of systematic review of patient outcomes*). For each device a narrative synthesis was conducted, with characteristics of the included trials, and their outcomes, described in the text and tabulated.

As stated, the analysis of BIS was based on trials included in an existing Cochrane review of BIS,[34] and supplemented by trials identified and included in the current systematic review. For each BIS outcome measure we present a narrative synthesis of the studies identified in the current systematic review, in addition to the pooled meta-analysis estimates from the Cochrane review. Where possible, we have updated the Cochrane meta-analyses for BIS with trials identified in the current review. However, the Cochrane BIS review included only trials of adults, and it was not considered appropriate to combine trials of children identified in our searches with the existing adult trials. We used Cochrane Review Manager 5.1.6 (The Cochrane Collaboration, The Nordic Cochrane Centre, Copenhagen, Denmark) to conduct the meta-analyses.

Systematic review of cost-effectiveness

Identification of studies

A comprehensive search strategy was developed, tested and refined by an experienced information scientist to identify studies of the cost-effectiveness of depth of anaesthesia monitoring. The MEDLINE search strategy is provided in *Appendix 2*.

A total of six electronic resources were searched. Searches were from database inception to November 2011 (an update search was done in February 2012). The following electronic databases were searched: MEDLINE (Ovid); MEDLINE In-Process & Other Non-Indexed Citations (MEIP); EMBASE; The Cochrane Library including CENTRAL and CDSR; CRD including HTA database, DARE and National Health Service Economic Evaluation Database (NHS EED); and EconLit. Bibliographies of retrieved articles were checked for any additional references, and the expert advisory group was contacted to identify additional published and unpublished studies.

Inclusion/exclusion criteria

Studies were selected for inclusion in the systematic review of cost-effectiveness through a two-stage process using predefined and explicit criteria. The full literature search results were independently screened by two reviewers to identify all citations that possibly met the inclusion criteria (*Table 2*).

Full papers of relevant studies were retrieved and assessed independently by two reviewers using a standardised eligibility form, using the same inclusion/exclusion criteria, except that only studies with standard treatment specified as 'no depth of anaesthesia monitor' were included. Studies reporting other outcomes (one or more of probability of intraoperative awareness, consumption of anaesthetic agents,

TABLE 2 Inclusion/exclusion criteria for screening titles and abstracts

Criterion	Eligibility
Population	Patients receiving general anaesthetic for surgery, including adults and children in whom the technology is licensed
Interventions	Any depth of anaesthesia monitoring device
Design	Economic evaluation (cost-consequence analysis, cost-effectiveness analysis, cost–utility analysis, cost–benefit analysis)
Outcomes	Cost per patient, cost per episode of intraoperative awareness or cost per QALY
Other	Exclude non-English language Exclude conference abstracts

QALY, quality-adjusted life-year.

postoperative morbidity or mortality, HRQoL) were not included in the review, but were retained to inform the development and population of the decision-analytic model.

Data extraction and critical appraisal methods

Data were extracted by one reviewer using a standard data extraction form (see *Appendix 6*) and checked by a second reviewer. At each stage, any disagreements between reviewers were resolved by consensus.

The quality of the included economic evaluations was assessed using a critical appraisal checklist based on that proposed by Drummond and colleagues[36] and Philips and colleagues[37] (see *Appendix 6*).

Method of data synthesis

Studies of cost-effectiveness were synthesised through a narrative review with tabulation of results of included studies, where appropriate.

Economic evaluation

We developed a decision-analytic model to assess the cost-effectiveness of depth of anaesthesia monitoring, compared with standard clinical monitoring, adopting the perspective of the UK NHS. Separate analyses are presented for each of the included technologies, compared with standard clinical monitoring – the included technologies are not compared with each other.

The scope issued by NICE identified a number of health outcomes, including morbidity and mortality from anaesthetic agents, pain-relieving drugs, antibiotics, antisickness drugs and muscle relaxants, as well as patient discomfort and sequelae resulting from intraoperative awareness. The model was developed to allow for the inclusion of these outcomes, if suitable data on baseline values and the effect of depth of anaesthesia monitoring on these outcomes was identified in our systematic review of patient outcomes. Outcomes in the model are expressed as quality-adjusted life-years (QALYs). The model evaluates costs from the perspective of the NHS and Personal Social Services. Costs are expressed in UK sterling (pounds, £) at a 2011 price base. Both costs and outcomes are discounted using a 3.5% annual discount rate, in line with current guidance.[38,39]

Analytical methods

Base case

A base-case analysis is presented for a general surgical population (at average risk of intraoperative awareness) and for a population assumed to be at high risk of intraoperative awareness. In the general

surgical population, additional potential benefits (in terms of reductions in anaesthetic dose and reduction in anaesthetic-related complications) that may be associated with depth of anaesthesia monitoring are included in the base-case analysis, based on data from our systematic review of patient outcomes. Where data from the systematic review of patient outcomes were insufficiently robust, or where no evidence specific to the technology being considered was identified, data derived for other included technologies were used to populate the model.

Deterministic sensitivity analysis

Uncertainties around the probability, resource use and cost estimates, as well as effect parameters derived in the systematic review of patient outcomes, were investigated by applying ranges around the point estimates used in the base-case analysis. Where possible the ranges used in the deterministic sensitivity analyses were based on 95% confidence intervals (CIs) estimated for each input parameter. The method adopted was univariate sensitivity analysis – that is, varying one parameter at a time, leaving all other variables unchanged. This is to highlight the impact, if any, of each selected parameter alone on the cost-effectiveness results.

Scenario analysis

Scenario analysis was used to address uncertainty associated with the choice of data source adopted for parameter values in the base case and for variables omitted from the model.

Commercial-in-confidence information

This report contains reference to confidential information provided as part of the NICE appraisal process. This information has been removed from the report, and the results, discussions and conclusions of the report do not include the confidential information. These sections are clearly marked in the report.

Chapter 3 Assessment results

Results of systematic review of patient outcomes

Quantity and quality of research available

In total, 776 bibliographic records were identified from electronic bibliographic databases and reference lists provided by the manufacturers of the BIS, E-Entropy and Narcotrend monitors (*Figure 1*).

Of these 776 records, 741 were excluded, based on information provided in the title and/or abstract. Full-text publications were obtained and assessed for the remaining 35 records, of which 10 were found on further scrutiny to not meet the inclusion criteria. Reasons for excluding the 10 full-text records were that they were not RCTs (five publications), they included an inappropriate or unclear comparator group (four publications) and, in one case, the publication was retracted by the journal (see *Appendix 4*).

The remaining 25 full-text publications reported 25 studies, which were eligible for inclusion in the systematic review. Four of the 25 RCTs were identified by our update searches in February 2012, all

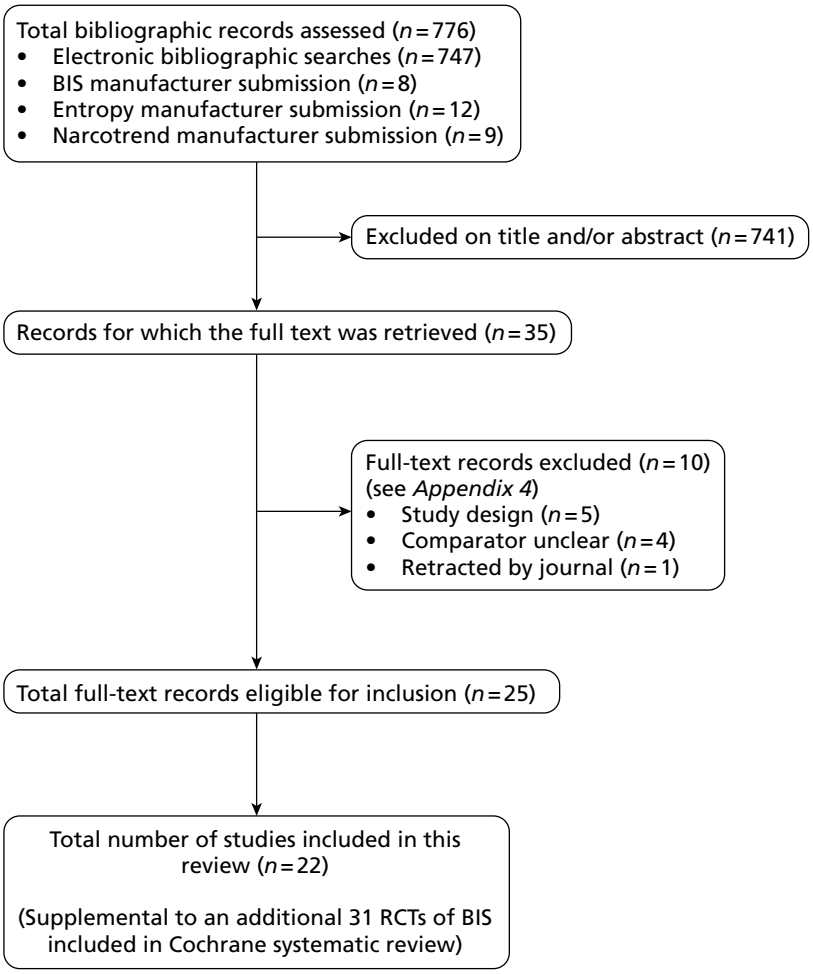


FIGURE 1 PRISMA (Preferred Reporting Items for Systematic Reviews and Meta Analyses) flow chart showing the study selection process for bibliographic records (excluding those already identified in a Cochrane systematic review of BIS studies).

evaluating BIS. Because of finite time and resources we prioritised the largest of these for inclusion in the review (a trial of around 5000 patients, specifically designed to assess intraoperative awareness[40]). The other three were smaller trials (80 patients,[41] 40 patients,[42] and 20 patients[43] respectively) and their inclusion in the review was unlikely to change the findings. In summary, a total of 22 RCTs were included in this systematic review.

The 22 included studies were all RCTs that included study arms for at least one relevant technology (BIS, E-Entropy or Narcotrend) and a comparator that reflected standard clinical monitoring.

The 22 included studies were two- or three-arm RCTs that compared the following technologies against standard clinical monitoring:

- BIS alone: 11 studies[40,44–53]
- E-Entropy alone: five studies[54–58]
- Narcotrend alone: two studies[59,60]
- BIS and E-Entropy: two studies[61,62]
- BIS and Narcotrend: two studies.[63,64]

These 22 studies provide 15 comparisons of BIS against standard clinical monitoring, seven comparisons of E-Entropy against standard monitoring and four comparisons of Narcotrend against standard monitoring (*Table 3*).

The 15 comparisons of BIS against standard monitoring supplement the Cochrane review,[34] which included 31 RCTs of BIS against standard clinical practice.[27,61,63–91]

Note that only 11 of the 15 BIS studies in the current review are presented in the following BIS subsections for the following reasons:

- One of the trials of BIS and E-Entropy compared with standard clinical monitoring was included in the Cochrane BIS review,[61] and therefore is described only within the E-Entropy subsections of this report (i.e. for the comparison of E-Entropy with standard clinical monitoring).
- Two of the trials of BIS and Narcotrend compared with standard clinical monitoring were included in the Cochrane BIS review,[63,64] and are therefore described only within the Narcotrend subsections of this report (i.e. for the comparison of Narcotrend with standard clinical monitoring).
- One of the BIS publications identified in the current systematic review (Leslie and colleagues[50]) is a long-term follow-up publication of one of the trials (the B-Aware trial by Myles and colleagues[79]) included in the Cochrane review.[73] We report the long-term results of this trial in this report (see *Assessment of outcomes: Bispectral Index*) but details of the characteristics of the trial (including the risk of bias judgement) can be found in the Cochrane review itself.

Risk of bias in Bispectral Index trials

Table 4 reports a summary of the risk of bias judgements for the trials of BIS included in this systematic review (NB. The risk of bias judgements for the 31 RCTs in the Cochrane BIS review are not tabulated in this report, but are summarised in the text below).

In many cases the risk of bias in the trials was unclear because of limitations in reporting of methodological details. Uncertainty was greatest in relation to concealment of the random allocation process, where details were unclear in all but two trials. In the Cochrane systematic review of BIS, 12 of the 31 (39%) trials were considered to have adequately concealed random allocation, with most of the remainder judged as unclear.

TABLE 3 Distribution of diagnostic technologies across the trials included in this review

Author	BIS	Entropy	Narcotrend
Aime *et al.*[61]	✓	✓	
Avidan *et al.*[44]	✓		
Bannister *et al.*[45]	✓		
Bhardwaj and Yaddanapudi[46]	✓		
Chan *et al.*[47]	✓		
Choi *et al.*[54]		✓	
Ellerkmann *et al.*[62]	✓	✓	
Gruenewald *et al.*[55]		✓	
Kamal *et al.*[48]	✓		
Kerssens *et al.*[49]	✓		
Kreuer *et al.*[63]	✓		✓
Kreuer *et al.*[64]	✓		✓
Lai *et al.*[59]			✓
Leslie *et al.*[50]	✓		
Liao *et al.*[51]	✓		
Messieha *et al.*[52]	✓		
Messieha *et al.*[53]	✓		
Rundshagen *et al.*[60]			✓
Talawar *et al.*[56]		✓	
Vakkuri *et al.*[57]		✓	
Wu *et al.*[58]		✓	
Zhang *et al.*[40]	✓		

Details of blinding of participants and trial personnel to trial arm were also generally unclear, as was the case of blinding of outcome assessors. In the Cochrane BIS review[34] just over half of the studies were judged to be of a low risk of bias because of blinding of outcome assessors (17/31; 55%).

Random sequence generation was one of the domains where risk of bias was lowest. However, although all studies were reported to be randomised trials, in six trials (46%) the method of randomisation was not given. In the Cochrane systematic review of BIS[34] just under half of the included studies (15/31; 48%) were judged to be of a low risk of bias because of adequate random sequence generation. Most of the remainder were unclear because of lack of details given in trial publications.

In general, there appeared to be low risk of bias in terms of selective reporting of outcomes, as could be judged from the details reported in the trial publications. This was also the case in the Cochrane BIS review.[34] Bias associated with incomplete outcome data was judged low in around half of the trials (and in just under half in the Cochrane BIS review,[34] 15/31; 48%). In the remainder it was unclear, and in one trial it was judged to be high because of an imbalance in the percentage of patients excluded from the analysis between trial arms.[62] In general, it was not considered that risk of other forms of bias were present. However, in one trial the risk was considered high because of the study being funded in part by the BIS module manufacturer.[45]

TABLE 4 Summary of risk of bias – BIS

Study	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting
Avidan *et al.*[44]	Low	Low	Unclear	Low	Low	Low
Bannister *et al.*[45]	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Bhardwaj and Yaddanapudi[46]	Low	Unclear	Unclear	Unclear	Low	Low
Chan *et al.*[47]	Unclear	Unclear	Unclear	Unclear	Unclear	Unclear
Ellerkmann *et al.*[62]	Low	Unclear	Unclear	Unclear	High	Low
Kamal *et al.*[48]	Unclear	Unclear	Unclear	Unclear	Low	Low
Kerssens *et al.*[49]	Low	Unclear	Unclear	Low	Unclear	Unclear
Liao *et al.*[51]	Unclear	Unclear	Unclear	Unclear	Low	Low
Messieha *et al.*[52]	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Messieha *et al.*[53]	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Zhang *et al.*[40]	Low	Unclear	Low	Low	Unclear	Low

The trials varied in terms of their sample sizes, from as low as 20 patients to over 6000. There were seven (46%)[45,46,48,52,53,59,62] that included fewer than 100 patients and five (33%)[48,51,61,63,64] that had between 101 and 200 patients. One trial included 921 patients,[47] another included 5309[40] and another, the largest, included 6041 patients.[44] In the Cochrane BIS review[34] the majority of trials included fewer than 100 patients (21/31; 68%). Seven trials (23%) included between 101 and 200 patients. Another study – the B-Unaware trial by Avidan and colleagues 2008 – included 1941 patients,[27] and the largest included 2463 patients.[79] (NB. The Cochrane BIS review appears to count two publications relating to this single trial as two separate studies. One publication reports the main trial results,[79] and a second publication focuses on recovery outcomes from the trial.[74])

Six (55%)[40,44,46,49,51,62] of the 11 BIS trials reported a statistical sample size calculation based on a nominated primary outcome, although one of these trials reported that the number of patients chosen was arbitrary rather than being based on a statistical calculation.[49] The Cochrane BIS review[34] did not comment on sample size power calculations in the studies included.

Six (55%)[40,44,46,48,49,62] of the 11 BIS trials reported patient attrition. The attrition rate varied from 1.5%[40] to 15% [49] of the total number of patients enrolled. Most of the studies reported the reasons for attrition, generally comprising exclusions from the analyses as a result of deviations from the study protocol. Given the nature of the procedure and the relatively short follow-up duration, loss to follow-up was rarely reported. In five (45%) studies it was reported by the authors that there was no attrition, or there did not appear to be any attrition.[45,47,51–53] Whether or not an intention-to-treat (ITT) analysis had been employed was rarely mentioned in the trial reports. Only two trials mentioned that patients had been analysed according to the procedure to which they had been randomised.[44,46]

Five of the BIS trials disclosed information about funding.[40,44,45,49,51] Funding for two of these trials was provided by medical research funding organisations and/or hospital departmental grants.[44,51] The other three trials reported varying financial associations with BIS manufacturers.[40,44,49] The trial by Bannister and colleagues[45] stated that Aspect Medical Systems supplied the BIS monitor, and that one author was employed by Aspect Medical Systems and another author was a paid consultant to Aspect Medical Systems. This funding therefore represents a conflict of interest. The trial by Kerssens and colleagues[49] reported that Aspect Medical Systems did not financially support the study, but that the lead author had

received an educational grant in support of her salary from Aspect Medical Systems, and one co-author was a paid consultant to Aspect Medical Systems. In the trial by Zhang and colleagues,[40] Aspect Medical Systems provided BIS electrodes, but no further detail on funding was given. None of the other BIS trials stated or appeared to have any major conflicts of interest. The Cochrane BIS review[34] did not report funding details of the included trials, or whether or not any of the trials had conflicts of interests.

Risk of bias in E-Entropy trials

Table 5 reports a summary of the risk of bias judgements for the trials of E-Entropy included in this systematic review.

The risk of bias in the E-Entropy trials was unclear in many cases because of limitations in the reporting of methodological details. Uncertainty was greatest concerning allocation concealment and the blinding of participants and personnel, which were not adequately reported in any of the seven E-Entropy trials.

Risk of bias because of random sequence generation was considered low in four of the trials, in which sequences were generated either by computer[56,57,61] or by drawing lots.[62] Risk of bias because of random sequence generation was deemed unclear in the remaining three trials, which provided no information on the method of sequence generation.

The method of allocation concealment was considered to pose unclear risk of bias in all seven of the trials, either because no relevant information was reported[58,61,62] or sealed envelopes were used for allocation codes, but it was not stated whether or not the envelopes were opaque.[54–57]

Anaesthetists who administered anaesthesia according to standard clinical monitoring were blinded to E-Entropy values. However, none of the studies unequivocally reported that study participants and personnel were blinded to the study groups. The risk of bias because of inadequate blinding in each of the E-Entropy studies was therefore judged to be unclear.

In three of the seven E-Entropy trials, the risk of attrition bias because of analysis of incomplete outcome data was considered low, as exclusions were a minor proportion of the sample size,[54] or were generally balanced between groups with generally similar reasons given,[61] or the analysis was conducted by ITT with no discernible attrition.[56] Two trials were considered at high risk of attrition bias because the rate of attrition was ≥10% in at least one of the study arms, and not balanced across the arms.[58,62] The remaining two trials were judged to have unclear risk of attrition bias because of incomplete outcome data, either because attrition was not reported at all[55] or it was not reported separately by study arm.[57]

TABLE 5 Summary of risk of bias – E-Entropy

Study	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting
Aime *et al.*[61]	Low	Unclear	Unclear	Unclear	Low	Low
Choi *et al.*[54]	Unclear	Unclear	Unclear	Unclear	Low	Low
Ellerkmann *et al.*[62]	Low	Unclear	Unclear	Unclear	High	Low
Gruenewald *et al.*[55]	Unclear	Unclear	Unclear	Unclear	Unclear	Low
Talawar *et al.*[56]	Low	Unclear	Unclear	Low	Low	Low
Vakkuri *et al.*[57]	Low	Unclear	Unclear	Unclear	Unclear	Unclear
Wu *et al.*[58]	Unclear	Unclear	Unclear	Unclear	Unclear	Low

The risk of bias because of selective reporting of outcomes was judged to be low for six of the seven E-Entropy trials, as there was no indication within the primary publications that more outcomes had been measured than were subsequently reported (in general, there was concordance between the outcomes specified in the methods and results sections of the publications). In the remaining trial,[57] risk of bias from selective reporting was considered unclear as several outcomes were reported narratively without any supporting quantitative data that could be checked by the reviewers.

One of the E-Entropy trials[56] reported that no external funding was used, and one trial[62] did not report whether or how the work was funded. Two trials were funded by non-commercial sponsors, which were a university[54] and a national science organisation.[58] The remaining three E-Entropy trials were supported by the E-Entropy device manufacturer (GE Healthcare; formerly Datex-Ohmeda), either through provision of equipment alone[55,61] or through provision of equipment, funding and also technical support.[57] The authors of this latter trial[57] included a research engineer, research scientist and chief scientist of the device manufacturer and two medical advisors to the device manufacturer. These three trials that involved support from the device manufacturer could be at risk of bias because of conflict of interests. The study that involved the most extensive links with the manufacturer[57] was deemed by the reviewers to be at high risk of bias because of a high likelihood of conflicting interests. In the four E-Entropy trials that were not supported by the E-Entropy device manufacturer, three did not refer to conflict of interests[54,58,62] and one stated that no conflicts were disclosed.[56]

The seven E-Entropy studies were published during 2005 to 2010 and ranged in their total sample size from 50 to 335 patients.[54–58,61,62] Five of the trials involved a two-arm comparison of E-Entropy against standard clinical monitoring.[54–58] One trial involved a three-arm comparison of BIS, E-Entropy and standard clinical monitoring.[61] The remaining trial was a three-arm comparison of E-Entropy, E-Entropy and BIS, and standard practice.[62] The number of patients randomised per arm ranged from 25 to 40 in six trials.[54–56,58,61,62] In the seventh (largest) trial, only the number per arm after attrition (160 patients) was reported.[57]

Only one of the E-Entropy trials did not report a sample size calculation.[58] Three trials calculated the sample size needed to detect a specified percentage difference in anaesthetic consumption for sevoflurane[54,61] or propofol.[62] The remaining three trials calculated the sample size needed to detect differences in patient recovery from anaesthesia, namely the time to eye opening,[55] time to awakening (not defined)[56] or the time to response to a verbal command.[57]

Overall, the range of attrition in the trials was 0–11% of the total population per trial, or 0–17% of the population per study arm. Attrition appeared to be zero in one trial,[56] and was not reported in one trial.[55] Among the remaining five trials, reasons for attrition were clearly reported separately by study group in two trials;[54,61] were reported only for aggregated data across study groups in one trial;[57] were vaguely specified as resulting from 'technical problems' in one trial;[54] and were not specified in the remaining trial.[58]

An analysis by ITT was explicitly reported in one trial and appears valid as no attrition was discernible in the study report.[56] Another trial[55] did not explicitly mention ITT analysis but appeared to have used an ITT approach, as it was stated that all patients were included in the final analysis, although attrition was not reported. A third trial[54] analysed nearly all the randomised patients [only 1/40 per group (2.5%) were excluded], which may be considered close to an ITT approach. The remaining four trials[57,58,61,62] did not follow the ITT principle as their analyses excluded from 4% to 17% of the randomised patients per study arm.

Risk of bias in Narcotrend trials

Table 6 reports a summary of the risk of bias judgements for the trials of Narcotrend included in this systematic review.

TABLE 6 Summary of risk of bias – Narcotrend

Study	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting
Kreuer *et al.*[63]	Low	Unclear	Unclear	Low	Low	Low
Kreuer *et al.*[64]	Low	Unclear	Unclear	Low	Low	Low
Lai *et al.*[59]	Unclear	Unclear	Unclear	Unclear	Low	Low
Rundshagen *et al.*[60]	Unclear	Unclear	Unclear	Unclear	Unclear	Low

In many cases the risk of bias in the trials was unclear because of limitations in reporting of methodological details. Uncertainty was greatest in relation to concealment of the random allocation process and blinding of participants and personnel, where details were unclear in all four trials.[59,60,63,64]

Both the method of random sequence generation and blinding of outcome assessment were unclear in two trials,[59,60] with low risk of bias for these domains in the other two trials.[63,64] Risk of bias because of incomplete outcome data was low in all but one trial in which details were unclear.[60]

In general, there appeared to be low risk of bias in terms of selective reporting of outcomes, as could be judged from the details reported in the trial publications. Other sources of bias were reported for only one study where the paper was translated from Chinese to English prior to publication and it is unclear whether or not any checks were made to ensure fidelity of the published version to the original work.[59]

The trials were conducted between 2003 and 2010 and trial sizes ranged from 120 patients[63,64] to 48 patients[60] and 40 patients.[59] All but the smallest study reported the use of a sample size calculation. No attrition was reported for three trials[59,63,64] and these studies conducted ITT analyses. The fourth trial[60] reported attrition although not by study group, and analyses did not include all patients who started but it is unclear whether or not attrition happened pre or post randomisation. All four trials[59,60,63,64] did not report any conflict of interest. Two studies[63,64] stated that the study was solely supported by departmental funding, one[59] did not report any details of the sponsor and the fourth[60] reported that the study was supported by a pharmaceutical company and a university institutional research grant.

Characteristics of included studies: Bispectral Index

The following subsections describe the key characteristics of the BIS trials included in this systematic review. The characteristics of the 31 trials included in the Cochrane BIS review are summarised alongside.

Study populations

Five of the 11 BIS trials were conducted in children, with mean ages of between 4 and 6 years, and age ranges from 2 to 18 years.[45,46,51–53] The remaining six studies were conducted in adults,[40,44,47–49,62] with mean ages ranging across the studies from 43 to 64 years. One study was conducted to investigate POCD in an elderly population, defined as >60 years (no further age information given) (conference abstract).[47] All of the trials included in the Cochrane BIS review[34] studied adult patients (the review's inclusion criteria specified adults over the age of 18 years).

All of the trials included mixed-sex populations. Generally, there was an even mix of males and females in the trials, though there was a higher percentage of males (i.e. >60%) in three studies.[46,48,51] One study did not report the sex of the included patients.[47]

All but one of the studies reported patients' weight.[47] The majority of studies reported weight in kilograms, ranging from a mean of 68–91 kg in the adult studies, and of 17–28 kg in the children studies. In addition to reporting weight in kilograms, one trial also reported body mass index (BMI), which was between 28

and 30 kg/m^2.[49] Another trial reported weight only in terms of BMI, with a mean of 30 kg/m^2, indicating an overweight/obese population.[44] The Cochrane BIS review[34] included one study of obese patients.

Racial origin was reported in only one trial, in which the population was predominantly (>80%) classified as white.[44] The countries in which the trials were conducted included the USA,[45,49,52,53] USA/Canada,[44] China,[40,47,51] Germany,[62] Egypt[48] and India.[46] In the Cochrane BIS review[34] the majority of studies were conducted in Europe or the USA. Seven of the trials were conducted in single centres,[45,46,48,51–53,62] with one trial taking place in two centres,[47] another taking place in three centres,[44] one trial taking place in 13 centres[40] and a trial not reporting the number of centres.[49]

The type of surgery reported in the adult trials varied: open heart,[44] major non-cardiac,[47] major orthopaedic,[49] orthopaedic,[62] and elective moderate abdominal surgery.[48] The surgical procedures in the trials of children included tonsillectomy and/or adenoidectomy,[45] urogenital/urological surgery[46,51] and dental rehabilitation.[52,53]

Only two of the trials reported patient risk factors for awareness.[40,44] To be included in the trial by Avidan and colleagues[44] patients had to be at high risk for intraoperative awareness, demonstrating one or more of the following risk factors: planned open heart surgery; aortic stenosis; pulmonary hypertension; use of opiates; use of benzodiazepines; use of anticonvulsant drugs; daily alcohol consumption; ASA status 4; end-stage lung disease; history of intraoperative awareness; history of, or anticipated, difficult intubation; cardiac ejection fraction of <40%; and marginal exercise tolerance. The trial by Zhang and colleagues[40] included patients receiving TIVA, which they cited as a risk factor for intraoperative awareness. The Cochrane BIS review[34] included four trials that were classified as including patients at high risk of intraoperative awareness.[27,78,79,82]

The eligibility criteria employed by the trials generally excluded patients with significant comorbidities, or factors that may interfere with EEG readings, including epilepsy, cerebrovascular disease, dementia, treatment with opioids and antipsychotic medication, and illicit drug use. Two of the studies permitted inclusion of children with mild cerebral palsy without significant neurological deficit.[52,53] The trials included in the Cochrane BIS review[34] also generally excluded patients with the above factors. Some of these trials also excluded patients considered obese, or patients with diabetes or impaired renal or hepatic function.

The ASA physical status classification of the patients in the trials was generally between I and II, indicating that they were generally healthy, with only mild disease. In three of the trials the ASA status was not reported[45,47,48] (although in one of these trials the inclusion criteria specified patients had to be within I–III[48]). In one trial the proportion of patients with ASA status I–II was 50%, and the remainder of patients were classified as III (severe systemic disease).[49] There was one trial in which patients were predominantly classified as III–IV (IV being classified as a patient with severe systemic disease that is a constant threat to life).[44]

Technologies

The trials varied in the level of detail given on the BIS module and monitors used. Two studies did not provide any information other than that a BIS module was used.[44,47] Most commonly reported was the BIS Monitor Model A-2000 as mentioned in four trials.[40,45,48,62] In one this was described as: 'IP X 2';[46] in another 'version XP, software version 4.0';[62] and in the third trial using Aspect Medical Systems 'Software program Datex-Ohmeda S/5 Collect (v4.0)' (Aspect Medical Systems Inc., Norwood, MA, USA).[48] One trial used BIS (version 3.3, Aspect Medical Systems) using an A-1050 EEG monitor,[45] while another used BIS monitor (XP, algorithm 3.4; Aspect Medical Systems).[49] A further two trials reported using BIS (Aspect Medical Systems), but gave no further information on the software version or the monitor used.[52,53] Although most studies reported using Aspect Medical Systems BIS, one trial reported using the BIS monitor as manufactured by Phillips but using 'Aspect Medical Systems' XP platform technology'.[51] Given the variability in reporting it is not clear how comparable the trials are in terms of the software and BIS algorithms used, which may have implications for the interpretation of the results of the trials.

All of the trials reported the target BIS values to be achieved during anaesthesia. In five trials the target was 40–60.[40,44,45,47,51] In one of these trials the target was increased to 60–70 during last 15 minutes of surgery.[45] In the remaining trials the target values were higher: 45–60;[46] 50 during maintenance (target value of 60 to facilitate rapid emergence from anaesthesia 15 minutes before expected end of surgery);[62] 50–60;[48,49] 55–65;[53] and 60–70.[52]

Although all of the trials compared BIS against standard clinical monitoring, the monitored parameters varied. Only one trial measured ETAC in order to detect possible intraoperative awareness.[44] Audible alarms sounded if the age-adjusted MAC fell outside of 0.7 to 1.3. The remaining nine trials used clinical signs to guide anaesthetic use. In general, a combination of signs were monitored in each trial, most commonly: blood pressure;[46,48,49,52,62,94] heart rate;[48,49,52,53,62] surgical stimulation;[52,53] sweating;[62] tear production;[62] and movement.[62]

Two trials did not explicitly define which signs were monitored other than that they were clinical signs and haemodynamic changes.[45,47] A further trial mentioned that the aim of standard clinical monitoring was to maintain haemodynamic stability while avoiding patient movement and achieving a rapid recovery.[51]

Some of the trials reported that clinical signs were also monitored in the BIS arm, suggesting that adjustment of anaesthesia was based on signs of inadequate anaesthesia as well as BIS values.[48,52,53,62] For example, in one trial[48] changes in anaesthesia were guided by the presence of clinical signs in relation to the BIS value. If the patient exhibited hypertension or tachycardia and the BIS was >60 then sevoflurane was increased. If BIS was in the target range of 50–60, then fentanyl was given. If BIS was <50 then sevoflurane was decreased and the patient checked for lack of analgesia. In the one trial that used ETAC as the comparator to BIS,[44] it was stated that both forms of monitoring were used as part of structured protocols. It was not intended that these protocols would prescribe or restrict the use of anaesthetic agents. Practitioners were able to increase or decrease anaesthetic administration at their discretion if a patient's condition was haemodynamically unstable. The protocols were designed to increase vigilance and to provide warnings that patients might be experiencing awareness. Some trials did not explicitly report whether or not clinical signs were monitored in the BIS arm, and it is possible that in these studies anaesthesia was adjusted based on BIS monitoring in conjunction with changes in clinical signs.

All trials reported that a BIS monitor was used in the standard clinical monitoring arm, but that the values were hidden from the anaesthetist, for example by placing it out of their line of sight, or using a curtain or cover, and also switching off any audible alarms.

The majority of trials did not explicitly report where or when monitoring commenced and ceased. Where details were provided, monitoring started prior to anaesthesia induction[45,46] and in the operating theatre.[46,51,62] Three studies reported cessation of monitoring: until patients achieved discharge criteria from the recovery room (Steward score of 6)[46] and until discharge from the PACU.[52,53]

The training and experience of the anaesthetist in using BIS was rarely mentioned in the trials. The trial by Avidan and colleagues[44] reported that summaries of BIS and ETAC protocols were given to the practitioners to provide education and to increase adherence. Furthermore, signs were affixed to anaesthesia machines to remind practitioners to check BIS/ETAC and consider patient awareness. One of the trials mentioned that the anaesthetist was experienced, but provided no further information.[62]

Anaesthetic agents and protocols

Five of the trials reported that an inhaled general anaesthetic was used for both induction and maintenance.[44,45,51–53] In all but one of these trials sevoflurane was the inhaled anaesthetic used.[53] Two of these trials also gave nitrous oxide in oxygen.[45,51] In the fifth trial patients either received isoflurane, sevoflurane or desflurane.[44] Three trials reported that both i.v. and inhalational general anaesthetic were used.[47–49] In two of these propofol was used for induction of anaesthesia and sevoflurane was given for maintenance.[48,49] The third trial implied that both propofol and an inhalational anaesthetic were given,

but did not provide any further detail.[47] Three trials reported that propofol was given for both induction and maintenance of general anaesthesisa.[40,46,62] One of these also used nitrous oxide in oxygen during the maintenance period.[46]

Only one trial stated that regional anaesthesia was used, although no information was provided on which agent was used.[62] One trial mentioned that regional anaesthesia was used for postoperative pain management.[49] In the remaining nine trials[44–48,51–53,62] it was either reported that regional anaesthesia was not used or the use of regional anaesthesia was not stated.

Use of analgesia at various points during surgery was reported by seven of the trials, including fentanyl,[49,51–53] fentanyl or morphine[45,46] or remifentanil (during induction).[62] One trial reported that analgesia was used at the discretion of the anaesthetist.[40] In three trials the use of analgesia was not stated.[44,47,48] Premedication with midazolam was used in seven trials.[40,44–46,52,53,62] In two of these trials ketamine was also used as premedication.[52,53]

Muscle relaxants were used in seven of the trials, including atracurium,[46,48] cisatracurium,[63] vecuronium bromide[49] and rocuronium bromide.[52,53] One trial did not specify which agent was used.[40]

Duration of anaesthesia was reported by five of the BIS trials[46,48,49,51,62] and ranged from a mean of 40 minutes (paediatric urological surgery)[51] to 126 minutes (major orthopaedic surgery in adults).[49] In the trials featuring adults, duration of anaesthesia was, in general, between 100 and 120 minutes. Duration of surgery was reported by seven of the BIS trials,[40,45,46,48,51–53] and ranged from around 30 minutes (in children undergoing tonsillectomy and/or adenoidectomy)[45] to 160 minutes (children undergoing dental surgery).[52] Not all trials reported both duration of anaesthesia and duration of surgery.

Outcomes

Table 7 illustrates the distribution of outcomes reported by the trials included in this systematic review. The table also shows the frequency of the outcomes in this review, the Cochrane BIS review[34] and the grand total for both reviews.

The most commonly reported outcome was anaesthetic consumption ($n = 30$ trials), followed by recovery outcomes such as time to extubation ($n = 26$ trials); time to eye opening (either spontaneously or in response to command) ($n = 22$ trials); and time to discharge from the PACU. Intraoperative analgesic consumption was reported in 11 trials.

Adverse outcomes were less commonly reported, such as PONV ($n = 3$ trials); and emergence delirium ($n = 1$ trial[59]). One trial, by Leslie and colleagues,[50] reported stroke, myocardial infarction, mortality for all surviving and available patients 30 days post operation. This is a long-term follow-up (median = 4.1 years) publication of the B-Aware trial [NB. A publication of the short-term results of this trial by Myles and colleagues 2004[79] (primary outcome: intraoperative awareness) was included in the Cochrane BIS review[34]].

Six of the 11 BIS trials[40,44,46,49,59,62] included in this systematic review specified a primary outcome measure. In two trials the primary outcome measure was anaesthetic consumption,[46,62] and in another trial the primary outcome measure was time to first movement response.[59] In the other three trials the primary outcome measure was intraoperative awareness.[40,44,49]

In the trial by Avidan and colleagues[44] – which recruited patients classified as at high risk of intraoperative awareness – the incidence of definite intraoperative awareness was the primary outcome measure. The incidence of definite or possible awareness was a secondary outcome. Awareness was assessed by a modified Brice questionnaire (references cited), and assessments were made 72 hours after surgery and 30 days after extubation. Patients who reported memories of the period between 'going to sleep' and 'waking up' were contacted by a different evaluator, who asked additional structured questions. Responses to the questionnaire from patients who had reported memories were reviewed by three independent

TABLE 7 Bispectral Index study outcomes

Study	Avidan[44]	[a]Bannister[45]	[a]Bhardwaj[46]	Chan[47]	Ellerkmann[62]	Leslie[50]	Kamal[48]	Kerssens[49]	[a]Liao[51]	[a]Messieha 2004[52]	[a]Messieha 2005[53]	[a]Zhang[40]	Total this review	Total Cochrane BIS review[34]	Grand total
Outcomes															
Anaesthetic consumption		✗	P	✗	P			✗	✗				6	24	30
Intraoperative awareness	P				✗		✗	P	✗			P	6	4	10
Distressing experience of awareness	✗												1	0	1
Analgesic consumption		✗	✗		✗		✗		✗				5	6	11
Muscle relaxant requirement													0	2	2
Time to response to commands			✗										1	12	13
Time to eye opening			✗				✗		✗				3	19	22
Time to extubation		✗	✗				✗			✗	✗		5	21	26
Time to laryngeal mask airway removal									✗				1	0	1
Time to first movement response		✗							P				1	0	1
Time to recovery of orientation									✗				1	7	8
Time to phonation									✗				1	0	1

continued

TABLE 7 Bispectral Index study outcomes (*continued*)

Study	Avidan[44]	[a]Bannister[45]	[a]Bhardwaj[46]	Chan[47]	Ellerkmann[62]	Leslie[50]	Kamal[48]	Kerssens[49]	[a]Liao[51]	[a]Messieha 2004[52]	[a]Messieha 2005[53]	[a]Zhang[40]	Total this review	Total Cochrane BIS review[34]	Grand total
PACU stay		✗	✗				✗			✗			4	12	16
Time to home readiness													0	7	7
Monitoring device values		✗						✗			✗		3	0	3
Postoperative nausea and vomiting									✗				1	2	3
Emergence delirium									✗				1	0	1
POCD				✗									1	1	2
Parental satisfaction									✗				1	0	1
Treatment of haemodynamic events													0	0	0
Haemodynamic profiles		✗	✗						✗				3	3	6
Stroke						✗							1	0	1
Myocardial infarction						✗							1	0	1
Mortality						✗							1	0	1

✗, stated secondary outcome measure/not stated whether primary or secondary outcome measure; P, primary outcome measure.

a Study of children.

experts, who determined whether the reported event involved definite awareness, possible awareness or no awareness. Where there was a difference in judgement over an awareness episode a fourth expert made the final determination. This study was designed specifically to evaluate the effects of BIS on intraoperative awareness, and to overcome methodological limitations of a previous single-centre trial by the same investigators (the B-Unaware trial[27] – included in the Cochrane BIS review[34]) by including a study sample sufficiently large enough to detect a relatively rare outcome such as awareness.

The trial by Zhang and colleagues[40] also reported incidence of confirmed awareness, or possible awareness, using a Brice questionnaire. Assessments were made on the first and fourth day following surgery. An independent evaluating committee was used to verify cases of awareness. The patients in this trial were noted to be at increased risk of intraoperative awareness after receiving TIVA.

The trial by Kerssens and colleagues[49] measured explicit awareness, via a patient interview, as well as implicit awareness, via a word recognition test. This is the only trial identified by the current systematic review that measured implicit awareness. The underlying hypothesis was that intraoperative memory could occur either because of insufficient anaesthetic or stress-induced learning mechanisms during unconsciousness (i.e. intraoperative memory could be dependent on and/or independent of depth of anaesthesia). Six hours after surgery, patients were interviewed using questions similar to the Brice interview, consisting of five questions, with additional questions asked as necessary. Following the interview a recognition memory test was performed. During anaesthesia, sequences of pre-determined neutral words were played to patients through headphones. The postoperative memory test involved playing pre-determined combinations of words that had been used during anaesthesia, and distractor words, to patients though headphones. Patients were instructed to listen to each test sequence and select the word played during surgery, or to guess if necessary. The main analysis of this study was the effect of study group assignment on recognition memory test performance, but given the low incidence of explicit recall (awareness) the study was not powered to detect differences in explicit recall. An arbitrary sample size of 100 patients was chosen to assess recognition memory.

Intraoperative awareness was also reported as a non-primary outcome by three other BIS trials included in this systematic review.[48,51,62] In these trials, awareness was one of a number of outcomes measured, and patients were not identified as being at particular risk. Awareness was assessed by a patient interview administered at various times up to 3 days post operation. In the trial by Ellerkmann and colleagues[62] interviews took place on the first and third postoperative days, in the trial by Kamal and colleagues[48] interviews took place on the first, second and third days postoperatively, and in the trial by Liao and colleagues[51] the timing was not specified. Little detail of the interviews was given other than that 'patients were questioned for recall of events, hearing vague sounds, feeling surgical instruments or dressing application, or dreaming';[48] or patients were asked 'whether they could recall any event or dreaming during the intraoperative period';[51] or that a 'standardised interview' was used (reference cited).[62]

The Cochrane BIS review conducted a meta-analysis of explicit intraoperative awareness, which included four RCTs.[27,78,79,82] The Cochrane review also included a further eight trials[61,63,66,83,84,87–89] that reported explicit intraoperative awareness, but the review did not classify these as featuring patients at higher risk. They were not included in any meta-analysis and the impact on awareness was not commented on by the Cochrane review. The Cochrane BIS review did not report whether any of the included trials measured implicit awareness or assessed awareness during surgery using techniques such as the isolated forearm technique.

Assessment of outcomes: Bispectral Index

The following sections report the results of the BIS trials included in this systematic review. Tabulated data are from the studies identified by this review (i.e. supplemental to the trials in the Cochrane BIS review). Where appropriate we have updated the meta-analyses of the Cochrane BIS review with studies from the current review, presented graphically in forest plots. Where it was not appropriate to update the Cochrane BIS meta-analysis we have presented the results of the meta-analysis narratively.

Intraoperative awareness

Table 8 gives the results of the six trials included in this systematic review which measured the impact of BIS monitoring on explicit intraoperative awareness, as assessed by patient interview.

No cases of awareness were reported at all in three trials,[48,51,62] and a very low number of cases were reported in a fourth trial.[49] As stated earlier, these trials were not specifically designed to detect the effect of depth of anaesthesia monitoring on awareness, and therefore are unlikely to have sufficiently large enough sample sizes for relatively rare awareness events. In the trial by Avidan and colleagues,[44] which included patients classified at higher risk for intraoperative awareness and was statistically powered for this outcome, there was a higher percentage of both definite awareness, and of definite or possible awareness cases, in the group who received BIS monitoring than the group who had standard clinical monitoring. However, these differences were not statistically significant. Avidan and colleagues[44] also reported patient distress and sequelae resulting from intraoperative awareness, as a post hoc secondary outcome. Distress was measured using the Michigan Awareness Classification tool (reference supplied) and was characterised by reports of fear, anxiety, suffocation, sense of doom or sense of impending death. There was a higher percentage of distress reported in the BIS-monitored group, but no statistically significant difference between groups.

In contrast to Avidan and colleagues,[44] Zhang and colleagues[40] reported a significantly lower incidence of confirmed intraoperative awareness in patients monitored by BIS than in those who received standard clinical monitoring. Incidence of possible awareness was also lower for BIS-monitored patients, although not statistically significant. The incidence of confirmed or possible awareness was significantly lower for BIS-monitored patients.

Intraoperative awareness was the primary outcome measure in the Cochrane BIS review.[34] However, as stated earlier, the review reported awareness outcomes only for trials in its set which were conducted with patients considered to be at higher risk of awareness ($n = 4$).[27,78,79,82] The Cochrane review combined these

TABLE 8 Intraoperative awareness during BIS monitoring (all patients, irrespective of risk of awareness)

Study	BIS	Standard clinical monitoring	Mean difference (95% CI), *p*-value
Avidan *et al.*,[44] *n/N* (%)			
Definite awareness	7/2861 (0.24)	2/2852 (0.07)	0.17 (–0.03 to 0.38), $p = 0.98$
Definite or possible awareness	19/2861 (0.66)	8/2852 (0.28)	0.38 (0.03 to 0.74), $p = 0.99$
Patient distress and sequelae resulting from intraoperative awareness	8/2861 (0.28)	1/2852 (0.04)	0.24 (0.04 to 0.45), $p = 0.99$
Ellerkmann *et al.*,[62] *n/N*	0/27	0/27	–
Kamal *et al.*,[48] *n/N*	0/28	0/29	–
Kerssens *et al.*,[49] *n/N* (%)	2/67 (3)	1/61 (2)	NR
[a]Liao *et al.*,[51] *n/N*	0/52	0/54	–
Zhang *et al.*,[40] *n/N* (%)			
Confirmed awareness	4/2919 (0.14)	15/2309 (0.65)	OR 0.21 (0.07 to 0.63), $p = 0.002$
Possible awareness	4/2919 (0.14)	6/2309 (0.26)	$p = 0.485$
Confirmed or possible awareness	8/2919 (0.27)	21/2309 (0.9)	$p = 0.01$

CI, confidence interval; NR, not reported.
a Study of children.

four trials in a fixed-effect meta-analysis, and we have updated this meta-analysis to include the two trials from our study set that featured higher risk patients.[40,44] *Figure 2* reports the results of this meta-analysis.

The meta-analysis included three subgroup analyses: trials that used inhaled GA only; trials that used a mixture of inhaled and i.v. anaesthesia; and trials that used TIVA. The original overall pooled Peto's odds ratio (OR) from the Cochrane review was 0.33 [95% confidence interval (CI) 0.13 to 0.84], indicating a statistically significant difference between groups favouring BIS. The addition of the trials by Avidan and colleagues[44] and Zhang and colleagues[40] increased the OR to 0.45 (95% CI 0.25 to 0.81). Caution is advised in the interpretation of this result as, overall, there was statistically significant heterogeneity ($p=0.009$; $I^2=79\%$). In the subgroup of trials that used only inhaled anaesthesia the Peto's OR was 1.79 (95% CI 0.63 to 5.11) in favour of standard clinical monitoring. This is in contrast with the other two subgroups, which favoured BIS monitoring.

Explicit intraoperative awareness was an outcome measured in a further eight trials included in the Cochrane BIS review. However, as stated earlier, the review did not report the results of these trials for this outcome. We examined these studies (data not formally extracted) and note that no patients in any of these eight trials reported experiencing intraoperative awareness. It is unlikely that these studies were adequately statistically powered to detect awareness.

The trial by Kersens and colleagues[49] was the only study to report implicit awareness, that is awareness that the patient does not necessarily recall experiencing. The probability of postoperatively selecting a word presented during anaesthesia (target) was higher in the BIS monitoring group (mean 0.371 ± 0.132) than in the standard clinical monitoring group (mean 0.323 ± 0.132). The probability of postoperatively selecting a word not presented during anaesthesia (distractor) was lower in the BIS monitoring group (mean 0.315 ± 0.117) than in the standard clinical monitoring group (mean 0.338 ± 0.119). It was not reported whether or not differences between study groups were statistically significant. Intragroup and overall differences between postoperative target and distractor word recall suggest that BIS-monitored patients were more likely to select words presented during anaesthesia than words not presented during anaesthesia, but standard clinical monitoring patients performed no better than chance in word selection (within-group difference in probability of selecting target word or distraction word: BIS: $p=0.001$; standard clinical monitoring: $p \geq 0.05$).

Anaesthetic consumption

Table 9 reports the impact of BIS monitoring on intraoperative general anaesthetic requirement.

Six of the 11 BIS trials included in this systematic review reported this outcome measure,[45–47,49,51,62] two of which reported it to be the primary outcome.[46,62] Three of the trials reported volatile anaesthetic consumption, all of which were for sevoflurane. Two of these three trials were conducted in children.[45,51] The mean end-tidal sevoflurane concentration (%) during maintenance of GA in each of these three trials was statistically significantly lower in the BIS-monitored group than in the standard clinical monitoring group. The other three trials[46,47,62] reported i.v. anaesthetic consumption, all of which used propofol. One of these trials was conducted with children.[46] In two of the three trials the maintenance dose was higher in BIS-monitored patients than standard clinical monitoring, but with no statistically significant differences between groups.[46,62] The third trial was reported in a conference abstract, and limited results are given, except that there was a 25.3% reduction in propofol consumption compared with standard clinical monitoring.[47]

The Cochrane BIS review[34] conducted random-effects meta-analyses for anaesthetic consumption, producing separate meta-analyses for volatile anaesthetic consumption and for propofol consumption. We have updated these meta-analyses with studies included in our systematic review. *Figure 3* shows the results of the meta-analysis of volatile anaesthetic consumption (sevoflurane).

Study or subgroup	BIS Events	BIS Total	Standard monitoring Events	Standard monitoring Total	Weight	Peto's OR Peto, fixed, 95% CI
Inhaled GA only						
Avidan 2008[27]	2	967	2	974	8.8%	1.01 (0.14 to 7.16)
Avidan 2011[44]	7	2861	2	2852	19.7%	3.03 (0.82 to 11.21)
Muralidhar 2008[78]	0	10	0	10		Not estimable
Puri 2003[82]	0	14	1	16	2.2%	0.15 (0.00 to 7.80)
Subtotal (95% CI)		3852		3852	30.7%	1.79 (0.63 to 5.11)
Total events	9		5			
Heterogeneity: χ^2 = 2.46, df = 2 (p = 0.29); I^2 = 19%						
Test for overall effect: z = 1.09 (p = 0.28)						
Mixed anaesthesia						
Myles 2004[79]	2	1225	11	1238	28.4%	0.25 (0.08 to 0.75)
Subtotal (95% CI)		1225		1238	28.4%	0.25 (0.08 to 0.75)
Total events	2		11			
Heterogeneity: Not applicable						
Test for overall effect: z = 2.48 (p = 0.01)						
Total intravaenous GA (TIVA)						
Muralidhar 2008[78]	0	10	0	10		Not estimable
Zhang 2011[40]	4	2919	15	2309	41.0%	0.24 (0.10 to 0.60)
Subtotal (95% CI)		2929		2319	41.0%	0.24 (0.10 to 0.60)
Total events	4		15			
Heterogeneity: Not applicable						
Test for overall effect: z = 3.06 (p = 0.002)						
Total (95% CI)		8006		7409	100.0%	0.45 (0.25 to 0.81)
Total events	15		31			
Heterogeneity: χ^2 = 11.99, df = 4 (p = 0.02); I^2 = 67%						
Test for overall effect: z = 2.68 (p = 0.007)						
Test for subgroup differences: χ^2 = 9.53, df = 2 (p = 0.009); I^2 = 79.0%						

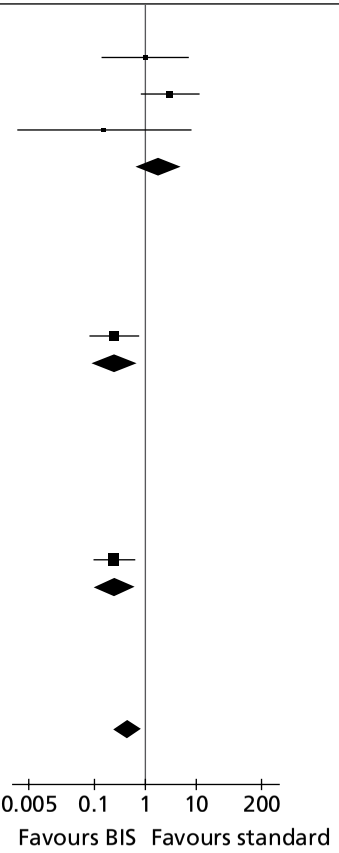


FIGURE 2 Meta-analysis of intraoperative awareness during BIS monitoring (patients classified at higher risk of awareness)

TABLE 9 Consumption/concentration of anaesthetic during BIS monitoring

Study	BIS	Standard clinical monitoring	Mean difference (95% CI), *p*-value
Volatile anaesthetic (sevoflurane), mean ± SD end-tidal sevoflurane concentration (%)			
[a]Bannister *et al.*[45]			
Maintenance of GA	1.8 ± 0.4	2.4 ± 0.6	$p < 0.05$
Last 15 minutes of GA	1.6 ± 0.6	2.1 ± 0.7	$p < 0.05$
End of procedure	1.1 ± 0.6	1.5 ± 0.7	NS
Kerssens *et al.*[49]			
Maintenance phase	1.31 ± 0.29	1.56 ± 0.29	$p < 0.001$
[a]Liao *et al.*[51]			
Maintenance	2.5 ± 0.4	2.9 ± 0.5	0.001;[b] $p < 0.01$[c]
Propofol consumption			
[a]Bhardwaj *et al.*[46]			
Maintenance phase µg/kg/minute, mean (SD)	108.6 (37.8)	106.6 (38.9)	Mean difference 1.9 (–19.9 to 23.7), *p*-value NR
Chan *et al.*[47]	25.3% reduction vs standard clinical monitoring[d]		
Ellerkmann *et al.*[62]			
Maintenance phase µg/kg/minute, mean (SD)	104 (20)	101 (22)	Entropy/BIS vs standard clinical monitoring, $p = 0.27$

NR, not reported; NS, not statistically significant; SD, standard deviation.

a Study of children.

b For three-group comparison (BIS; auto-regressive index; standard clinical monitoring).

c Post hoc comparison BIS vs standard clinical monitoring.

d Assumed that this comparison was between BIS and standard clinical monitoring; however, the wording of the results does not rule out that the comparison may instead have been between BIS and a matched 'control' group.

As stated, two of the three studies measuring sevoflurane consumption in our systematic review were conducted in children. The Cochrane BIS review[34] only included studies of adults, therefore we have only updated their meta-analysis with the one study of adults from our set (Kerssens and colleagues[49]). The original mean difference in MAC equivalents from the Cochrane review for sevoflurane consumption was –0.16 (–0.29 to –0.04), indicating a statistically significant difference in favour of BIS. Updating the meta-analysis with the trial by Kerssens and colleagues[49] reduced the mean difference slightly to –0.15 (95% CI –0.25 to –0.06), but remained statistically significant. However, caution is advised because of a high degree of unexplained statistical heterogeneity ($p < 0.00001$; $I^2 = 85\%$).

Figure 4 shows the results of the meta-analysis of propofol consumption.

As stated, one of the three studies measuring propofol consumption in our systematic review was conducted in children.[46] As the Cochrane BIS review[34] only included studies of adults, therefore we have updated their meta-analysis with one of the two studies of adults from our set.[62] (NB. The other adult study[47] was only reported in a conference abstract and the results were not reported in a format amenable to meta-analysis.) The original mean difference propofol consumption (mg/kg/minute) in the Cochrane review was –1.44 (–1.95 to –0.93), indicating a statistically significant difference in favour of BIS. Updating the meta-analysis with the trial by Ellerkmann and colleagues[62] reduced the mean difference slightly to –1.30 (95% CI –1.83 to –0.76), but remained statistically significant. Again, caution is required because of highly significant unexplained statistical heterogeneity ($p < 0.00001$; $I^2 = 80\%$).

Study or subgroup	BIS Mean	BIS SD	BIS Total	Standard monitoring Mean	Standard monitoring SD	Standard monitoring Total	Weight	Mean difference IV, random, 95% CI
Ahmad 2003[65]	1.19	0.14	49	1.21	0.48	48	12.6%	–0.02 (–0.16 to 0.12)
Basar 2003[67]	0.81	0.11	30	0.84	0.14	30	16.0%	–0.03 (–0.09 to 0.03)
Boztug 2006[68]	0.39	0.11	24	0.49	0.11	23	16.0%	–0.10 (–0.16 to –0.04)
Kerssens 2009[49]	0.72	0.16	62	0.86	0.16	47	16.1%	–0.14 (–0.20 to –0.08)
Nelskyla 2001[80]	0.28	1.11	32	0.27	1.61	30	1.8%	0.01 (–0.68 to 0.70)
Paventi 2001[81]	0.46	2.42	45	1.12	2.42	45	0.9%	–0.66 (–1.66 to 0.34)
Song 1997[84]	0.50	0.17	15	1.00	0.17	15	13.6%	–0.50 (–0.62 to –0.38)
Tufano 2000[86]	0.47	0.11	40	0.78	0.78	40	8.3%	–0.31 (–0.55 to –0.07)
Zohar 2006[89]	0.25	0.15	25	0.31	0.20	25	14.6%	–0.06 (–0.16 to 0.04)
Total (95% CI)			322			303	100.0%	–0.15 (–0.25 to –0.06)

Heterogeneity: $\tau^2 = 0.01$; $\chi^2 = 52.83$, df = 8 ($p < 0.00001$); $I^2 = 85\%$
Test for overall effect: $z = 3.08$ ($p = 0.002$)

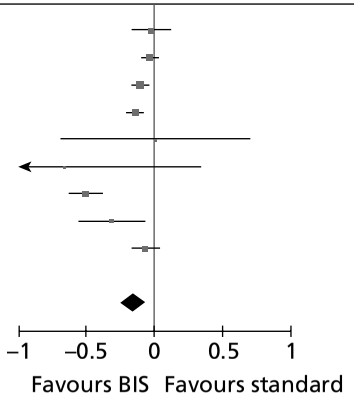


FIGURE 3 Meta-analysis of volatile anaesthetic consumption (sevoflurane) during BIS monitoring, MAC equivalents. IV, inverse variance, SD, standard deviation.

Study or subgroup	BIS Mean	BIS SD	BIS Total	Standard monitoring Mean	Standard monitoring SD	Standard monitoring Total	Weight	Mean difference IV, random, 95% CI
Chiu 2007[70]	2.90	2.20	10	6.00	1.93	10	5.0%	–3.10 (–4.91 to –1.29)
Anez 2001[66]	8.04	2.52	20	11.94	2.28	19	6.1%	–3.90 (–5.41 to –2.39)
Struys 2001[85]	6.39	1.13	10	6.48	1.59	10	7.4%	–0.09 (–1.30 to 1.12)
Tufano 2000[86]	5.88	1.20	40	7.80	2.72	40	8.8%	–1.92 (–2.84 to –1.00)
Muralidhar 2008[78]	3.38	0.99	10	5.07	0.70	10	9.6%	–1.69 (–2.44 to –0.94)
Mayer 2007[76]	4.70	1.10	22	5.88	1.20	22	10.0%	–1.18 (–1.86 to –0.50)
Ellerkmann 2010[62]	6.24	1.20	30	6.06	1.30	30	10.2%	0.18 (–0.45 to 0.81)
Masuda 2002[75]	4.30	1.10	20	4.90	0.80	19	10.3%	–0.60 (–1.20 to 0.00)
Luginbuhl 2003[91]	6.03	1.40	40	6.64	0.90	40	10.7%	–0.61 (–1.13 to –0.09)
Kreuer 2003[63]	4.80	1.00	40	6.80	1.20	40	10.8%	–2.00 (–2.48 to –1.52)
Gan 1997[71]	6.96	1.98	115	8.04	1.74	125	10.9%	–1.08 (–1.55 to –0.61)
Total (95% CI)			357			365	100.0%	–1.30 (–1.83 to –0.76)

Heterogeneity: $\tau^2 = 0.63$; $\chi^2 = 61.12$, df = 10 ($p < 0.00001$); $I^2 = 84\%$
Test for overall effect: $z = 4.74$ ($p < 0.00001$)

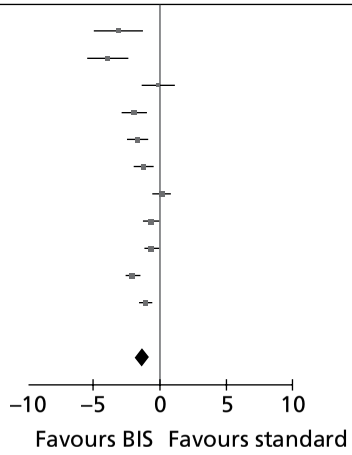


FIGURE 4 Meta-analysis of propofol consumption during BIS monitoring, mg/kg/minute. IV, inverse variance; SD, standard deviation.

Outcomes related to postanaesthesia care unit stay

Five of the 11 BIS trials[45,46,48,52,53] in our systematic review reported this outcome, of which four were conducted with children.[45,46,52,53] In none of the trials was use of PACU a primary outcome. All of the studies appear to have reported the time to discharge from the PACU. However, it was not always clear exactly when the time to discharge began (e.g. from the end of skin closure, termination of anaesthetic or from admittance to the PACU). Bannister and colleagues[45] reported time from end of surgery to PACU discharge, whereas Kamal and colleagues[48] and both the trials by Messieha and colleagues[52,53] stated measuring the end of general anaesthetic to PACU discharge (although in one of these trials[48] data do not appear to be reported for that outcome). Bhardwaj and colleagues[46] did not provide any detail on timing. Detail of discharge criteria varied between the trials. Bannister and colleagues[45]and Kamal and colleagues[48] both used the Aldrete scoring system (score of >9), whereas Bhardwaj and colleagues[46] used the Steward recovery scoring system (eligibility = score of 6). Messieha and colleagues[52,53] did not report use of discharge criteria.

Table 10 shows the results of the trials relating to stay in the PACU.

In all trials, time to discharge from the PACU was statistically significantly greater in the standard clinical monitoring group than in the BIS monitoring group, with mean differences in the range of 6.7–30 minutes. One trial did not report data for this outcome, mentioning that time to discharge was comparable between groups. There was also a statistically significant difference in the one trial that measured time to arrival at the PACU, with reduction of 4.7 minutes for BIS monitoring.[48] The two trials that reported duration of stay in the PACU both reported statistically significant differences in favour of BIS.[52,53]

Eligibility for discharge from the PACU unit was one of the secondary outcomes from the Cochrane BIS review.[34] The review meta-analysed the outcome 'PACU' stay, including data from 12 trials. Examination of characteristics of the trials included in this meta-analysis, as summarised in the Cochrane review, show that some of the trials reported time to arrival in the PACU, time to discharge from the PACU and length of stay in the PACU. These all appear to have been included in the same meta-analysis, and there is no discussion

TABLE 10 Postanaesthesia care unit stay outcomes following BIS monitoring

Study	BIS	Standard clinical monitoring	Mean difference (95% CI), *p*-value
[a]Bannister *et al.*[45]			
Time to discharge from the PACU minutes mean (SD)	20.0 (± 7.9)	26.7 (± 11.2)	$p<0.05$
[a]Bhardwaj *et al.*[46]	Time to achieve a Steward recovery score of 6 (for discharge from the recovery room) reported to be comparable in the two groups		
Kamal *et al.*[48]			
Arrival at PACU (minutes), mean (SD)	9.4 (± 1.9)	14.1 (± 2.8)	$p<0.01$
PACU discharge (minutes), mean (SD)	53.9 (± 14.7)	78.6 (± 21.5)	$p<0.01$
[a]Messieha *et al.*[52]			
Time to PACU discharge (minutes), mean (SD)	60 (± 13)	90 (± 11)	$p<0.001$
Duration of PACU stay (minutes), mean (SD)	45 (± 8)	71 (± 9)	$p<0.001$
[a]Messieha *et al.*[53]			
Duration of PACU stay (minutes), mean (SD)	47 (± 17)	63 (± 17)	$p = 0.02$

a Study of children.

about how timings may differ according to these different outcomes. Given this lack of clarity, and the fact that the Cochrane review included only trials of adults, we decided not to update this meta-analysis with data from trials identified in the current review. The pooled random-effects mean difference reported in the Cochrane review was –7.63 minutes (95% CI –12.50 to –2.76 minutes) in favour of BIS. However, caution is advised for the reasons given above, as well as a high degree of statistical heterogeneity ($p < 0.00001$; $I^2 = 82\%$). The results of the meta-analysis are similar to the results of the trials included in the current review (i.e. showing a benefit for BIS monitoring).

Time to recovery from anaesthesia

The trials included in the current systematic review reported a variety of outcomes relating to recovery from anaesthesia, including time to tracheal extubation, time to eye opening and movement responses.

Table 11 reports the time to tracheal extubation following surgery.

Five of the 11 BIS trials included in the current systematic review measured time to extubation, of which four were conducted with children.[45,46,52,53] None of these studies considered this to be a statistically powered primary outcome measure. Timing was reported to have begun from end of surgery in three studies,[45,52,53] and from termination of anaesthetic in two studies.[46,48] Extubation times were shorter for BIS-monitored patients than for those receiving standard clinical monitoring by as much as 5 minutes or as little as 0.5 minutes. Differences between groups were reported to be statistically significant in two trials,[45,53] but not in two other trials.[48,52] One trial did not report numerical data, stating that times were comparable between groups.[46]

A sixth study, conducted with children, reported time to laryngeal mask airway removal following surgery as an outcome.[51] The mean time [standard deviation (SD)] in minutes was 1.8 (1.6) in the BIS-monitored group, and 2.1 (2.4) in the standard clinical monitoring group ($p = 0.93$), indicating no statistically significant differences between groups.

Time to extubation was one of the secondary outcomes from the Cochrane BIS review.[34] The review meta-analysed data from 21 trials. Given that four of the five trials[45,46,52,53] in the current systematic review were conducted in children and the Cochrane review was restricted to trials of adults, we have not updated their meta-analysis. The overall random-effects mean difference in time to extubation was –2.87 minutes (95% CI –3.74 to –1.99 minutes), indicating a statistically significant difference in favour of BIS. Caution is advised as there was a high degree of statistical heterogeneity ($p < 0.00001$; $I^2 = 79\%$).

Table 12 reports time to eye opening following surgery.

TABLE 11 Time to extubation following BIS monitoring

Study	BIS	Standard clinical monitoring	Mean difference (95% CI), *p*-value
Mean (SD) time to extubation (minutes)			
[a]Bannister *et al.*[45]	7.1 (3.7)	11.3 (5.9)	$p < 0.05$
[a]Bhardwaj *et al.*[46]	Time to extubation reported to be comparable in the two groups		
Kamal *et al.*[48]	4.3 (2.1)	4.8 (2.3)	$p > 0.05$
[a]Messieha *et al.*[52]	9 (5)	13 (5)	$p = 0.07$
[a]Messieha *et al.*[53]	5 (2)	10 (7)	$p = 0.04$

a Study of children.

Three trials included in the current systematic review reported time to eye opening, two of which were conducted with children.[46,51] Timing was reported to have begun immediately after the last surgical stitch in two studies[48,51] and from the end of surgery in one trial.[46] Times were shorter in BIS-monitored patients, although by modest duration (up to 1 minute) and there were no statistically significant differences between groups. One trial provided only narrative results, reporting comparable times between groups.

Time to eye opening was one of the secondary outcomes from the Cochrane BIS review.[34] The review meta-analysed data from 19 trials. Given that two of the three trials in the current systematic review were conducted in children and the Cochrane review was restricted to trials of adults, we have not updated their meta-analysis. The overall random-effects mean difference in time to extubation was –2.14 minutes (95% CI –2.99 to –1.29 minutes), indicating a statistically significant difference in favour of BIS. Caution is advised as there was a high degree of statistical heterogeneity ($p < 0.00001$; $I^2 = 83\%$). The results of the meta-analysis are more conclusive than those of the relatively smaller number of trials included in the current review.

Table 13 reports the results of three trials that reported other recovery outcomes.

All three of the trials[45,46,51] reporting other recovery outcomes were conducted with children. Bannister and colleagues[45] reported mean time to first movement, with a statistically significant reduction for BIS-monitored patients of 2.8 minutes. Similarly, Liao and colleagues[51] reported a statistically significant

TABLE 12 Time to eye opening following BIS monitoring

Study	BIS	Standard clinical monitoring	Mean difference (95% CI), *p*-value
Mean (SD) time to eye opening (minutes)			
[a]Bhardwaj *et al.*[46]	Time to eye opening reported to be comparable in the two groups		
Kamal *et al.*[48]	4.1 (1.6)	4.4 (1.9)	$p > 0.05$
[a]Liao *et al.*[51]	15.0 (16.4)	16.1 (11.3)	$p = 0.17$[b]

a Study of children.
b For three-group comparison (BIS; auto-regressive index; standard clinical monitoring).

TABLE 13 Time to other recovery outcomes

Study	BIS	Standard clinical monitoring	Mean difference (95% CI), *p*-value
[a]Bannister *et al.*[45]			
Mean ± SD time to first movement response (minutes)	4.2 ± 3.7	7.0 ± 3.9	$p < 0.05$
[a]Bhardwaj *et al.*[46]			
Time to response commands	Time to response to commands reported to be comparable in the two groups		
[a]Liao *et al.*,[51] mean ± SD time to emergence from anaesthesia (minutes)			
Spontaneous movement	3.6 ± 2.7	6.1 ± 5.7	0.02[b]; $p < 0.05$[c]
Phonation	8.4 ± 5.2	12.9 ± 9.0	0.11[b]

a Study of children.
b For three-group comparison (BIS; auto-regressive index; standard clinical monitoring).
c Post hoc comparison BIS vs standard clinical monitoring.

reduction in time to first spontaneous movement of 2.5 minutes. This trial[51] also reported a shorter time to phonation (making a vocal sound) of 4.5 minutes, but this was not statistically significant. Bhardwaj and colleagues[46] reported time to response to commands, commenting that this was comparable in the two groups but not reporting any numerical data.

Postoperative nausea and vomiting

Postoperative nausea and vomiting was reported by only one of the trials included in the current systematic review, the trial by Liao and colleagues.[51] There was no difference between patients in the BIS and standard clinical monitoring groups in terms of nausea [$n = 5$ (10%); $n = 6$ (11%), respectively, $p = 0.95$] or vomiting [$n = 2$ (4%); $n = 3$ (6%), respectively, $p = 0.88$]. Postoperative nausea and vomiting was not reported by the Cochrane BIS review.[34]

Emergence delirium

Liao and colleagues[51] also reported the incidence of emergence delirium, as measured by the Paediatric Anaesthetic Emergence Delirium (PAED) instrument (noted to be valid and reliable by the authors, reference cited). Assessment took place by a trained observer in the PACU every 5 minutes after awakening for 30 minutes. The highest score during this period was used in the final PAED score. (NB. A description of the instrument and what the scores mean is not given.) There was no statistically significant difference between BIS and standard clinical practice monitored patients [median (interquartile range) score 18 (14–16); 15 (13–15), respectively, $p = 0.94$].

Postoperative cognitive dysfunction

The only trial to report postoperative cognitive dysfunction was that of Chan and colleagues, who studied an elderly patient population.[47] Cognitive dysfunction was assessed by a battery of eight neuropsychology tests before and at 1 and 3 weeks after surgery (no information on the tests reported). POCD was confirmed when two or more test parameters or the combined z-value > 1.96 (no further information given). There was no statistically significant difference between BIS and standard clinical monitoring in rates of dysfunction at 1 week post surgery [146 (32.5%); 177 (39.1%), respectively, $p = 0.07$]. However, the difference between groups become significant at 3 months post surgery [36 (8.1%); 54 (12%), respectively, $p = 0.03$; OR 1.6 (95% CI 1.0 to 2.4)]. Caution is advised as this trial[47] was reported in a conference abstract therefore detail of its characteristics are lacking, prohibiting a thorough appraisal of its methodological quality. As the abstract was published in 2010 a full publication potentially may be available in the near future.

Mortality, myocardial infarction and stroke

One trial, by Leslie and colleagues,[50] reported stroke, myocardial infarction and mortality for all surviving and available patients 30 days post operation (*Table 14*). This is a long-term follow-up (median = 4.1 years) publication of the B-Aware trial[79] in patients classified at higher risk of intraoperative awareness because of factors such as type of surgery (e.g. high-risk cardiac surgery), health status (e.g. cardiovascular impairment) and lifestyle (e.g. heavy alcohol intake). [NB. A publication of the short-term results of this trial by Myles and colleagues[79] (primary outcome: intraoperative awareness) was included in the Cochrane BIS review.[34] Results of this trial are presented earlier in this report.]

TABLE 14 Mortality, myocardial infarction and stroke

Outcome	Group 1 BIS	Group 2 Routine care	OR or HR (95% CI), *p*-value
Mortality rate per 1000 patient-years (95% CI)	67 (60 to 76)	70 (62 to 79)	HR 0.86 (0.72 to 1.01), $p = 0.07$
Myocardial infarction, *n* (%)	105 (9)	111 (9)	OR 0.85 (0.64 to 1.14), $p = 0.28$
Stroke, *n* (%)	53 (4)	62 (5)	OR 0.79 (0.54 to 1.16), $p = 0.22$

HR, hazard ratio, based on multivariate analyses; OR, odds ratio.

There was no statistically significant difference between BIS-monitored patients and patients who received routine care in mortality, myocardial infarction or stroke.

Summary of Bispectral Index assessment

- Six trials included in this systematic review measured the impact of BIS monitoring on explicit intraoperative awareness. Four of these trials reported few or no cases of awareness; however, they were not statistically powered to detect this outcome. The other two trials were powered to detect awareness and we added them to the meta-analysis from the Cochrane BIS review (restricted to patients considered to be at higher risk of awareness). The pooled Peto's OR was 0.45 (95% CI 0.25 to 0.81), in favour of BIS. However, there was statistically significant heterogeneity and a non-significant difference in the subgroup of trials in which only inhaled GA was used.
- Three trials included in this systematic review reported changes in sevoflurane consumption, all of which were statistically significantly lower with BIS monitoring. We updated the Cochrane meta-analysis with one of these trials, producing a pooled mean difference of –0.15 (95% CI –0.25 to –0.06) MAC equivalents in favour of BIS (with unexplained statistically significant heterogeneity).
- Three trials included in this systematic review reported changes in propofol consumption. In two of these the maintenance dose was higher in BIS-monitored patients than standard clinical monitoring, but not statistically significant. In the third trial propofol consumption was lower for BIS. We updated the Cochrane meta-analysis with one of these trials, producing a pooled mean difference of –1.33 mg/kg/minute (95% CI –1.82 to –0.84 mg/kg/minute), in favour of BIS (with unexplained statistically significant heterogeneity).
- Five trials included in this systematic review reported time to discharge from the PACU, all of which appeared to be secondary outcomes. In all trials time to discharge was statistically significantly shorter in BIS-monitored patients, with mean differences in the range of 6.7–30 minutes. The Cochrane BIS review did a meta-analysis of the outcome 'PACU stay' (including time to arrival in the PACU, time to discharge from the PACU, and length of stay in the PACU). The pooled mean difference was –7.63 minutes (95% CI –12.50 to –2.76 minutes) in favour of BIS (with unexplained statistically significant heterogeneity).
- Five trials included in this systematic review measured time to tracheal extubation, as a secondary outcome. Extubation times were shorter for BIS-monitored patients compared with standard clinical monitoring by as much as 5 minutes, and as little as 0.5 minutes, but not always statistically significant. The pooled mean difference in the Cochrane review for this outcome was –2.87 minutes (95% CI –3.74 to –1.99 minutes) in favour of BIS (with unexplained statistically significant heterogeneity).
- Three trials included in the current systematic review reported time to eye opening as a secondary outcome. Times were shorter in BIS-monitored patients, although by modest duration (up to 1 minute), and there were no statistically significant differences between groups. The pooled mean difference in the Cochrane review for this outcome was –2.14 minutes (95% CI –2.99 to –1.29 minutes), indicating a statistically significant difference in favour of BIS (with unexplained statistically significant heterogeneity).
- Postoperative nausea and vomiting was reported by only one trial. Incidence of nausea and vomiting was low (around 10% or less) and there was no statistically significant difference between groups.
- Only one trial reported the incidence of postoperative cognitive dysfunction. There was no statistically significant difference between groups in rates of dysfunction at 1 week post surgery. By 3 months post surgery, incidence had fallen to around 8–12%, with a significant difference in favour of BIS. This study was reported only as a conference abstract and it is not clear whether or not this outcome was adequately statistically powered.
- Longer-term postoperative outcomes of stroke, myocardial infarction and mortality were reported by only one trial (median of 4.1 years post operation), as secondary outcomes. Mortality was lower in BIS-monitored patients, although not statistically significant. Incidence of stroke and myocardial infarction was similar between groups.
- In summary, BIS monitoring was associated with overall lower rates of explicit intraoperative awareness (limited to patients classified at higher risk of awareness, and non-significant effects in the subgroup

of patients receiving only inhaled anaesthesia), lower general anaesthetic consumption and shorter recovery times (e.g. PACU discharge, time to extubation, time to eye opening). Generally, there was little difference between BIS and standard clinical monitoring in complications arising from excessive anaesthetic dose (e.g. nausea, vomiting and cognitive dysfunction). Caution is advised in the interpretation of the results as not all outcomes appeared to be adequately statistically powered, and there was significant heterogeneity. There was much variation between the trials in terms of patient characteristics and surgical procedures.

Characteristics of included studies: E-Entropy

Study populations

Two of the seven E-entropy trials were conducted with children, with median age 4–6 years (range 3–12 years).[54,56] The remaining five trials were in adults, with the mean age of patients ranging from 33 years[55] to 69 years.[58] The trials varied in their sex composition. One trial was entirely on adult women,[55] whereas another trial was almost entirely on young boys (the trial included 12% girls in one study arm only).[56] One trial included more elderly men than women (men–women ratio approximately 4 : 1),[58] whereas another trial included more middle-aged women than men (male–female ratio approximately 1 : 3). The remaining three E-Entropy trials included a more even balance of males and females.[54,61,62] In all seven trials the mean body weight of patients appeared to be within the normal range, with mean weights ranging from 16 kg to 22 kg in the child studies and from 65 kg to 82 kg in the adult studies. One trial was conducted at six centres in three countries (Finland, Sweden and Norway).[57] The remaining trials appeared to be single-centre studies (not explicitly stated in two trials) that were each carried out in one country: Germany,[55,62] France,[61] India,[56] South Korea[54] and Taiwan.[58] None of the E-Entropy trials reported the ethnicity of their participants.

Four of the E-Entropy trials were in patients undergoing a mix of abdominal, urological, gynaecological and/or orthopaedic surgical procedures,[56,61,62] which also included breast and thyroid surgery in one trial.[57] One trial specifically involved children undergoing tonsillectomy or adenoidectomy.[54] Another trial was carried out specifically in women undergoing laparoscopic gynaecological procedures.[55] The remaining trial focused on total knee replacement surgery.[58] Only one of the E-Entropy trials was clearly limited to day surgery patients.[56] None of the seven trials identified any specific risk factors for intraoperative awareness among their populations and none reported whether or not patients had any comorbidities that affect EEG monitoring. However, all the E-Entropy trials stated that they excluded patients with any history of cerebrovascular and/or neurological disorders. The ASA grade of patients was I–II in four of the trials,[54–56,58] and I-III in the remaining three trials.[57,61,62] The proportion of grade III patients varied by study groups within these three trials, ranging 1–3%,[57] 11–15%[61] and 3–26%.[62]

Technologies

Four of the seven E-Entropy trials reported that they used the E-Entropy module manufactured by GE Healthcare,[55,57,61,62] and six of the trials reported that they used the S/5TM monitor (Datex-Ohmeda).[54–58,61] Very little other information about the modules and monitors was provided: only one trial mentioned the version of the S/5 monitor used (Avance),[56] and none of the studies stated the version of the E-Entropy algorithm software used.

The target E-Entropy values during anaesthesia maintenance were mostly in the range 40–65. Four trials specified target ranges for state entropy, which were either 40–60[54,55] or 45–65.[56,57] A further trial specified a specific state entropy target of 50.[62] The remaining two trials specified target ranges for both state entropy and response entropy, which were 35–45[58] and 40–60.[61] Four of the trials that specified target values for state entropy permitted an increase in the state entropy value during the last 15 minutes of surgery. During this period, the target values were specified as 60,[62] 65–70,[56] 'ideally 65, but not > 70'[57] and '> 60 acceptable'.[55] In addition to the target values of state and response entropy, three trials also specified target values of the difference between response entropy and state entropy: these were < 10 in two trials[55,57] and 5–10 in the remaining trial.[58]

Two of the seven E-Entropy trials reported that E-Entropy monitoring for anaesthesia delivery was done in conjunction with monitoring haemodynamic changes. One of these trials specified that heart rate and blood pressure were to be kept within ± 20% of their baseline (preoperative visit) values.[57] The second trial stated that E-Entropy was used to guide anaesthesia unless (unspecified) haemodynamic changes of 30% persisted for > 5 minutes.

In addition to titrating anaesthesia to maintain the specified target entropy values, two trials specified corrective action if target values were exceeded. One trial specified intermittent provision of a sufentanil bolus if the response entropy–state entropy difference exceeded 10 for > 2 minutes.[61] The other trial specified administration of a propofol bolus if the state entropy value increased suddenly above 65.

In all seven of the E-Entropy trials, the E-Entropy monitoring was initiated in the operating theatre. Two trials stated[57] or implied[58] that E-Entropy monitoring was started before anaesthesia induction, and two trials stated that E-Entropy monitoring began after anaesthesia induction.[55,56] The remaining three trials did not report whether E-Entropy monitoring commenced before or after anaesthesia induction.

Comparators

Standard clinical monitoring was based on blood pressure and heart rate in three trials.[54,57,61] As well as blood pressure and heart rate, a further two trials also monitored sweating, lacrimation or movement,[62] or coughing, chewing, grimacing or purposeful movement.[55] The remaining trials monitored heart rate, mean arterial pressure and lacrimation, and either movement in response to surgical stimulation,[56] or sweating, flushing or wrinkling of frontal facial muscles, together with monitoring the end-tidal anaesthetic concentration.[58] Quantitative thresholds for the clinical parameters that were used to guide anaesthesia titration were specified in five of the seven E-Entropy trials.[54–58]

In addition to titrating anaesthesia according to the clinical parameters, in one trial[58] the ETAC was adjusted to maintain mean arterial pressure and heart rate fluctuations to within ± 30% of the baseline values. In another trial, i.v. fentanyl was given if clinical parameters were not stabilised after increasing the anaesthetic concentration to 1.3 MAC.[56]

Anaesthetic agents and protocols

Three of the seven trials used i.v. propofol for anaesthesia induction.[55,61,62] One trial used i.v. propofol with alfentanil analgesic for induction.[57] A further trial employed propofol if patients had an i.v. line, but otherwise used inhaled sevoflurane for induction.[56] The remaining two trials both used inhaled sevoflurane for induction in all their patients.[54,58]

For maintenance of anaesthesia, three trials used inhaled sevoflurane,[54,58,61] and one trial used inhaled isoflurane.[56] The remaining trials used i.v. delivery of propofol,[62] propofol and remifentanil,[55] or propofol and alfentanil analgesic.[57]

Overall, two trials used the same inhaled agent (sevoflurane) for both induction and maintenance;[54,58] three trials used i.v. agents (all included propofol) for both induction and maintenance;[55,57,62] and two trials used an i.v. anaesthetic for induction followed by an inhaled anaesthetic for maintenance.[56,61]

Regional anaesthesia was only clearly reported in one of the E-Entropy trials, in which a caudal block was placed with bupivacaine.[56] Two trials stated that regional anaesthesia was not used.[58,61] One trial referred to regional anaesthesia in the publication abstract but did not provide details.[62] The remaining three trials did not refer to regional anaesthesia.

One of the E-Entropy trials stated that analgesics were not used during induction or maintenance of anaesthesia, although ketorolac was used after anaesthetic cessation.[54] One trial used i.v. sufentanil during induction and maintenance, with morphine during the last 15 minutes of surgery, followed by paracetamol, nefopam or non-steroidal anti-inflammatory drugs postoperatively.[61] Two trials used fentanyl

during anaesthesia maintenance. Of these, one also used fentanyl and lidocaine during induction,[58] whereas the other used fentanyl postoperatively, according to the patient's pain score.[56] One trial used piritramide during the last 15 minutes of surgery only.[55] The remaining two trials did not refer to analgesia either during induction, maintenance or post surgery.[57,62]

Premedication was reported in five of the E-Entropy trials. The agents used were oral hydroxyzine,[61] oral midazolam alone,[62] oral midazolam with a benzodiazepine,[55] i.v. midazolam[54] and oral diazepam (in five of six study centres).[57] The remaining two trials did not specify whether or not premedication was used.

All of the E-Entropy studies except one[56] used muscle relaxants. The muscle relaxants were atracurium,[58,61,62] rocuronium[54,55] or were not specified a priori but were chosen at the anaesthetist's discretion when needed.[57]

In five trials anaesthesia was administered in the operating theatre.[56–58,61,62] The two remaining trials did not report where anaesthetics were administered.

The mean duration of anaesthesia was reported in six studies and ranged from 64.3 minutes for tonsillectomy or adenoidectomy procedures in children[54] to 190.8 minutes for general surgical procedures in adults.[61] The remaining study reported median duration of anaesthesia which was 68–72 minutes (range 32–180 minutes) for lower abdominal or urological surgical procedures in children.[56]

Duration of the surgery itself was reported less precisely than the duration of anaesthesia. Surgical duration was described as a minimum of 1 hour,[55,61] approximately 1.5 hours,[58] a mean of 41.4–48.1 minutes,[54] or a median of 29–30 minutes (range 15–95 minutes)[56] or was not reported.[57,62]

The training and experience of the anaesthetists in E-Entropy module use was reported in four of the seven E-Entropy trials.[55,57,61,62] One trial stated that anaesthetists were allowed to accustom themselves to the use of E-Entropy monitoring for 3 weeks, and all participants had substantial previous experience with EEG-based depth of anaesthesia monitors.[57] In the remaining three trials the descriptions provided for training or experience were only superficial: 'more than 3 months of routine use';[61] 'experienced anaesthesiologist';[62] and 'anaesthesia was supervised by an experienced staff anaesthetist'.[55]

Outcomes

Anaesthetic consumption was the primary outcome in four of the seven E-Entropy studies (*Table 15*).[54,58,61,62] The method of assessing anaesthetic consumption was by weighing the vaporiser,[61] measuring the end-tidal concentration,[54] using data from the S/5 anaesthetic delivery system[58] or was not reported.[62] In the remaining three trials the primary outcomes were time to eye opening[55,56] and time to response to a verbal command,[57] after cessation of anaesthesia.

The most frequently reported outcomes overall for which quantitative results were reported were: anaesthetic consumption (a primary outcome in four trials[54,58,61,62] and a secondary outcome in three trials[55–57]); entropy values (a secondary outcome in all seven trials); time to eye opening (a primary outcome in two trials[55,56] and a secondary outcome in four trials[54,57,61,62]); intraoperative awareness (a secondary outcome in all except one trial[56]); haemodynamic profiles (a secondary outcome in all except one trial[62]); time to extubation (a secondary outcome in three trials[54,57,61]); and postoperative pain (a secondary outcome in two trials[55,56]). Other outcomes that were reported quantitatively in one trial each were postoperative pain, analgesia consumption, PONV, time to recovery based on Aldrete or Steward scores, time spent with adverse haemodynamic profiles, probability of emergence and (in a study with children) parental satisfaction. Some of the trials provided only a narrative report of outcomes. These outcomes were not extracted from the primary trials as no estimates of effect or variance could be determined. For example, two trials[57,58] stated narratively that pain scores, analgesic use and incidence of PONV did not differ between E-Entropy and clinical practice groups but no quantitative results were reported for these outcomes and so these are not included in *Table 15*.

TABLE 15 E-Entropy study outcomes

Outcome	Study						
	Aime *et al.*[61]	[a]Choi *et al.*[54]	Ellerkmann *et al.*[62]	Gruenewald *et al.*[55]	[a]Talawar *et al.*[56]	Vakkuri *et al.*[57]	Wu *et al.*[58]
Anaesthetic consumption	P	P	P	✗	✗	✗	P
Intraoperative awareness	✗	✗	✗	✗		✗	✗
Analgesic consumption	✗						
Time to response to commands						P	
Time to eye opening	✗	✗	✗	P	P	✗	
Time to extubation	✗	✗				✗	
Time to recovery of orientation		✗				✗	
Time to PACU admission or discharge					✗	✗	
Monitoring device values	✗	✗	✗	✗	✗	✗	✗
Postoperative nausea and vomiting				✗			
Parental satisfaction				✗			
Treatment of haemodynamic events	✗						
Haemodynamic profiles	✗	✗		✗	✗	✗	✗
% time with adverse haemodynamic profiles	✗						
Time to complete recovery (Aldrete score of at least 9)		✗					
Recovery score (modified Steward recovery score)					✗		
Postoperative pain				✗	✗		
Patient satisfaction				✗			

✗, stated secondary outcome measure/not stated whether primary or secondary outcome measure; P, primary outcome measure.

a Study of children.

Three of the six trials that measured intraoperative awareness employed versions of standard patient questionnaires published by Brice and colleagues[24] (two studies[57,61]) or Nordström and colleagues[96] (one study[62]). The three remaining trials stated only that intraoperative recall was assessed by independent nurses;[54] patients were questioned about memory and awareness;[55] or the level of awareness was assessed.[58] Four trials reported the timing of the intraoperative awareness assessments, which were 24 hours after surgery,[55] on the first postoperative day,[62] in the PACU and on the first day post surgery,[57] or on the first and third days post surgery.[61] The remaining two trials did not specify the timing of the awareness outcome assessments. No further details of the methods for assessing intraoperative awareness were reported.

Length of follow-up was relatively short in all the trials, being 1 day post surgery (for intraoperative awareness) in three trials,[54,55,57] 3 days post surgery (for intraoperative awareness) in three trials,[58,61,62] and only 2 hours post surgery (for pain assessment) in the remaining trial.[56] The duration of follow-up would not have been adequate for detecting delayed onset of awareness recall, which may occur more than 1 week post surgery.

Assessment of outcomes: E-Entropy

Intraoperative awareness

Only one case of intraoperative awareness was reported in the six trials that measured this outcome (*Table 16*). This was experienced by an adult woman in the standard clinical practice group of the trial by Gruenewald and colleagues.[55] It should be noted that the sample sizes of these studies may have been too small to detect rare events such as intraoperative awareness.

Anaesthetic consumption

Four trials that assessed volatile anaesthetic consumption either as the primary outcome for sevoflurane[54,58,61] or a secondary outcome for isoflurane,[56] all demonstrated statistically significant reductions in the E-Entropy-guided anaesthesia group compared with the standard clinical monitoring group (*Table 17*). In the trial by Aime and colleagues,[61] the rates of sevoflurane consumption, but not the total amount consumed, were significantly lower in the E-Entropy group. In this trial[61] the difference in sevoflurane consumption rates between groups was more pronounced when the consumption rate was normalised to patients' body weight.

Three trials that assessed consumption of i.v. anaesthetics[55,57,62] showed mixed results (*Table 17*). Propofol consumption in the E-Entropy group was statistically significantly lower than in the standard clinical practice group in two trials that assessed anaesthetic consumption as secondary outcomes,[55,57] but not in

TABLE 16 Intraoperative awareness during E-Entropy monitoring

Study	Entropy	Standard clinical monitoring	Mean difference (95% CI), *p*-value
Aime *et al.*,[61] *n/N* (%)	0/40 (0)	0/60 (0)	NR
[a]Choi *et al.*,[54] *n/N* (%)	0/39 (0)	0/39 (0)	NR
Ellerkmann *et al.*,[62] *n/N* (%)	0/30 (0)	0/30 (0)	NR
Gruenewald *et al.*,[55] *n/N* (%)	0/37 (0)	1/35 (2.8)	NR
[b]Vakkuri *et al.*,[57] *n/N* (%)	0/160 (0)	0/160 (0)	NR
Wu *et al.*,[58] *n/N* (%)	0/34 (0)	0/31 (0)	NR

NR, not reported.
a Study of children.
b Number reported only after attrition.

TABLE 17 Consumption of anaesthetic during E-Entropy monitoring

Study	Entropy	Standard clinical monitoring	Mean difference (95% CI), *p*-value
Volatile anaesthetic consumption (sevoflurane), mean ± SD vaporiser weight change			
Aime *et al.*[61]			
Total (g)	22.8 ± 14.4	25.6 ± 17.2	$p = 0.49$
Rate (g/hour)	7.8 ± 3.4	9.4 ± 5.6	$p = 0.07$
[a]Rate normalised (g/kg/hour)	0.10 ± 0.05	0.14 ± 0.09	$p = 0.003$
Volatile anaesthetic consumption (sevoflurane), mean ± SD end-tidal sevoflurane concentration (%)			
[b] Choi *et al.*[54]	2.2 ± 0.3	2.6 ± 0.4	$p < 0.05$
Volatile anaesthetic consumption (sevoflurane), mean ± SD total sevoflurane consumption recorded by S/5 monitor			
Wu *et al.*[58]			
Total consumption (ml)	27.79 ± 7.4	31.42 ± 6.9	$p = 0.023$
Volatile anaesthetic consumption (isoflurane), mean end-tidal isoflurane concentration (%)			
[b]Talawar *et al.*[56]			
Immediately before LMAI	0.81	1.24	$p < 0.05$
15 seconds after LMAI	0.78	1.24	$p < 0.05$
15 seconds after caudal analgesia	0.69	0.84	$p < 0.05$
15 seconds after skin incision	0.68	0.78	$p < 0.05$
5 minutes after skin incision	0.68	0.79	$p < 0.05$
Immediately before LMAR	0.35	0.38	$p \geq 0.05$
Intravenous anaesthetic consumption (propofol and remifentanil), mean ± SD consumption rate and number (%) requiring propofol bolus based on E-Entropy			
Ellerkmann *et al.*[62]			
Propofol (μg/kg/minute)	106 ± 24	101 ± 22	$p = 0.27$
Remifentanil (μg/kg/minute)	0.08 ± 0.02	0.09 ± 0.02	$p = 0.56$
Requiring bolus, *n/N* (%)	12/30 (40)	10/30 (33)	NR
Gruenewald[55]			
Propofol (μg/kg/minute)	81 ± 22	95 ± 14	$p < 0.01$
Remifentanil (μg/kg/minute)	0.46 ± 0.08	0.39 ± 0.08	$p < 0.001$
Intravenous anaesthetic consumption (propofol and alfentanil), median (range) consumption rate			
[c]Vakkuri[57]			
Propofol (mg/kg/minute)	0.10 (0.04–0.23)	0.11 (0.03–0.21)	$p < 0.001$
Alfentanil (μg/kg/minute)	0.60 (0.12–2.2)	0.57 (0.16–1.6)	$p = 0.54$

LMAR, laryngeal mask airway removal; LMAI, laryngeal mask airway insertion; NR, not reported.

a Normalised to patient body weight and anaesthetic duration.

b Study of children.

c Unclear whether data are for whole operation or last 15 minutes (*p*-value the same for both).

a trial that assessed anaesthetic consumption as the primary outcome.[62] Remifentanil consumption was significantly higher in the E-Entropy group in one trial that assessed this as a secondary outcome,[55] but did not differ between groups in the trial that assessed this as the primary outcome.[62] Alfentanil consumption, assessed as a secondary outcome in one trial, did not differ significantly between the study groups.[57]

The trials that assessed anaesthetic consumption measured outcomes in different ways, expressed their outcomes in different units (total consumption or rates) and, as noted above, differed in the patient populations that they included. These differences would preclude the meaningful pooling of the anaesthetic consumption outcomes that were reported (*Table 17*).

Time to recovery from anaesthesia

Results are summarised in *Table 18* for the trials that reported time to eye opening;[54–57,61,62] extubation;[54,57,61] spontaneous breathing;[57] recovery of orientation;[54,57] response to commands;[57] recovery defined by Aldrete score;[54] and recovery defined by modified Steward score.[56]

Time to eye opening was significantly shorter, by approximately 2–4 minutes, in the E-Entropy group than in the standard clinical practice group in two of six trials.[56,57] One of these assessed this as a primary outcome in children[56] and the other assessed it as a secondary outcome in adults.[57] In the remaining four trials[54,55,61,62] (one of which specified this as a primary outcome[55]) the time to eye opening did not differ between the study groups (*Table 18*).

Time to extubation (a secondary outcome) was shorter, by approximately 3–4 minutes, in the E-Entropy group than in the standard clinical monitoring group in all three trials that assessed this outcome.[54,57,61] The differences were stated as statistically significant in two of the trials[54,57] but statistical significance was not reported in the remaining trial[61] (see *Table 18*).

The times to spontaneous breathing (a secondary outcome);[57] recovery of orientation (a secondary outcome);[54,57] response to commands (a primary outcome);[57] and recovery defined by an Aldrete score of at least 9 (a secondary outcome)[54] were each significantly shorter in the E-Entropy group than the standard clinical practice group in the two trials[54,57] that reported these outcomes (see *Table 18*). However, the time to recovery as defined by reaching a Steward score of 6 (a secondary outcome) did not differ between the study groups in one trial that assessed this outcome.[56]

Outcomes related to postanaesthesia care unit stay

The time from discharge from the operating room to the PACU was shorter by approximately 3–4 minutes in the E-Entropy group than the standard clinical practice group in the two trials that monitored these outcomes[56,57] (*Table 19*). The differences in both trials were statistically significant, although only marginally so in one of the trials.[56]

The time to discharge from the PACU was shorter in the E-Entropy group than the standard clinical monitoring in the only trial that assessed this outcome,[57] although the difference was not statistically significant. The time from which discharge from the PACU was measured was not reported, however, which makes interpretation of this outcome unclear[57] (see *Table 19*).

Postoperative pain

Two trials reported postoperative pain, using different rating scales (*Table 20*). Pain was assessed as a score on a 0–10 scale[55] or using the Children's Hospital of Eastern Ontario Pain Score (CHEOPS).[56] Pain scores were significantly lower in the E-Entropy group than standard clinical practice for the adult population.[55] In the paediatric population, the CHEOPS scores were significantly lower in the E-Entropy group at 60, 90 and 120 minutes after arrival in the PACU but not at 30 minutes after arrival.[56]

TABLE 18 Time to recovery from anaesthesia (before discharge to PACU)

Study	Entropy	Standard clinical monitoring	Mean difference (95% CI), *p*-value
Time (minutes) to eye opening, mean ± SD or median (range) [interquartile range] time since cessation of anaesthetic (or time from last suture[62])			
Aime *et al.*[61]	7.6 ± 4.1	7.2 ± 4.7	NR
[a]Choi *et al.*[54]	14.3 ± 3.6	18.0 ± 3.3	Stated not significant
Ellerkmann *et al.*[62]	9.2 ± 3.9	7.3 ± 2.9	Not reported
[b]Gruenewald *et al.*[55]	3 (0–9) [1–5]	4 (0–14) [3–6]	Stated not significant
[a,b]Talawar *et al.*[56]	8.2 ± 4.49, 7 (3–18)	10.96 ± 3.86, 10 (5–21)	2.72 (0.34–5.1), $p = 0.017$
Vakkuri *et al.*[57]	6.08 (0.15–37.5)	10.8 (2.23–43.2)	$p < 0.001$
Time (minutes) to extubation, mean ± SD or median (range) time since cessation of anaesthetic (or start time not reported[57])			
Aime *et al.*[61]	11.5 ± 5.8	14.2 ± 9.0	NR
[a]Choi *et al.*[54]	8.3 ± 1.4	11.9 ± 2.5	$p < 0.05$
Vakkuri *et al.*[57]	5.80 (3.00–27.3)	9.16 (1.67–32.3)	$p < 0.001$
Time (minutes) to spontaneous breathing, median (range) (start time not reported)			
Vakkuri *et al.*[57]	4.74 (0.00–18.0)	7.07 (–1.00–28.5)	$p < 0.001$
Time (minutes) to recovery of orientation, mean ± SD or median (range) time since cessation of anaesthetic (or start time not reported[57])			
[a]Choi *et al.*[54]	18.2 ± 4.0	23.3 ± 5.0	$p < 0.05$
Vakkuri *et al.*[57]	10.3 (1.17–48.7)	15.1 (4.08–113)	$p < 0.001$
Time (minutes) to response to commands, median (range) time to hand squeezing (start time not reported)			
Vakkuri *et al.*[57]	8.60 (1.17–47.4)	12.7 (2.43–48.1)	$p < 0.001$
Time (minutes) to complete recovery (Aldrete score ≥9), mean ± SD time since cessation of anaesthetic			
[a]Choi *et al.*[54]	24.3 ± 7.3	28.8 ± 5.7	$p < 0.05$
Time (minutes) to recovery (Steward score of 6), mean ± SD time since cessation of anaesthetic			
[a]Talawar *et al.*[56]	7.08 ± 3.78, 6 (1–15)	8.36 ± 4.8, 8 (2–24)	1.3 (–1.2–3.7), $p = 0.464$

NR, not reported.
a Study of children.
b Primary outcome.

TABLE 19 Time for discharge to/from PACU

Study	Entropy	Standard clinical monitoring	Mean difference (95% CI); *p*-value
Time (minutes) from discharge from operating room to PACU admission, mean ± SD or median (range) time since cessation of anaesthetic[56] or since discharge from operating room[57]			
[a]Talawar *et al.*[56]	15.32 ± 6.6, 15 (5–31)	19.32 ± 7.12, 19 (10–40)	4.0 (0.07–7.9), $p = 0.045$
Vakkuri *et al.*[57]	10.3 (3.83–42.4)	13.0 (5.00–49.8)	$p < 0.001$
Time (minutes) to discharge from PACU, median (range) – not stated whether time since discharge from operating room or since admission to PACU			
Vakkuri *et al.*[57]	134 (50–1293)	150 (7–1020)	$p = 0.21$

a Study of children.

Analgesic consumption

Only one E-Entropy trial assessed analgesic consumption.[61] Consumption of sufentanil was slightly lower in the E-Entropy group than the standard clinical monitoring group during both induction and maintenance of anaesthesia, but the differences were not statistically significant (*Table 21*).

Postoperative nausea and vomiting

One trial that assessed PONV after arrival in the recovery room[55] reported similar frequencies in the E-Entropy and standard clinical monitoring that did not differ significantly (*Table 22*).

In addition to the outcomes reported above, the E-Entropy trials reported that the following outcomes did not differ between E-Entropy and standard clinical practice groups (data not extracted): patient satisfaction scores;[55] parent satisfaction scores for children at 24 hours post surgery;[55] time spent by patients with adverse haemodynamic profiles;[61] and treatment for haemodynamic events.[61]

TABLE 20 Postoperative pain

Study	Entropy	Standard clinical monitoring	Mean difference (95% CI), *p*-value
Pain intensity score on arrival in recovery room (0–10 scale; no other details), median (range) [interquartile range]			
Gruenewald *et al.*[55]	6 (2–10) [4–7]	4 (1–10) [3–5]	$p = 0.03$
Pain intensity score based on CHEOPS scale, mean (standard error)			
[a,b]Talawar *et al.*[56]			
After 30 minutes in PACU	4.88 (0.319)	4.76 (0.09)	0.12 (–0.53 to 0.77), $p = 0.71$
After 60 minutes in PACU	4.48 (0.10)	4.76 (0.08)	–0.28 (4.59 to 4.92), $p = 0.01$
After 90 minutes in PACU	4.56 (0.10)	4.76 (0.08)	–0.2 (4.59 to 4.92), $p = 0.01$
After 120 minutes in PACU	4.88 (0.21)	5.44 (0.33)	–0.56 (4.77 to 6.09), $p = 0.01$

a Study of children.

b The confidence intervals appear to have been calculated differently for 30 minutes compared with 60, 90 and 120 minutes and their interpretation is unclear (not explained in the primary publication).

TABLE 21 Analgesic consumption during E-Entropy monitoring

Study	Entropy	Standard clinical monitoring	Mean difference (95% CI), *p*-value
Sufentanil consumption per patient, mean ± SD			
Aime *et al.*[61]			
Induction dose (μg/kg)	0.21 ± 0.05	0.23 ± 0.06	$p = 0.18$
Maintenance consumption (μg/hour)	13.6 ± 6.1	14.9 ± 8.3	$p = 0.66$
Maintenance consumption (μg/kg/hour)	0.18 ± 0.09	0.22 ± 12	$p = 0.26$

TABLE 22 Postoperative nausea and vomiting

Study	Entropy	Standard clinical monitoring	Mean difference (95% CI), *p*-value
Nausea and vomiting on arrival in recovery room,* n/N *(%)			
Gruenewald *et al.*[55]	15/37 (41)	13/35 (37)	Stated not significant

Summary of E-Entropy assessment

- Six trials monitored intraoperative awareness in adults and children receiving different volatile and i.v. anaesthetics. Only one case of awareness occurred, in the standard clinical practice group of one trial. However, sample sizes were relatively small in these trials.
- Four trials monitored consumption of volatile anaesthetic (three monitored sevoflurane as a primary outcome, one monitored isoflurane as a secondary outcome). Consumption was significantly lower in the E-Entropy monitoring than standard clinical practice groups of all trials, with the proviso that in one of these trials the difference in sevoflurane consumption was statistically significant for rates of consumption but not for total anaesthetic dose.
- Three trials that monitored consumption of i.v. anaesthetic yielded mixed results. Trials that monitored consumption of propofol, remifentanil and alfentanil as primary outcomes found no statistically significant differences between the study groups. However, significantly lower consumption of propofol and remifentanil in the E-Entropy group was reported in trials that assessed these as secondary outcomes.
- Time to eye opening was significantly shorter in the E-Entropy group than the standard clinical practice group in two of six trials, one of which assessed this as a primary outcome, but did not differ in the remaining four trials.
- Time to extubation (a secondary outcome) was shorter in the E-Entropy group than the standard practice group in all three trials that assessed this outcome. The differences were stated as statistically significant in two of the trials but statistical significance was not reported in the remaining trial.
- The times to spontaneous breathing (a secondary outcome), recovery of orientation (a secondary outcome), response to commands (a primary outcome) and recovery defined by an Aldrete score of at least 9 (a secondary outcome) were each significantly shorter in the E-Entropy group than in the standard clinical practice group. Except for time to orientation (two trials), these outcomes were reported by only one trial each. The time to recovery as defined by reaching a Steward score of 6 (a secondary outcome) did not differ between the study groups in one trial that assessed this outcome.
- The limited evidence available (from two trials which assessed secondary outcomes only) suggests that E-Entropy monitoring favours shorter time to discharge to and from the PACU, but it is unclear whether or not the time gains are clinically important.
- No firm conclusions can be drawn about effects of E-Entropy monitoring on postoperative pain because the only two trials that assessed this used different rating scales, and the effect of E-Entropy monitoring on pain scores was temporally variable in one of the trials. Analgesic consumption and frequency of PONV were assessed in one trial each and did not differ between the E-Entropy and standard clinical practice groups. Postoperative pain, nausea and vomiting, and analgesic consumption were assessed only as secondary outcomes in these trials.
- In summary, compared with standard clinical monitoring, E-Entropy monitoring favoured: lower consumption of volatile anaesthetics and some, but not all, i.v. anaesthetics; and shorter times to recovery and discharge to and from the PACU, assessed by various measures. E-Entropy monitoring had no consistent impact on other outcomes that were monitored, including intraoperative awareness, but the small sample sizes in the trials may not have provided adequate statistical power to detect meaningful differences in rare events. Pooled effect estimates would not be estimable reliably for these outcomes, because of the uniqueness of the individual studies (which included different populations in terms of age, sex and ethnicity, undergoing different surgical procedures) and differences between the trials in the way that outcomes were assessed and reported. Also, the majority of the outcomes were secondary and may not have been adequately powered statistically to detect clinically relevant differences between the E-Entropy and standard clinical practice groups.

Characteristics of included studies: Narcotrend

Study populations

In all trials of Narcotrend, the study population was adults (mean age 40–50 years) and 33–50% of participants were males for the three studies reporting sex. Mean weight ranged from 60 kg to about

84 kg. All studies appeared to be single-centred studies with three conducted in Germany and one in China.[59] Ethnicity of participants was not reported in any study.

The type of surgery was minor orthopaedic surgery,[63,64] microwave coagulation for liver cancer,[59] and all kinds of elective surgery, including surgery for 'malignoma' and peripheral vascular surgery.[60] No trial reported risk factors for awareness. Comorbidities were reported in two trials:[59,60] hypertension was reported in both of these and one trial[60] also reported cardiac arrhythmia, diabetes type II, asthma and miscellaneous comorbidities. Three trials included the number of participants with ASA grade I, II or III, with most grade II and fewest grade III; the fourth trial[59] reported only that participants were ASA grade II or III.

Technologies

The Narcotrend monitor with software version 2.0 AF was used in three trials,[60,63,64] whereas in the fourth trial[59] no details of the software version are reported. Two trials report using the MonitorTechnik (Germany) with Blue Sensor (Denmark).[59,60]

The Narcotrend target value during maintenance anaesthesia was D_0 and then adjusted to C_1 15 minutes before the expected end of surgery in two studies,[63,64] and D_2–E_0 during maintenance adjusted to D_0–D_1 10 minutes before the end of surgery in one study.[59] In the fourth study[60] the Narcotrend target value was D_2–E_0 with no further details given. The two studies[59,60] using Narcotrend target values of D_2–E_0 therefore used deeper levels of anaesthesia and hypnosis than the other two studies. Monitoring started in the operating theatre in two studies,[63,64] in the computed tomography department where surgery took place in one study[59] and was not reported in the fourth study.[60]

Only one trial[59] explicitly stated that observational indices of electrocardiography (ECG), heart rate and mean arterial blood pressure were continuously monitored alongside Narcotrend scores. The other three studies did not explicitly state whether or not standard clinical monitoring took place in addition to Narcotrend. However, as signs of inadequate anaesthesia were based on vital signs and clinical parameters it can be assumed that it did. For example, signs of inadequate anaesthesia were hypertension, tachycardia or patient movement, eye opening, swallowing, grimacing, lacrimation and sweating.[63,64] Vital clinical parameters of heart rate, pulse oximetry readings, rectal temperature and end-expiratory carbon dioxide were continuously measured in the fourth study.[60]

Comparators

Standard clinical continuous monitoring included heart rate, systemic arterial blood pressure, respiratory rate, oxygen saturation and end-tidal concentrations of carbon dioxide[63,64] plus end-tidal desflurane[64] and heart rate, pulse oximetry readings, rectal temperature and end-expiratory carbon dioxide.[60] In one study[59] heart rate, blood pressure and body movement were used for monitoring.

Anaesthetic agents and protocols

Three studies used total i.v. anaesthesia: two[60,63] used propofol–remifentanil for induction and maintenance anaesthesia; one used propofol–fentanyl induction and propofol anaesthesia maintenance.[59] The fourth study used desflurane–remifentanil anaesthesia.[64] Regional anaesthesia was not reported in any of the studies. Premedication was used in three studies in the form of midazolam[60,64] and diazepam.[63] Analgesia included metamizol with sodium chloride,[63,64] fentanyl[59] and novaminsulfone, piritramide or morphine.[60] Muscle relaxants used included atracurium,[64] cisatracurium[63] and rocuronium.[60]

Mean duration of anaesthesia ranged from 113 to 125 minutes,[64] from 108 to 127 minutes,[63] from 88 to 91 minutes[59] and from 105 to 111 minutes[60] in the four trials, with no significant differences between groups within each study. Duration of surgery was not reported in any study. Three studies[60,63,64] reported that all patients were anaesthetised by the same experienced anaesthesiologist, one of which mentions specific experience in Narcotrend.[63] No details are given for the length of experience/training of the anaesthetist in the fourth study.[59]

Outcomes

The primary outcome (statistically powered) specified in three trials was time to eye opening[63,64] and time to extubation[60] (*Table 23*). Time to tracheal extubation was also an outcome in two other studies.[63,64] All four studies[59,60,63,64] report anaesthetic consumption and intraoperative awareness. Other reported outcomes include time to arousal time[59] (defined as the time between cessation of drugs and the patient being able to open their eyes on command) and time to recovery of orientation (defined as the time between a patient opening their eyes on command and the restoration of orientation).[59] Two studies[63,64] report time to discharge to the PACU and two report PONV.[59,60]

Assessment of outcomes: Narcotrend

Intraoperative awareness

No patients in any of the trials of Narcotrend reported intraoperative awareness as explicit memory during anaesthesia, although two patients (8%) receiving Narcotrend anaesthetic monitoring recalled dreaming during anaesthesia.[60]

Anaesthetic consumption

Three studies report consumption of propofol; two[59,63] found a statistically significant reduction in the group receiving Narcotrend monitoring compared with standard clinical monitoring, whereas the third[60] found no difference in consumption between groups (*Table 24*).

Three studies reported remifentanil consumption and all found no statistically significant difference between Narcotrend and standard clinical monitoring.[60,63,64]

Desflurane consumption per patient was not different between the Narcotrend monitoring group and standard anaesthetic practice, although desflurane consumption per patient per minute was statistically significantly lower in the Narcotrend group.[64]

Time to arrival at postanaesthetia care unit

Two studies reported time to arrival at PACU and found statistically significantly shorter times in the Narcotrend monitoring group compared with the standard care monitoring group[63,64] (*Table 25*).

TABLE 23 Narcotrend study outcomes

	Study			
Outcomes	**Kreuer *et al.*[64]**	**Kreuer *et al.*[63]**	**Lai *et al.*[59]**	**Rundshagen *et al.*[60]**
Anaesthetic consumption	✗	✗	✗	✗
Intraoperative awareness	✗	✗	✗	✗
Analgesic consumption			✗	✗
Time to response to commands			✗	
Time to eye opening	P	P		
Time to extubation	✗	✗		P
Time to recovery of orientation			✗	
Time to arrival at PACU	✗	✗		
PONV			✗	✗

✗, stated secondary outcome measure/not stated whether primary or secondary outcome measure; P, primary outcome measure.

TABLE 24 Anaesthetic consumption

Study	Narcotrend	Standard clinical monitoring	*p*-value
Propofol consumption per patient			
Kreuer *et al.*[63]			
Mean ± SD (mg)	721.3 ± 401.2	970.5 ± 384.4	$p<0.05$
Mean ± SD (mg/kg/hour)	4.5 ± 1.1	6.8 ± 1.2	$p<0.001$
Lai *et al.*[59]			
Mean ± SD (mg)	380 ± 35	460 ± 30	$p<0.01$
Rundshagen *et al.*[60]			
Mean ± SD (µg/kg/minute)	0.093 ± 0.042	0.114 ± 0.035	$p=0.089$
Remifentanil consumption per patient			
Kreuer *et al.*[64]			
Mean ± SD normalised remifentanil infusion rate (µg/kg/minute)	0.22 ± 0.06	0.23 ± 0.07	NS
Kreuer *et al.*[63]			
Mean ± SD normalised remifentanil infusion rate (µg/kg/minute)	0.21 ± 0.07	0.20 ± 0.07	NS
Rundshagen *et al.*[60]			
Mean ± SD remifentanil dose (µg/kg/minute)	0.31 ± 0.10	0.34 ± 0.11	NS
Desflurane consumption per patient			
Kreuer *et al.*[64]			
Mean ± SD (mg)	4655.9 ± 2891.7	5547.3 ± 2396.4	NS
Mean ± SD (mg/minute)	374.6 ± 124.2	443.6 ± 71.2	$p<0.05$

NS, not statistically significant.

TABLE 25 Time to arrival at PACU

	Time to arrival at PACU (minutes)		
Study	Narcotrend	Standard clinical monitoring	*p*-value
Kreuer *et al.*,[64]			
Mean ± SD	8.0 ± 1.9	9.4 ± 2.4	$p<0.05$
Kreuer *et al.*[63]			
Mean ± SD	6.6 ± 2.8	12.4 ± 5.7	$p<0.001$

Time to eye opening

Time to eye opening was the primary outcome in two trials and results between the studies differ (*Table 26*). One trial[64] reported no statistically significant difference between Narcotrend monitoring and standard clinical monitoring, whereas the other trial[63] reported a statistically significant reduction in time to eye opening of 5.9 minutes in the Narcotrend group compared with standard care.

Time to extubation

Time to tracheal extubation was the primary outcome in one study[60] and no difference was found between monitoring of anaesthesia by Narcotrend and standard clinical monitoring (*Table 27*). In contrast,

TABLE 26 Time to eye opening

Study	Time to eye opening (minutes)		*p*-value
	Narcotrend	Standard clinical monitoring	
Kreuer *et al.*[64]			
Mean ± SD	3.7 ± 2.0	4.7 ± 2.2	NS
Kreuer *et al.*[63]			
Mean ± SD	3.4 ± 2.2	9.3 ± 5.2	$p<0.001$

NS, not statistically significant.

TABLE 27 Time to extubation

Study	Time to extubation (minutes)		*p*-value
	Narcotrend	Standard clinical monitoring	
Rundshagen *et al.*[60]			
Mean ± SD	10.6 ± 7.19	9.29 ± 6.23	NS
Kreuer *et al.*[64]			
Mean ± SD	3.6 ± 2.0	5.0 ± 2.4	$p<0.05$
Kreuer *et al.*[63]			
Mean ± SD	3.7 ± 2.2	9.7 ± 5.3	$p<0.001$

NS, not statistically significant.

TABLE 28 Time to emergence from anaesthesia

Study	Time to emergence from anaesthesia (minutes)		*p*-value
	Narcotrend	Standard clinical monitoring	
Lai *et al.*[59]			
Mean ± SD time to arousal	4.9 ± 2.2	9.5 ± 2.9	$p<0.01$
Mean ± SD orientation recovery	6.6 ± 3.2	12.2 ± 3.5	$p<0.01$

two other studies that reported time to extubation found statistically significant reductions in time to extubation of between 1.4 to 6 minutes with Narcotrend monitoring compared with standard clinical monitoring.[63,64]

Other measures of time to emergence from anaesthesia

Time to arousal (defined as the time between cessation of drugs and the patient being able to open their eyes on command) was statistically significantly shorter in the group receiving Narcotrend monitoring than the group receiving standard clinical monitoring.[59] Duration of orientation recovery was also shorter with Narcotrend monitoring (*Table 28*).[59]

Postoperative nausea and vomiting

One study found that no nausea or vomiting was reported after surgery in either group.[59] Another study[60] reported that nausea scores were statistically significantly higher in the group receiving anaesthesia monitoring by standard clinical practice than by Narcotrend at 10 minutes after extubation (mean ± SD,

24.06 ± 34.04 vs 6.88 ± 15.2, respectively, $p = 0.005$); however, there were no significant differences at other time points.

Analgesic consumption

Two studies[59,60] reported consumption of pain-relieving drugs and found no statistically significant differences between Narcotrend and standard care monitoring groups.

Summary of Narcotrend assessment

- Four trials[59,60,63,64] monitored intraoperative awareness in adults receiving different volatile and i.v. anaesthetics; no patients reported explicit memory during anaesthesia although two patients receiving Narcotrend monitoring recalled dreaming during anaesthesia.
- Three studies[59,60,63] that measured consumption of propofol reported different results; significantly lower consumption was found in the Narcotrend group in two studies, whereas no difference was reported between groups in the third study.
- Three studies[60,63,64] found no significant difference between groups in remifentanil or desflurane consumption.
- Two studies[63,64] reported time to arrival at PACU and found statistically significantly shorter times in the Narcotrend group compared with standard care.
- Time to eye opening was the primary outcome in two studies[63,64] which yielded conflicting results; one reported a significantly lower time in the Narcotrend group than with standard care and the other reported no difference between groups.
- Time to extubation was the primary outcome in one study[60] which found no difference between groups; two other studies[63,64] that reported this measure as a secondary outcome found significantly shorter time to extubation with Narcotrend monitoring than with standard care.
- Time to arousal and duration of orientation recovery were reported to be shorter with Narcotrend monitoring compared with standard care in the one study[59] reporting these outcomes.
- Results suggest that there are no differences between groups in PONV after surgery or analgesic consumption from the two studies that report these outcomes.
- In summary, Narcotrend monitoring compared with standard practice during minor orthopaedic surgery resulted in shorter recovery times (eye opening, arrival at PACU and time to extubation) and reduced propofol consumption. It was also associated with lower doses of propofol and shorter recovery during TIVA with propofol and fentanyl in liver cancer microwave coagulation. Narcotrend-assisted propofol–remifentanil anaesthesia did not reduce propofol or remifentanil consumption or time to extubation compared with standard clinical assessment in patients undergoing a range of elective surgery. The majority of the outcomes reported in the studies of Narcotrend were secondary and may not have been adequately powered statistically to detect clinically relevant differences. Also, the trial results are applicable to the specific patient groups included in the studies for the type of anaesthesia used and are not generalisable beyond this.

Results of systematic review of cost-effectiveness

The aim of this section is to assess the current state of evidence on the cost-effectiveness of depth of anaesthesia monitoring compared with standard clinical monitoring through a systematic review of the literature. The methods used for the search strategy are described in *Identification of studies*, and inclusion criteria are shown in *Inclusion/exclusion criteria*. Included studies were evaluated for their quality and for generalisability to the UK. This section concludes a statement on the current state of evidence on the cost-effectiveness of depth of anaesthesia monitoring and a discussion of key issues arising from included studies. The full data extraction forms for included studies are shown in *Appendix 6*.

Quantity and quality of research available

A total of 134 potentially relevant references were identified in the cost-effectiveness searches. Of these, the full text of 14 papers was retrieved and one study[97] met all of the a priori inclusion criteria. A summary

of the selection process and the reasons for exclusion are presented in *Figure 5* – a list of excluded studies can be found in *Appendix 7*.

The excluded studies were predominantly cost analyses, completed as part of BIS trials, which reported the difference in drug cost between the BIS and control arms. An update search, conducted in February 2012, identified six possible studies. These were all excluded on the basis of title and abstract as either not being full economic evaluations or did not include the specified interventions and comparators. The included study was simple calculation models of BIS monitoring compared with standard treatment. The completed checklist for quality assessment of the included study is shown in *Table 29*.

Characteristics and results of included studies

The included study was a simple calculation model of BIS monitoring compared with standard treatment. Characteristics of the study are shown in *Table 30* and a full data extraction form can be found in *Appendix 6*.

The included study employed a relevant comparator and similar patient group to the UK NHS. However, the study was of poor quality, with limited information reported on the methods, and sources used for the model parameters. Assumptions were not justified. The study did not include health-related quality of life (HRQoL) or investigate uncertainty through sensitivity analyses.

Abenstein[97] used a simple calculation model to compare GA with BIS monitoring to GA for high- and general-risk patients. The cost per avoided intraoperative recall is:

$$\frac{\text{Cost per patient of BIS}}{\text{Incidence BIS} - \text{Incidence GA}} \quad (1)$$

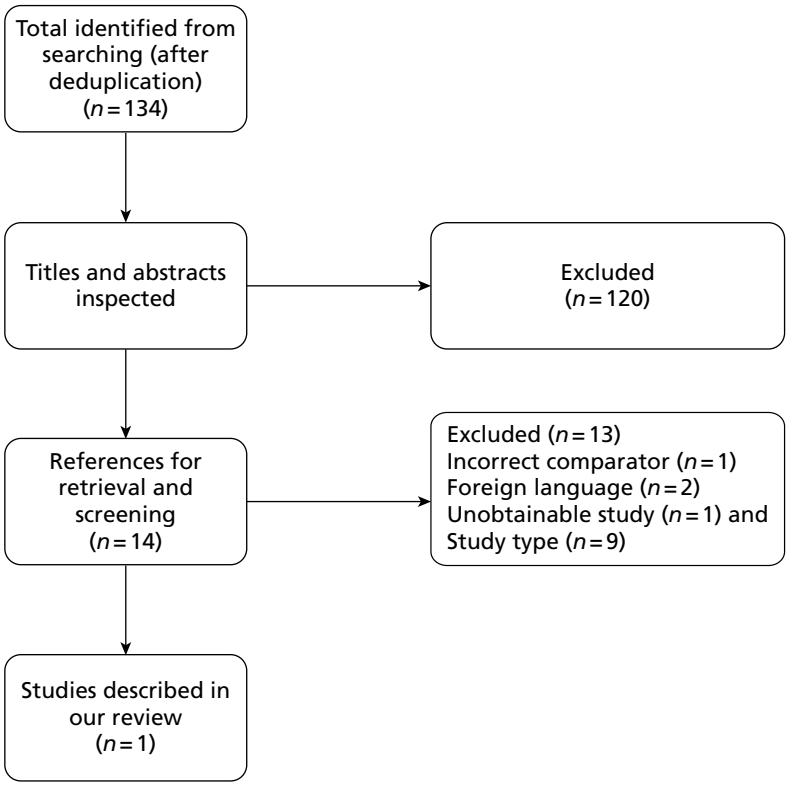


FIGURE 5 Flow chart of identification of studies for inclusion in the review of cost-effectiveness.

TABLE 29 Critical appraisal checklist of economic evaluation (questions in this checklist based on Philips and colleagues[37])

	Item	Abenstein[97]
1	Is there a clear statement of the decision problem?	Y
2	Is the comparator routinely used in the UK NHS?	Y
3	Is the patient group in the study similar to those of interest in the UK NHS?	Y
4	Is the health-care system comparable to the UK?	N
5	Is the setting comparable to the UK?	Y
6	Is the perspective of the model clearly stated?	N
7	Is the study type appropriate?	Y
8	Is the modelling methodology appropriate?	Y
9	Is the model structure described and does it reflect the disease process?	Y
10	Are assumptions about model structure listed and justified?	N
11	Are the data inputs for the model described and justified?	?
12	Is the effectiveness of the intervention established based on a systematic review?	N
13	Are health benefits measured in QALY?	N
14	Are health benefits measured using a standardised and validated generic instrument?	N
15	Are the resource costs described and justified?	?
16	Have the costs and outcomes been discounted?	N
17	Has uncertainty been assessed?	N
18	Has the model been validated?	N

N, no; Y, yes; ?, unclear.

TABLE 30 Characteristics of included economic evaluation

Author	Abenstein[97]
Publication year	2009
Country	USA
Study type	Cost-effectiveness analysis
Intervention(s)	BIS
Model type	Simple calculation
Intervention effect	Reduction in awareness for all patients from 0.18% to 0.04%
Base-case results	Cost of preventing each episode of awareness is US$11,294 for all patients

The cost per patient of BIS monitoring consisted of the cost of the sensors (US$17 each) and the cost of the monitor. The monitor was assumed to cost US$9000 and have a lifespan of 7 years, and be used by four patients per day for 300 days per year (US$1.07 per patient). The incidence of intraoperative recall for patients of general risk was taken from a prospective study by Ekman and colleagues[98] who reported a recall rate of 0.04% (GA with BIS) compared with 0.18% (GA). The cost per avoided intraoperative recall was US$11,294. Abenstein[97] estimated the cost per avoided intraoperative recall for high-risk patients to be US$4410 per avoided intraoperative recall. They used estimates of the incidence of intraoperative recall by averaging the difference between the studies by Myles and colleagues[79] and Avidan and colleagues,[27]

which gave a reduction in incidence of intraoperative recall from 0.59% to 0.18%. The authors concluded that the general use of BIS monitoring does not seem warranted and appears not to be cost-effective.

Summary

One cost-effectiveness analysis[97] was included in this systematic review, which compared BIS with standard clinical monitoring, using a simple calculation model. The study concluded that addition of BIS to GA was not cost-effective. However, the results and conclusions should be viewed with caution because of the poor methodological and reporting quality.

Model structure, model parameterisation and results of economic evaluation

Description of decision-analytic model

Overview

A decision-analytic model was developed to assess the cost-effectiveness of depth of anaesthesia monitoring, compared with standard clinical monitoring, in accordance with the scope of the appraisal issued by NICE. Separate analyses are presented for each of the three included technologies (the included technologies are not compared with each other).

The model was structured to include outcomes identified in the scope issued by NICE for this appraisal, where suitable data on the relative effectiveness of included technologies was identified in our systematic review of patient outcomes (see *Results of systematic review of patient outcomes*). The model evaluates costs (UK pounds using a 2011 price base) from the perspective of the NHS and Personal Social Services. Outcomes in the model are expressed as QALY. Both costs and outcomes are discounted using a 3.5% annual discount rate, in line with current guidance.[38,39]

Modelling approach and model structure

The model developed for this assessment was a simple decision tree, which accounted for patients' risk of experiencing short-term anaesthetic-related complications (such as PONV) and more serious complications that may be associated with risk of morbidity or mortality. These were included, in addition to a risk of experiencing intraoperative awareness, see *Figure 6*.

Each of the short-term anaesthetic-related complications could be associated with additional treatment costs (such as antiemetic medication for patients experiencing PONV, whereas for patients experiencing POCD there may be in-hospital costs of managing the condition, additional days of hospital stay and, for longer-term cases, additional costs of managing the condition following discharge). No direct cost-consequences for intraoperative awareness are included in the model. However, it is assumed that a proportion of patients who experience awareness will suffer psychological symptoms arising from the awareness episode and that a proportion of those will develop PTSD and may seek treatment.

We assumed that monitoring of basic clinical signs, including blood pressure and heart rate (mandatory worldwide), would be common components to standard clinical monitoring and to depth of anaesthesia monitoring using EEG devices (as discussed in *Description of technologies under assessment*) and, therefore, these have not been costed in the model. The key cost components identified for the standard clinical monitoring branch of the model are the costs of anaesthesia, costs of anaesthesia-related complications and costs of managing long-term sequelae of intraoperative awareness, with baseline levels (unit costs, estimated baseline consumption of anaesthetics and estimated baseline incidence of anaesthesia-related complications/intraoperative awareness) defined at the root node of the tree. The effects of EEG-based depth of anaesthesia monitoring (using the included technologies) compared with

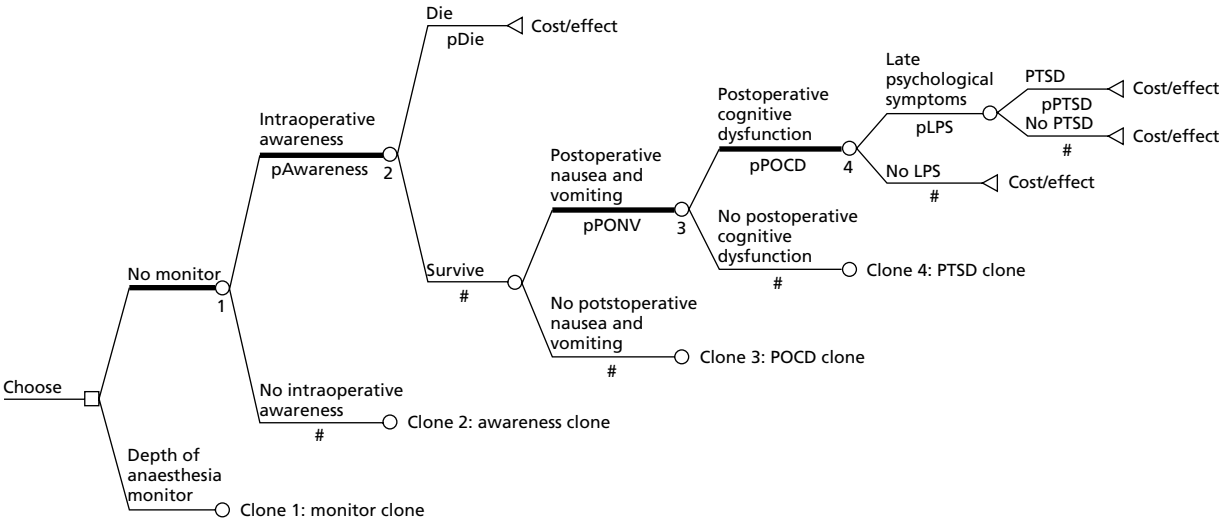


FIGURE 6 Decision tree evaluating cost-effectiveness of depth of anaesthesia monitoring compared with standard clinical monitoring.

standard clinical monitoring, which have been identified and assessed in the systematic review of patient outcomes, are applied to the baseline estimates, at the depth of anaesthesia monitoring node. These are applied as proportionate changes or OR/relative risks.

No quality-of-life (QoL) impact (utility loss) is included in the model for short-term anaesthesia-related complications (such as PONV) as these are expected to be of limited duration. Similarly, the model does not include an estimate of the QoL impact (utility loss) for an intraoperative awareness episode. The most significant quality-of-life QoL impact of any intraoperative awareness experience is assumed to be captured by estimating the incidence of psychological symptoms arising as a result of the awareness episode (including cases of PTSD).

As indicated, data population of the model required the estimation of baseline risks for a number of parameters in addition to the effectiveness estimates drawn from the systematic review of patient outcomes. The following section identifies the model parameters and the data sources used in the model.

Model parameters

Cost of depth of anaesthesia monitoring

The costs of depth of anaesthesia monitoring consist of the capital costs associated with acquisition of the module and recurring costs associated with sensors which are attached to the patient. *Table 31* summarises the costs supplied by manufacturers for each of the modules included in the assessment.

Equivalent annual costs for each module (assuming a 5-year useful life for the equipment and a discount rate of 3.5%) are presented in *Table 32*.

The annual throughput of patients for each module is assumed to be 1000 patients per year (equivalent to four patients per day for a working year of 250 days) if used for patients at average risk of intraoperative awareness, based on discussion with clinical experts. We assumed that throughput would be halved if depth of anaesthesia monitoring was limited only to patients at high risk of intraoperative awareness (equivalent to two patients per day for a working year of 250 days) – the impact of assumptions regarding patient throughput on the unit costs for DoA modules is tested in scenario analyses.

TABLE 31 Costs of depth of anaesthesia modules

Depth of anaesthesia model	Manufacturer	Cost of depth of anaesthesia monitor (£)	Sensor cost, per patient (£)
E-Entropy module	GE Healthcare	5352	8.68[a]
Vista module (BIS)	Covidien	4350[b]	14.50[c]
Compact M monitor	Narcotrend	8572–11,998[d]	0.56[e]

a Based on manufacturer's price of £217 for box of 25 sensors (one sensor per patient).
b Manufacturer's price for BIS Vista module.
c Based on manufacturer's price of £362.50 for box of 25 sensors (one sensor per patient for Vista module).
d Range of prices quoted, dependent on model. Manufacturer stated that these prices are approximate.
e Based on manufacturer's price of £0.14 per sensor (three required for one-channel recording and five required for two-channel recording. Manufacturer stated that these prices are approximate.

TABLE 32 Equivalent annual costs of depth of anaesthesia modules

Depth of anaesthesia module	Equivalent annual cost (£)
E-Entropy module	1185
BIS module	963
Narcotrend monitor[a]	2278

a Based on the mid-point of the range quoted by the manufacturer, £10,285.

Additional costs

The manufacturers' submissions to NICE indicate minimal additional power consumption associated with the modules. Therefore no additional costs were added to account for this.

The need for additional training for staff to operate the monitor appears to vary by model, according to the industry submissions. Narcotrend models require a day for the delivery of a lecture and training in the operating theatre or intensive care unit (ICU). The technical part of the training (handling the Narcotrend device, electrode placing) requires < 1 hour. The manufacturer of the E-Entropy model states that a 30-minute introductory training session is required in placement of the sensors, whereas no additional training is required for the use of a BIS monitor. This is not currently accounted for in the model.

The Narcotrend device included in this assessment is a stand-alone monitor [although the manufacturer's submission states that it can also send data in real time to other anaesthesia monitors (makes and coverage not specified)], whereas BIS and E-Entropy are modules designed to operate with other anaesthesia monitors. BIS is compatible with a range of monitoring platforms.

(Commercial-in-confidence information has been removed.) E-Entropy is compatible with GE Healthcare's most recent monitor range (CARESCAPE Monitors B850 and B650), but not older software levels (in GE Healthcare monitors) or with monitors produced by other manufacturers. The manufacturer's submission estimates that 45% of all UK operating theatres would be compatible with E-Entropy – for the remaining 55% significant investment in new monitoring equipment may be required for compatibility. Costs based on *Table 31* would not be representative for facilities requiring such investment in new monitoring equipment.

The manufacturers did not supply any information on maintenance costs or costs of maintenance contracts for any of the depth of anaesthesia modules. As a result, the base case excludes any costs for recurrent maintenance. The potential impact of maintenance costs are examined in scenario analyses using

assumptions regarding maintenance costs (annual maintenance costs estimated at 10% and 20% of the module acquisition cost).

Summary of unit costs for depth of anaesthesia modules

Unit costs for DoA modules include acquisition costs for the module (annualised, assuming a 5-year effective life, and converted to an average cost per patient using assumptions on patient throughput) and recurring costs arising from the single-use sensors attached to the patient.

Unit costs included in the base case do not include estimates of the cost of formal training or familiarisation with equipment or maintenance costs.

Anaesthetic dose

Baseline value

We undertook targeted searches for studies reporting costs of anaesthetics or estimates of anaesthetic consumption against duration of anaesthesia. Elliott and colleagues[99] reported a national survey of anaesthetic practice for paediatric and adult day surgery in UK and undertook a prospective RCT comparing the cost-effectiveness of anaesthetic regimens in adults (general, orthopaedic and gynaecology patients) and paediatric cases (general plus ear, nose and throat patients). They reported total costs (broken down by variable, semi-fixed and fixed components) for four anaesthetic regimens. The included regimens were TIVA (propofol induction, propofol maintenance), i.v./inhalational anaesthesia (propofol induction, isoflurane/N_2O maintenance or propofol induction, sevoflurane/N_2O maintenance) and total inhalational anaesthesia (sevoflurane induction, sevoflurane/N_2O maintenance).

A total of 1063 adult patients remained in the study until hospital discharge (265 propofol/propofol, 267 propofol/isoflurane, 280 propofol/sevoflurane, 251 sevoflurane/sevoflurane). The mean total and variable costs reported for the RCTs are shown in *Table 33*.

Variable costs included for each anaesthetic regimen in the trial were reported as being primarily drug costs (including anaesthetic agent use), but also included other items such as disposable equipment and therefore may not be the best basis for estimating savings that may be realised by reducing anaesthetic use associated with DoA monitoring.

Baseline consumption of inhaled anaesthetic agents in the economic model was estimated using an equation reported by Chernin,[100] based on a formula originally presented by Dion.[101]

$$\text{Cost per MAC unit time} = (\text{concentration} \times \text{FGF} \times \text{duration} \times \text{MW} \times \text{cost/ml})/(2412 \times D) \quad (2)$$

where concentration is the concentration (%) of gas delivered, FGF is the fresh gas flow rate in litres/minute, duration is duration of inhaled anaesthetic delivery in minutes, MW is molecular weight in grams, D is density in grams/ml and 2412 is a factor to account for the molar volume of a gas at 21 °C. If duration is set to 60 minutes, the above formula would estimate the cost per MAC hour for a given inhaled anaesthetic agent. *Table 34* presents the required values for calculating the cost per MAC hour of isoflurane, desflurane and sevoflurane at fresh gas flow rates of 2 l/minute for maintenance of anaesthesia.

TABLE 33 Mean total and variable costs, by anaesthetic regime, reported for CESA RCT

Cost	Propofol/ propofol	Propofol/ isoflurane	Propofol/ sevoflurane	Sevoflurane/ sevoflurane	Total
Mean total cost (£)	131.70	118.70	123.40	131.30	126.10
Mean variable cost (£)	21.10	7.10	13.80	15.30	14.40

TABLE 34 Estimated consumption of inhaled anaesthetic agents, ml per MAC hour

Input	Units	Sevoflurane	Isoflurane	Desflurane
Anaesthetic concentration	%	1.80	1.15	6.60
Fresh gas flow	l/minute	2	2	2
Duration	Minutes	60	60	60
Molecular weight of anaesthetic	g	200.00	184.50	168.00
Density	g/ml	1.52	1.50	1.45
Cost	£/ml	0.5920	0.2280	0.3040
ml per MAC hour	ml	11.78	7.04	38.04
Cost per MAC hour	£	6.98	1.60	11.57

Consumption of i.v. anaesthetic (e.g. propofol) will be based on reported total consumption in included trials. Where this is not reported consumption will be estimated based on normalised rates (mg/kg/hour or μg/kg/hour where appropriate), average patient weight and duration of anaesthesia.

Change in anaesthetic consumption associated with depth of anaesthesia monitoring

The summary values reproduced in *Table 35* below are taken from the systematic review of patient outcomes reported earlier in *Results of systematic review of patient outcomes*.

Consumption of anaesthetic drugs used in TIVA, for the comparison of E-Entropy and standard clinical monitoring, is based on data reported in two clinical trials[55,62] that were modelled separately, as we considered them unsuitable for pooling, given substantial differences in the patient populations.

Unit cost of anaesthetic agents

Unit costs for propofol are taken from the *British National Formulary* (BNF[33]). Unit costs for volatile inhaled anaesthetic gases are not available in the BNF. As a result, these costs have been provided by University Hospital Southampton NHS Foundation Trust. The unit costs reported for inhaled anaesthetic gases are based on currently quoted wholesale prices and do not reflect any discounts that may be available to NHS purchasers (*Table 36*).

Estimated baseline (standard clinical monitoring) cost of anaesthetic agents adopted in the model

Table 37 presents a summary of estimated baseline costs, change in anaesthetic consumption and cost of anaesthetic associated with use of DoA monitoring, based on assumptions presented in *Tables 34* and *Table 36*.

Postoperative nausea and vomiting

Our systematic review of patient outcomes identified limited evidence of the impact of DoA monitoring on the risk of PONV. A baseline risk of PONV (30%)[102–104] for standard clinical monitoring and DoA monitoring has been included in the model. The sensitivity of the results to the potential impact of depth of anaesthesia monitoring on the risk of PONV is explored in a scenario analysis using data from a meta-analysis on the effectiveness of BIS on a range of outcomes including PONV by Liu.[105] We assumed that all treatments (such as prophylaxis against PONV) were the same for each monitoring group, and that all patients experiencing PONV were treated using 4 mg ondansetron by intramuscular or slow i.v. injection (unit cost = £5.39; BNF[33]).

TABLE 35 Change in anaesthetic usage associated with depth of anaesthesia monitoring – mean difference and 95% CI

Technology	Anaesthetic agent	Population	Number of trials	Mean difference (95% CI)	Proportionate change (95% CI)
BIS vs standard clinical monitoring	Sevoflurane	General surgical	9[49,65,67,68,80,81,84,86,89]	–0.15 (–0.25 to –0.06)	–0.202[a] (–0.330 to –0.074)
	Propofol	General surgical	11[62,63,66,70,71,75,76,78,85,86,91]	–1.30 (–1.83 to –0.76)	–0.193[b] (–0.272 to –0.113)
E-Entropy vs standard clinical monitoring	Sevoflurane	General surgical	1[61]	–0.04[c] (–0.07 to –0.01)	–0.286 (–0.492 to –0.079)
	Propofol	Orthopaedic surgery	1[62]	5[d] (–7.54 to 17.54)	0.050 (–0.075 to 0.174)
	Remifentanil			–0.01[e] (–0.02 to 0.00)	–0.111 (–0.232 to 0.010)
	Propofol	Elective gynaecological laparoscopy	1[55]	–14[f] (–22.47 to –5.53)	–0.147 (–0.237 to –0.058)
	Remifentanil			0.07[g] (0.03 to 0.11)	0.179 (0.085 to 0.274)
Narcotrend vs standard clinical monitoring	Desflurane	Orthopaedic surgery	1[64]	–69[h] (–113.37 to –24.63)	–0.156 (–0.256 to –0.056)
	Remifentanil			–0.01[i] (–0.04 to 0.02)	–0.043 (–0.168 to 0.081)
	Propofol	Minor orthopaedic surgery	2[60,63]	–1.99[j] (–2.922 to –1.06)	–0.292 (–0.429 to –0.155)
	Remifentanil			–0.01[k] (–0.04 to 0.01)	–0.054 (–0.158 to 0.050)

a Mean difference divided by weighted mean consumption (MAC equivalents) in standard monitoring arm (meta-analysis weights).

b Mean difference divided by weighted mean normalised consumption (mg/kg/hour) in standard monitoring arm (meta-analysis weights).

c Mean difference in patient-normalised consumption (g/kg/hour). Mean normalised consumption in standard monitoring arm of trial was 0.14 g/kg/hour.

d Mean difference in patient-normalised consumption (µg/kg/minute). Mean normalised consumption in standard monitoring arm of trial was 101 µg/kg/minute.

e Mean difference in patient-normalised consumption (µg/kg/minute). Mean normalised consumption in standard monitoring arm of trial was 0.09 µg/kg/minute.

f Mean difference in patient-normalised consumption (µg/kg/hour). Mean normalised consumption in standard monitoring arm of trial was 95 µg/kg/minute.

g Mean difference in patient-normalised consumption (µg/kg/minute). Mean normalised consumption in standard monitoring arm of trial was 0.39 µg/kg/minute.

h Mean difference in patient-normalised consumption (mg/kg/minute). Mean normalised consumption in standard monitoring arm of trial was 443.60 mg/kg/minute.

i Mean difference in patient-normalised consumption (µg/kg/minute). Mean normalised consumption in standard monitoring arm of trial was 0.23 µg/kg/minute.

j Mean difference in patient-normalised consumption (mg/kg/hour), pooled across two trials (see *Appendix 8*). The mean-normalised consumption pooled across the standard monitoring arms of the trials was 6.81 mg/kg/hour.

k Mean difference in patient-normalised consumption (µg/kg/minute), pooled across two trials (see *Appendix 8*). The mean-normalised consumption pooled across the standard monitoring arms of the trial was 0.25 µg/kg/minute.

TABLE 36 Unit costs of general anaesthetics

Anaesthetic agent	Unit	Cost (£)	Cost (£)/ml
Isoflurane	250-ml bottle	57.00[a]	0.228
Desflurane	250-ml bottle	76.00[a]	0.304
Sevoflurane	250-ml bottle	148.00[a]	0.592
Propofol (1% injection, 10 mg/ml)	50-ml bottle	10.10[b]	0.202

a University Hospital Southampton NHS Foundation Trust.
b BNF.[33]

TABLE 37 Estimated baseline cost, estimated change in consumption and cost of anaesthetic associated with depth of anaesthetic monitoring

Comparison	Source	Agent	Cost (£)	Proportionate change	Estimated cost with depth monitoring (£)
BIS vs standard clinical monitoring	Meta-analysis	Sevoflurane	11.04[a]	–0.202	8.81
		Propofol	20.92	–0.193	16.90
Entropy vs standard clinical monitoring	Aime *et al.*[61]	Sevoflurane	15.93[c]	–0.286	11.38
	Ellerkmann *et al.*[62]	Propofol	18.85[d]	0.050	19.78
		Remifentanil	4.26[e]	–0.111	3.78
	Gruenewald *et al.*[55]	Propofol	14.35[f]	–0.147	12.24
		Remifentanil	14.94[g]	0.179	17.62
Narcotrend vs standard clinical monitoring	Kreuer *et al.*[64]	Desflurane	24.09[h]	–0.156	20.35
		Remifentanil	11.63[i]	–0.043	11.12
	Kreuer *et al.*[63] and Rundshagen *et al.*[60]	Propofol	19.39[j]	–0.292	13.72
		Remifentanil	10.79[k]	–0.054	10.20

a Anaesthetic duration of 1.6 hours.
b Normalised consumption of 6.73 mg/kg/hour, patient weight of 77 kg, anaesthetic duration of 2 hours.[62]
c Anaesthetic duration of 2.3 hours.[61]
d Normalised consumption of 6.06 mg/kg/hour, patient weight of 77 kg, anaesthetic duration of 2 hours.[62]
e Normalised consumption of 0.005 mg/kg/hour, patient weight of 77 kg, anaesthetic duration of 2 hours.[62]
f Normalised consumption of 5.70 mg/kg/hour, patient weight of 68 kg, anaesthetic duration of 1.8 hours.[55]
g Normalised consumption of 0.023 mg/kg/hour, patient weight of 68 kg, anaesthetic duration of 1.8 hours.[55]
h Anaesthetic duration of 2.1 hours.[64]
i Normalised consumption of 0.014 mg/kg/hour, patient weight of 79 kg, anaesthetic duration of 2.1 hours.[64]
j Normalised consumption of 6.81 mg/kg/hour, patient weight of 79 kg, anaesthetic duration of 1.8 hours.[60,63]
k Normalised consumption of 0.015 mg/kg/hour, patient weight of 79 kg, anaesthetic duration of 1.8 hours.[60,63]

Postoperative cognitive dysfunction

Baseline value

Our systematic review of patient outcomes identified limited evidence of the impact of depth of anaesthesia monitoring on the risk of POCD. One study, conducted in an elderly population (> 60 years old) available as an abstract, reported a reduction in POCD for BIS-monitored patients at 7 days and 3 months, although the difference at 7 days was reported to not be statistically significant. There is disagreement over the true incidence of POCD, with some authors arguing that this may be

underestimated because of loss to follow-up for the most severe cases,[106] whereas others argue that it may be overestimated by identifying as POCD what was a pre-existing cognitive decline. Duration of POCD was estimated using data reported for the International Study of Post-Operative Cognitive Dysfunction (ISPOCD).[107] This study recruited people over the age of 60 years who were presenting for major abdominal, non-cardiac thoracic or orthopaedic surgery under GA. Subjects with Mini Mental State Examination (MMSE) score of <23 at baseline were excluded. Incidence of POCD at 1 week after surgery was 25.8% and was present in 9.9% of subjects at 3 months. This compared with 3.4% at 1 week and 2.8% at 3 months in non-surgical controls. Longer-term follow-up of subjects in the ISPOCD study,[108] between 1 and 2 years, reported cognitive dysfunction in 10.4% of patients and 10.6% of controls, although there was considerable attrition of the cohort (336 of the original 1218 subjects followed up between 1 and 2 years). For this assessment we have assumed that the excess (22.4% at 1 week and 7.1% at 3 months) represents cognitive dysfunction attributable to undergoing GA, which will then gradually reduce to zero (at 18 months). Using these proportions (22.4% at 1 week, 7.1% at 3 months and zero at 18 months), we used the area under the curve to estimate the mean duration of POCD at 29.65 days for patients over the age of 60 years. We estimated the proportion of surgical patients experiencing POCD using data on the proportion of patients undergoing any procedure available from Health Episode Statistics (HES) online,[109] which reported that 45% of patients were age 60 years and above.

Change in postoperative cognitive dysfunction associated with depth of anaesthesia monitoring

Odds ratios for POCD at 7 days and at 3 months were estimated using data tabulated in the abstract by Chan and colleagues[47] (*Table 38*).

The ORs were applied to the baseline proportions with cognitive dysfunction at 7 days and 3 months, and mean duration of POCD associated with BIS monitoring was estimated at 21.10 days.

Quality of life impact of postoperative cognitive dysfunction associated with depth of anaesthesia monitoring

The QoL impact of POCD was based on the utility decrement reported by Jonsson and colleagues[110] for the difference between a MMSE evaluation score of > (no dysfunction), which had a utility of 0.69, and a MMSE evaluation score of between 21 and 25 (indicating mild cognitive impairment), which had a utility of 0.64.

Intraoperative awareness

Baseline value

Awareness (defined as postoperative recollection of events occurring during GA) has generally been described as a rare occurrence, with an incidence of 0.1–0.2% in the general surgical population. Although still rare, the risk of awareness has historically been greater (up to 1%) in particular types of surgery (cardiac surgery, caesarean section and trauma surgery).[79,111,112]

We conducted targeted searches for studies reporting incidence of intraoperative awareness in general surgical populations and in those populations identified as being at greater risk of awareness. *Table 39* reports the studies identified by the searches, the study populations as well as the methods used to assess and measure awareness. The majority of studies reported using the Brice interview[24] or modified

TABLE 38 Estimated OR for POCD at 7 days and 3 months estimated from Chan *et al*.[47]

Time	Routine care (*n* = 452)	BIS-guided anaesthesia (*n* = 449)	Estimated OR
1 week	39.1%	32.5%	0.750
3 months	12.0%	8.1%	0.646

TABLE 39 Studies reporting incidence of awareness in general surgical and high-risk populations – summary of characteristics, methods and results

Study	Study design (dates)	Population description	Sample size	Measure of awareness	Timing and frequency of measure	Incidence of awareness, *n* (%)
Liu *et al.*[113]	One centre, prospective (2/1990–4/1990)	Patients aged 16 years or older undergoing GA[a]	1000	Brice interview	Single interview between 20 and 36 hours after surgery	2 (0.2)
Ranta *et al.*[6]	One centre prospective (8/1994–8/1995)	Patients aged 12 years or older undergoing GA	2612[b]	Brice interview	Twice: in PACU; re-interviewed the same day or day after	10 (0.38)[c]
Sandin *et al.*[79]	Two centre prospective (1998–1998)	Patients undergoing GA	11,785	Modified Brice interview	Three times: in PACU; 1–3 days and 7–14 days later	18 (0.15)
Myles *et al.*[12]	QA programme, single centre (not stated)		10,811	Not stated	Once: 'first day after operation'	12 (0.11)
Sebel *et al.*[13]	Multicentre, cohort (4/2001 – 12/2002)	Patients undergoing GA	19,575	Modified Brice interview	Twice: in PACU; and then ≥1 week later	25 (0.13; includes 13 BIS-monitored cases)
Pollard *et al.*[14]	Quality assurance programme, eight centres (2002–4)	Patients aged 18 years or older undergoing GA	87,361[d]	Modified Brice interview	Twice: in PACU; and within 1–2 days of anaesthesia	6 (0.0068)
Errando *et al.*[18]	One centre, prospective (4/1995–11/2001)[e]	Patients undergoing GA	4001	Structured interview – does not appear to include Brice questions	Three times: in PACU; 7 days and 30 days later	39 (0.99)[f]
Lyons *et al.*[111]	(1982–9)	Patients undergoing GA for caesarean section	3000	Unclear	Unclear	8 (0.93)
Ranta *et al.*[112]		Cardiac surgery patients	204	Unclear	Unclear	3 (1.5)
[g]Puri *et al.*[82]	Multicentre RCT	High-risk patients	16[h]	NR	NR	1 (6.25)
Myles *et al.*[79]	Multicentre RCT	High-risk patients	1238[i]	Structured questionnaire, not defined	Three times: 2–6 hours, 24–36 hours and 30 days after surgery	11 (0.89)

NR, not reported.

Study	Study design (dates)	Population description	Sample size	Measure of awareness	Timing and frequency of measure	Incidence of awareness, *n* (%)
Avidan *et al.*[27]	RCT	High risk – at least one major criterion	974[j]	Brice questionnaire	Three times: within 24 hours, between 24 and 72 hours and at 30 days after extubation	2 (0.20)
[g]Muralidhar *et al.*[78]		High risk	20[k]	NR	NR	0 (0.00)
Avidan *et al.*[44]	Multicentre RCT	High-risk patients – at least one risk factor	2852[l]	Modified Brice interview. Michigan awareness classification for assigning to possible or definite awareness	Twice: 72 hours and 30 days after extubation	2 (0.07)[m]

NR, not reported.

a Excluded patients undergoing obstetric or intracranial surgery.

b Captured 54% (2612/4818 eligible cases).

c Reported 10 definite and nine possible awareness (incidence of 0.73% if possibles included).

d Follow-up in main database was 83.1% (177,468/211,842).

e Data collection was not continuous over the whole period. Actual data collection periods were April 1995 to April 1997 and from December 1998 to November 2001.

f Denominator for incidence calculation in report is 3921 (no explanation why this is lower than stated sample size of 4001). If 'high-risk' patients (emergency surgery, intraoperative hypotension-shock and caesarean section) were excluded, the incidence reduced to 0.8% (28/3477). At the 7-day interview, six patients previously classified as aware denied awareness, leading to an incidence of 0.6% (22/3477).

g From Cochrane Review systematic review of BIS.[34]

h In routine care (clinical signs) arm – overall trial population 30.

i In routine care (clinical signs) arm – overall trial population 2463.

j In routine care (end-tidal anaesthetic gas as guide) arm – overall trial population 1941.

k In routine care (end-tidal anaesthetic gas as guide) arm – overall trial population 40.

l In routine care (structured end-tidal anaesthetic-agent concentration protocol) arm – overall trial population 6041.

m In this trial the incidence of awareness in the standard care arm (0.07%, 2/2852) was lower than in the BIS arm (0.24%, 7/2861).

versions of the Brice interview administered on at least two occasions (with the first interview in the PACU). Three comparatively large studies (sample sizes between 10,000 and 20,000 patients) in general surgical populations estimated similar incidences and are commonly cited in support of the previously quoted incidence of 0.1–0.2%. However, two more recent studies have suggested wildly divergent incidence in the general surgical population (from 0.007% up to 0.99%). Although the authors of the study[18] indicating the highest incidence in a general surgical population reported lower values when excluding high-risk cases (emergency surgery, intraoperative hypotension-shock and caesarean section) and those patients who (in subsequent interviews) denied experiencing awareness, the reported incidence remained substantially in excess of the assumed risk for the general surgical population and closer to that assumed for high-risk patients.

A pooled estimate from all these studies gives a cumulative incidence of awareness of 0.21% (95% CI 0.06% to 0.45%) assuming random effects [Cochran's $Q = 212.55$ (df = 5); $p < 0.0001$; $I^2 = 97.6\%$ for fixed-effect model; see *Appendix 9* for details]. Excluding the two outlying studies (Pollard and colleagues[14] and Errando and colleagues[18]) yields a slightly lower estimate, with narrower CI (0.16%; 95% CI 0.10% to 0.23%) assuming random effects [Cochran's $Q = 7.85$ (df = 3); $p = 0.0493$; $I^2 = 61.8\%$ for fixed-effect model].

The incidence of awareness in high-risk patients has been calculated from the standard clinical monitoring arms of RCTs in this group of patients from our systematic review of patient outcomes (*Results of systematic review of patient outcomes*). Pooling these estimates gives a cumulative incidence of awareness of 0.45% (95% CI 0.06% to 1.19%) assuming random effects [Cochran's $Q = 19.97$ (df = 4); $p = 0.0005$; $I^2 = 80.0\%$ for fixed-effect model; see *Appendix 9* for details].

In the model we apply the pooled estimates of 0.16% (95% CI 0.10% to 0.23%) for risk of awareness in the base case for general surgical patients and 0.45% (95% CI 0.06% to 1.19%) for high-risk patients. The lowest incidence (0.007%), reported by Pollard and colleagues,[14] and highest incidence (0.99%), reported by Errando and colleagues,[18] are used in scenario analyses for general surgical patients. A high value of 1% is used in scenario analyses for high-risk patients.

Change in incidence of intraoperative awareness associated with depth of anaesthesia monitoring

The summary values reproduced in *Table 40* are taken from the systematic review of patient outcomes reported earlier in *Results of systematic review of patient outcomes*. There are no entries for E-Entropy and Narcotrend in this table as insufficient data were identified in the systematic review of patient outcomes to derive robust results. As a result the relevant OR derived for BIS were used in the model to estimate the impact on intraoperative awareness of depth of anaesthesia monitoring with E-Entropy and Narcotrend.

TABLE 40 Effectiveness of depth of anaesthesia monitoring on risk of awareness – Peto's OR and 95% CI from systematic review of patient outcomes

Model of general anaesthetic	Population	Number of trials	Peto's OR	95% CI
Mixed anaesthesia (includes both patients undergoing TIVA and patients undergoing inhaled GA)[a]	High risk	1	0.25	0.08 to 0.75
Inhaled GA only	High risk	4	1.79	0.63 to 5.11
TIVA	High risk	2	0.24	0.10 to 0.60
Pooled effect	High risk	7	0.45	0.25 to 0.81

a In this trial the choice of anaesthesia was left to the discretion of the anaesthetist – some had TIVA (approximately 42%) and others had inhaled anaesthetics (with or without i.v. anaesthetic).

In addition, the systematic review did not identify any robust data on the effect of depth of anaesthesia monitoring on the incidence of intraoperative awareness in patients considered at average risk of awareness. Consequently, the relevant OR derived for high-risk patients were used in the model to estimate the impact on intraoperative awareness of depth of anaesthesia monitoring for general surgical patients considered at average risk of awareness.

Sequelae of intraoperative awareness

Incidence of psychological sequelae A targeted search for studies reporting symptoms of patients who had reported awareness during surgery was undertaken in order to understand the health-related consequences of intraoperative awareness.

Eight studies were identified[5,7,8,10,11,19,114,115] (*Table 41*). These suggested that the patients who had experienced intraoperative awareness fall into three groups: those who do not experience any sequelae, those who experience 'late psychological symptoms' and those who go on to suffer from PTSD. Late psychological symptoms (LPS) comprise anxiety, chronic fear, nightmares, flashbacks, indifference, loneliness and a lack of confidence in future life. Anxiety, nightmares and flashbacks appeared to be the predominant symptoms in the study by Samuelsson and colleagues[115] in patients with an LPS duration of < 2 months; those experiencing symptoms for a longer duration reported nightmares and flashbacks alone. A diagnosis of PTSD is made if all six criteria of the clinician-administered PTSD scale (CAPS) are positive. These include symptoms of re-experiencing trauma, avoidance, hyper-arousal, significant distress and the duration of symptoms lasting longer than 1 month.[8]

Just two of the studies had a prospective design.[8,115] The study by Samuelsson and colleagues[115] reported 46 awareness cases in a cohort of 2681 interviewed after surgery. This is therefore the strongest evidence for development of PTSD and LPS that was identified in the targeted search. Leslie and colleagues,[8] although reporting a small cohort, were the only authors among those identified to report time to onset and duration of symptoms. However, some cases of PTSD reported were ongoing, and it is unclear how this may impact on the duration results. The two prospective studies were used to inform the baseline data inputs, for the states of LPS and PTSD, into the model, as presented in *Table 42* below. The six remaining studies were small, with limited usefulness for understanding the prevalence of psychological symptoms associated with awareness, because of retrospective design, participant recruitment methods or low recruitment levels.[5,7,10,11,19,114]

Duration of post-traumatic stress disorder Leslie and colleagues[8] reported a median duration of 4.7 years (range 4.4 to 5.6 years) for patients experiencing symptoms of PTSD. No further information on the distribution is provided so it unclear how well the median approximates to the mean duration of symptoms, as these cases of PTSD reported were ongoing, and it is unclear how this may impact on the duration results. Targeted searches did not identify any other studies reporting duration of PTSD symptoms associated with intraoperative awareness. One study was identified which reported duration of PTSD (median duration and survival curves) in a non-institutionalised, civilian population aged 15–54 years, conducted in the USA.[118] These data were from the National Comorbidity Survey (a survey designed to study the distribution, correlates and consequences of psychiatric disorder in the USA) and included 5877 respondents from 48 states. Response rates to part 2 of the survey, which included components related to PTSD, were between 98.1% (for those screening positive for any lifetime diagnosis in part 1 of the survey) and 99% (for a random subsample of those not screening positive in part 1 of the survey). The median duration of symptoms for respondents who had ever sought professional treatment ($n = 266$) was 36 months and for those who had not sought professional treatment ($n = 193$) was 64 months. We estimated the mean duration of PTSD symptoms for the population who had not sought professional treatment, by fitting a regression (assuming a Weibull distribution for the survival function; see *Appendix 10* for details) to the reported survival curves. The mean duration of PTSD symptoms derived in this analysis was 152 months (12.7 years) (*Figure 7*).

TABLE 41 Studies reporting incidence of LPS and PTSD in patients who experienced awareness – summary of characteristics, methods and results

Study	Date	Method of recruitment	Identification/classification of LPS and PTSD	Aware	LPS		PTSD	
				n	*n*	%	*n*	%
Evans[116]	1987	Advertisement in four British newspapers		27				
Moerman *et al.*[5]	1993	Referral from university hospital anaesthesiology department	Response to (open-ended) interview question – 'have there been any consequences?' (of the identified awareness episode). Patients reported sleep disturbance, dreams and nightmares, flashbacks and anxiety during the day	26	18	69	NR	
Schwender *et al.*[114]	1998	Advertisements in four German papers and on internet ($n = 21$) or referral from three hospital anaesthesia departments ($n = 24$)	Response to questionnaire items on after effects (including anxiety and nightmares). No definition for PTSD reported (simply states 'whether … PTSD syndrome developed')	45	22	49	3	7
Domino *et al.*[19]	1999	Retrospective analysis of American Society of Anaesthetists Closed Claims Project (malpractice claims) – data from 1961 to 1995[a]	No definitions – reports states '% (*n*) sustained temporary emotional distress, whereas in % (*n*) post-traumatic stress disorder developed'	61	51	84	6	10
Osterman *et al.*[10]	2001	Advertisement in newspapers, fliers in hospitals, self-referral following print and TV news stories or referral by anaesthetist	PTSD defined using CAPS	16	NR		9	56
Lennmarken *et al.*[11]	2002	18 patients identified as experiencing awareness during GA in two hospitals (reported by Sandin and colleagues[9]) were followed up for interview regarding psychological symptoms[b]	PTSD defined using diagnostic criteria A1-F in the DSM-IV, American Psychiatry Association[117]	9	7	78	4	44
Samuelsson *et al.*[115]	2007	Consecutive patients who had undergone GA were interviewed regarding awareness during previous GA	LPS were any one of: anxiety, chronic fear, nightmares, flashbacks, indifference, loneliness and lack of confidence in future life (each rated on a scale from zero to two). PTSD appears to be defined on basis of existing clinical diagnosis (not specifically identified or classified in study)	46	15[c]	33	1	2
Ghoneim *et al.*[7]	2009	Data extracted from published case reports on 'awareness' and 'anaesthesia' – from PubMed between 1950 and August 2005	No definition of LPS	271	NR	22	NR	

TABLE 41 Studies reporting incidence of LPS and PTSD in patients who experienced awareness – summary of characteristics, methods and results (*continued*)

Study	Date	Method of recruitment	Identification/classification of LPS and PTSD	Aware	LPS		PTSD	
				n	*n*	%	*n*	%
Leslie *et al.*[8]	2010	13 patients identified as experiencing awareness in the B-Aware trial (reported by Myles and colleagues[79]) were followed up for interview regarding psychological symptoms[d]	PTSD defined using CAPS	7	NR		5	71

DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders*-Fourth Edition; NR, not reported.

a Total claims for adverse outcomes between 1961 and 1995 in closed claims project was 4183.

b Of the 18 patients experiencing awareness identified by Sandin and colleagues,[9] two could not be contacted, six declined to participate and one had died.

c Samuelsson identified eight (17%) patients as having a total symptom score (summed across seven symptoms) of > 2 (no rationale for this threshold).

d Of the 13 patients experiencing awareness in the B-Aware trial,[79] six had died.

Leslie and colleagues[8] pooled their estimate of PTSD with the estimates of Lennmarken and colleagues[11] and Samuelsson and colleagues[115] for severe psychological sequelae ($n = 8$, 17%) to derive an incidence of 26% (95% CI 15% to 37%).

TABLE 42 Baseline values for probability of LPS and PTSD in patients experiencing awareness

Sequelae	Value (95% CI)	Method	Source
LPS			
Probability, given awareness	0.326 (0.195 to 0.480)	15/46 patients with awareness	Samuelsson *et al.*[115]
PTSD			
Probability, given awareness	0.177 (0.113 to 0.230)	Pooled proportion of subjects with LPS having PTSD or severe symptoms, from (two) studies reporting this proportion, applied to probability of LPS Pooled estimate based on 0.571[11] (4/7) and 0.0533[115] (8/15) = 0.542 (95% CI 0.345 to 0.733) Probability PTSD = (15/46)×0.542	Samuelsson *et al.*[115] and Lennmarken *et al.*[11]

Quality of life impact of psychological sequelae A review of the HRQoL of patients with PTSD was undertaken in order to explore the differences in scores between PTSD patients and those who had also experienced trauma, but had not gone on to develop PTSD. These scores were used to inform those in the model for patients experiencing awareness and developing psychological symptoms.

Methods

A systematic search was undertaken in order to identify studies reporting utility values associated with PTSD. The details of the search strategy are documented in *Appendix 11*. A total of 334 studies were initially identified by the search. The abstracts were screened by two independent reviewers and 21 full papers were retrieved (*Figure 8*). These were assessed against the inclusion criteria detailed in *Table 43*.

Characteristics of the included studies

Two papers[119,120] met the inclusion criteria for the review. The study design and population baseline characteristics are shown in *Table 44*.

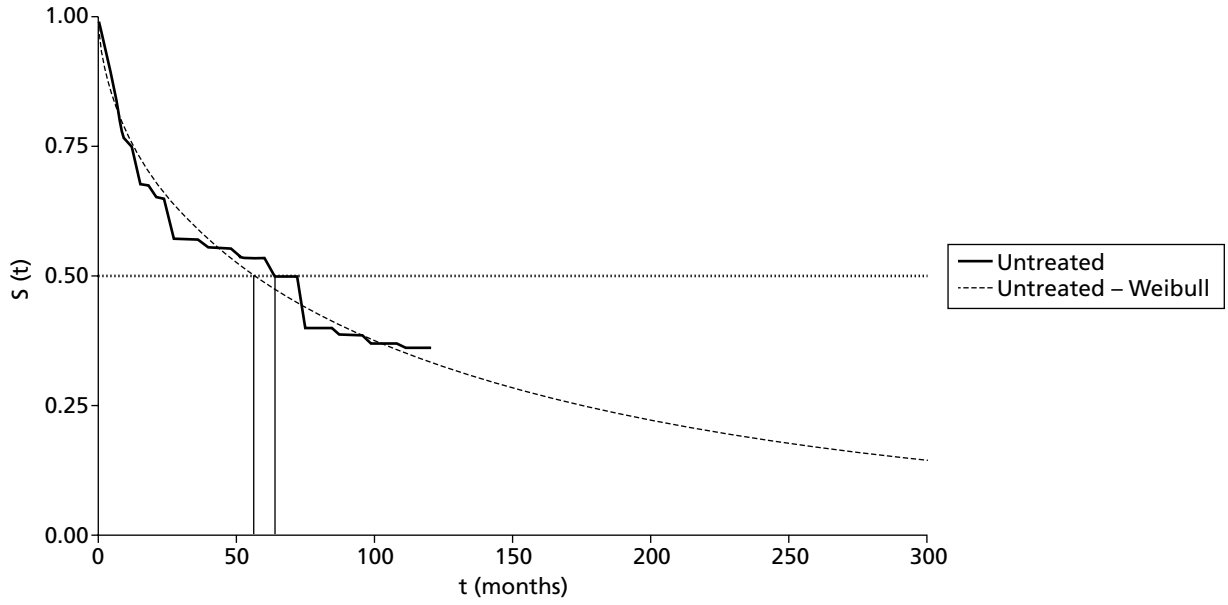


FIGURE 7 Survival curve based on duration of symptoms for respondents who did not seek treatment for PTSD, reported by Kessler and colleagues,[118] and fitted Weibull model.

TABLE 43 Inclusion criteria for QoL review

Criteria	Include	Exclude
Participants	Adults with PTSD	Studies related to or concerning specific morbidities, with the exception of psychiatric (or related) illness
Design	Studies that report a utility value, based on generic preference based measures for QoL, such as EQ-5D, SF-6D, or other standard valuation technique such as standard gamble or TTO	
Interventions	Any	
Other	Articles published in English	Articles in languages other than English Conference abstracts

EQ-5D, European Quality of Life-5 Dimensions; SF-6D, Short Form questionnaire-6 Dimensions; TTO, time trade-off.

The two included studies[119,120] were both undertaken in patients with PTSD. This population is diverse and there are a range of types of trauma that can trigger the disorder, such as domestic abuse, natural disaster or serious illness.[122–124] Freed and colleagues interviewed veterans with PTSD,[120] whereas Doctor and colleagues[125] interviewed a sample of patients taking part in a trial of treatments for chronic PTSD at baseline.

The two studies[119,120] were considerably different in size, with the Freed study[120] having approximately four times as many respondents. The two studies reported differing populations in respect to both age and sex, which may have contributed to the differing results. In the study by Freed and colleagues,[120] female patients constituted 21% of the sample, and the average age was 60 years. In the Doctor and colleagues study[125] the sample was on average younger, with a mean age of 37, and 76% of the respondents were female.

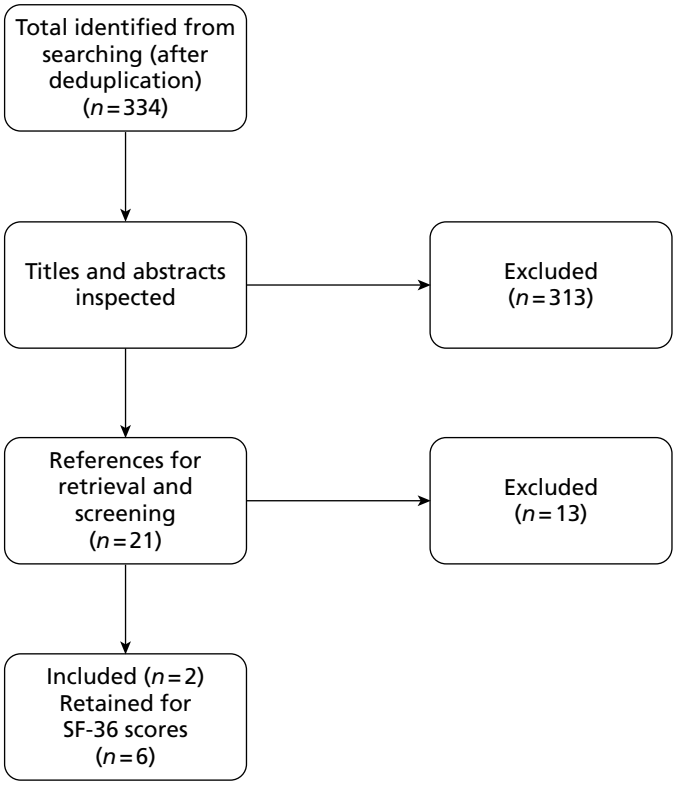


FIGURE 8 Flow chart of identification of QoL studies for inclusion in the review.

TABLE 44 Characteristics of included QoL studies

Criteria	Doctor *et al.*[119]	Freed *et al.*[120]
Patient group	Patients with PTSD	Veterans with PTSD
Country and setting	US, multicentre trial, setting not reported	US study, British sample, primary care clinics
Sample size	184	840
Duration of symptoms	Patients were a minimum of 12 weeks from the traumatic event	NR
Age, mean ± SD (years)	37.31 ± 11.33	60 ± 12
Sex (F)	141 (76%)	176 (21%)
QoL instrument	SG/TTO/VAS	SF-36[a]

NR, not reported; SF-36, Short Form questionnaire-36 items; SG, standard gamble; TTO, time trade-off; VAS, visual analogue scale.

a SF-36 scores transformed to utility scores using Brazier *et al.*[121]

In addition, the two studies generated the results using different valuation tools and methods. Neither of the included studies was based on the European Quality of Life-5 Dimensions (EQ-5D) questionnaire, as prescribed by the NICE reference case.[38] Doctor and colleagues[119] asked respondents to respond using standard gamble (SG), visual analogue scale (VAS), and time trade-off (TTO) techniques, the last of which is recommended as an alternative.[38] Freed and colleagues[126] used the Short Form questionnaire-36 items (SF-36) responses from a previous study[127] and converted these to preference weighted health scores (PWHSs) using the formula developed by Brazier and colleagues.[121]

Both studies included the results of statistical models generated in order to identify predictors for worsening or improvement of utility scores.[119,120]

Quality of the included studies Doctor and colleagues[119] clearly reported inclusion and exclusion criteria for patients entering the trial, which appeared appropriate. The methods employed to elicit utility scores were clearly described, although the description of TTO does not appear to be correct, which could undermine the results. Freed and colleagues[120] have based their analysis on the results of a previous study, the sources for the analysis are clearly stated, and the interview methods and scales employed are adequately described. The sample is of British veterans, which is relevant to the UK, but the generalisability of the HRQoL of veterans to different patient populations is unclear. Freed and colleagues[120] have also carried out ordinary least squares regressions (OLS) in order to allow researchers to adjust the estimates of patients' PWHS. The methods for these were adequately described, but contradictory results are reported: the PWHS is reported to increase if a patient has both a PTSD diagnosis and increasing severity of symptoms on the PTSD checklist (PCL). These contradictions are not fully considered or explained, and therefore limit the usefulness of the regression results in estimating HRQoL in patients with PTSD.

Results

The mean utility scores reported in each of the included studies are presented in *Table 45*.

The scores for veterans in the Freed study[120] with PTSD is lower than that of veterans without PTSD, with a difference of 0.11, suggesting that PTSD does negatively impact on HRQoL.

Doctor and colleagues[119] report three separate scores according to the valuation method. The scores for TTO and VAS are similar (0.66 and 0.64 respectively), whereas the score for SG appears high at 0.87. The authors argue that the mixed-effect model employed has accounted for possible bias in SG methods (SG requires the participants to state the probability that they would accept a treatment that has a certain probability of conferring full health, with the concomitant probability of immediate death). However, they also state that the TTO method has a lower risk of bias, although justification for this is not reported.[119] TTO is also recommended by NICE where EQ-5D scores are unavailable.[38] The study does not provide a raw comparable score for a group without PTSD using these methods, and therefore it is not possible to draw conclusions as to the decrement in utility resulting from developing PTSD from this paper.

Studies reporting Short Form questionnaire-36 items scores

A further six studies[122–124,129–131] that did not meet our inclusion criteria, but which reported the eight subscales of the SF-36, were identified. A preference-based utility score can be estimated from studies that report scores for the eight subscales of the SF-36.[128] Preference-based health-related utilities from these results have been estimated by SHTAC in order to assess the robustness of the estimates in the study by Freed and colleagues.[120] These were converted using the algorithm published by Ara and Brazier,[128] and are reported in *Table 46*.

Scores derived using the SF-36 do not meet the NICE reference case,[38] which recommends the EQ-5D, and that values generated from the Short Form questionnaire-6 Dimensions (SF-6D) [132] be employed in the

TABLE 45 Utility scores reported in the included QoL studies

	Doctor *et al.*[119]			
Patient group	SG	TTO	VAS	Freed *et al.*[120]
HRQoL score in patients with PTSD (mean ± SD)	0.87 ± 0.25	0.66 ± 0.28	0.64 ± 0.2	0.535[a]
HRQoL score in patients without PTSD (mean ± SD)	NR	NR	NR	0.652[a]

NR, not reported.

a Transformed from SF-36 using Brazier *et al.*[121]

TABLE 46 Health-related utilities estimated from SF-36 scores

Study	Patient group	Utility		
		PTSD	No PTSD	Difference
Laffaye *et al.*[122]	Women experiencing domestic abuse	0.634	0.748	0.114
Meeske *et al.*[124]	Young adult survivors of childhood cancer	0.666	0.799	0.132
Berger *et al.*[129]	Male ambulance workers	0.705	0.790	0.085
Shiner *et al.*[130]	Veterans	0.508	–	–
Tsai *et al.*[123]	Earthquake survivors (0.5 years post)[a]	0.649	0.783	0.134
Evren *et al.*[131]	Alcohol-dependent men with history of emotional abuse	0.592	0.659	0.068

a Three years post earthquake and delayed PTSD and recovery scores also reported in Tsai *et al*.[123]

sensitivity analysis. The studies reporting the SF-36 scores were carried out in diverse groups, with differing traumatic triggers for PTSD. Furthermore, caution should be exercised in the interpretation of this table as these studies have not been fully data extracted or quality assessed. However, the scores consistently indicate similar differences in HRQoL between groups of patients who have similar experiences who go on to develop PTSD, and those who do not, and the differences are consistent with those reported by Freed and colleagues.[120] On average across these papers the difference is 0.10. These results lend weight to the estimates of decrement in utility as a result of PTSD, and may also be useful for sensitivity analysis. However, the results for the utility scores in patients with and without PTSD are generally higher than those reported by Freed and colleagues,[120] with the exception of those reported by Shiner and colleagues[130] also elicited from veterans.

Summary

- Two papers met the inclusion criteria for this review of utility scores in PTSD. Six other papers reporting SF-36 scores for people with PTSD were also retained.
- Neither of the studies meeting the inclusion criteria (reporting a utility value based on a generic, preference-based measure) was based on the EQ-5D.
- One study reported a utility score for patients with PTSD based on TTO.[119] However, no score for patients without PTSD was reported, and therefore no difference in these can be derived.
- The second study reported scores for patients both with and without PTSD, but these were based on the SF-36 and converted to a utility score.[120]
- Therefore the evidence base for HRQoL in patients with PTSD is limited.
- Six further studies provide SF-36 scores, which have been transformed into utility values. These can provide context and values for sensitivity analysis.

Post-traumatic stress disorder costs The costs of treating PTSD have been estimated based on assumptions contained in the national cost impact report[133] associated with NICE Clinical Guideline no. 26[133] on the management of PTSD in adults and children in primary and secondary care.[134] The costing report acknowledged that there has been little systematic collection of information about PTSD, on services provided to people with PTSD or on uptake of these services. This limited the feasibility of developing a comprehensive bottom-up costing model and resulted in the costing being based on a series of assumptions – developed and validated through discussion with members of the Guideline Development Group (GDG) and key clinical practitioners in the NHS. These assumptions, in terms of uptake and services available, are summarised in *Figure 9* and are discussed below.

Data from the adult psychiatric morbidity survey,[135] which reported that 24% respondents assessed as having a neurotic disorder were receiving treatment of some kind at the time of interview, were used as the basis for estimating the current proportion of people with PTSD who seek treatment. On the basis

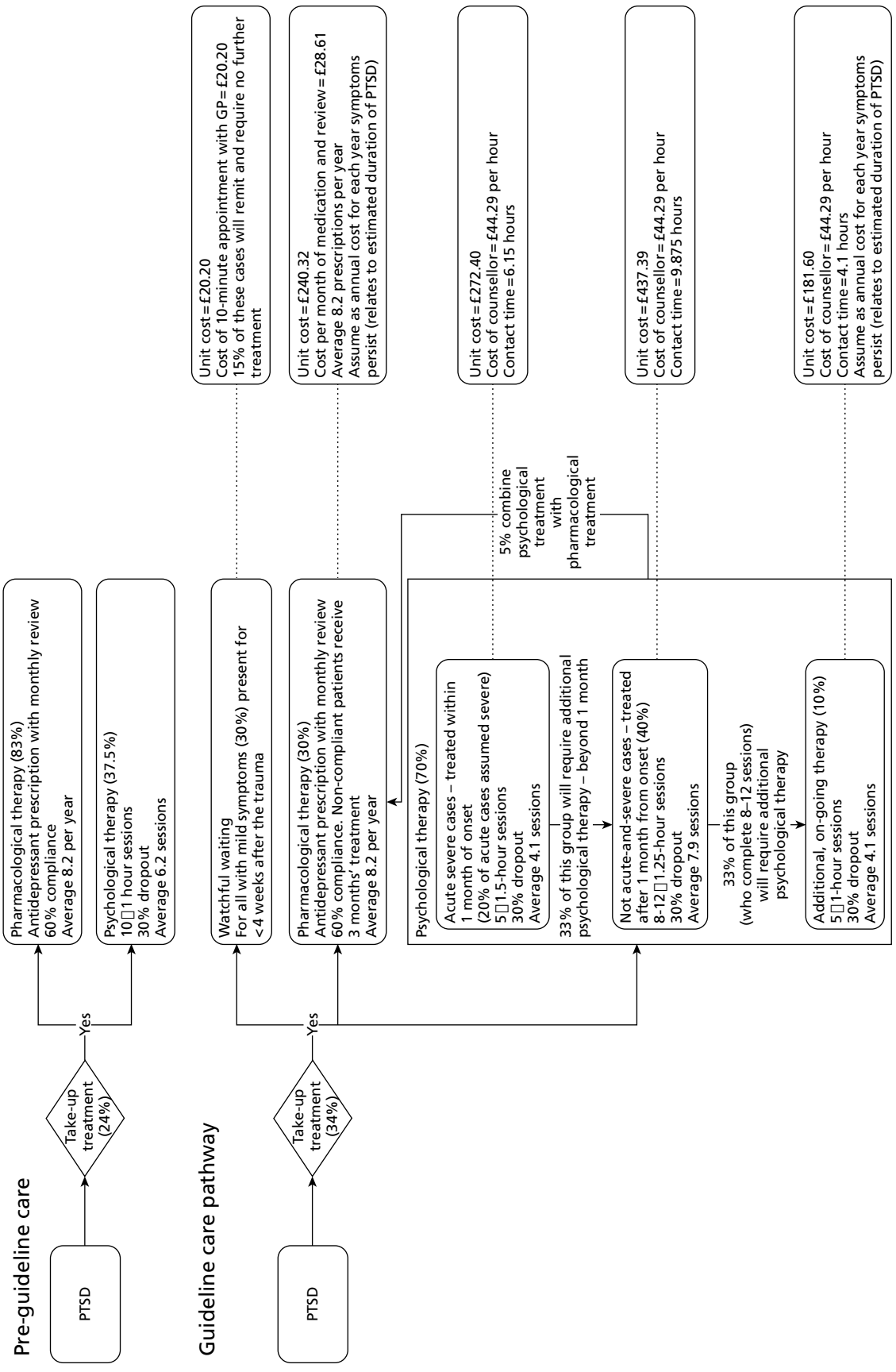


FIGURE 9 Care pathways and costing assumptions developed for NICE.

of additional data from the same survey, 62.5% of these were assumed to be receiving pharmacological therapy alone, 16.7% were receiving counselling or therapy alone and 20.8% were receiving both. It was assumed that, following implementation of the guideline, the proportion receiving treatment would increase by 10%, to 34%. Moreover, the guideline proposed a substantially different care pathway with significantly fewer patients expected to receive medication (with a recommendation that drug treatments not be offered routinely as first line, but to provide trauma-focused psychological treatment to more patients with PTSD symptoms). We estimated an average cost for management of PTSD using the assumptions regarding take-up of treatment options (70% of patients accept psychological treatment, and 30% initially accept pharmacological treatment) and severity (30% patients have mild symptoms and are initially managed through watchful waiting, and 20% have severe symptoms and are offered trauma-focused psychological treatment within the first month after the traumatic event). *Table 47* summarises the unit costs, assumptions regarding the proportion of patients receiving each treatment and the overall cost estimated for PTSD.

The NICE guideline[133] does not include any estimates for inpatient care for people with PTSD. Targeted searches did not identify any UK studies of health service use, in particular use of secondary services and inpatient care for people with PTSD. One US study identified by the searches reported health-care utilisation, derived from electronic medical records, for civilian primary care patients, including a proportion who had current PTSD.[136] This study reported an incidence rate ratio of 2.22 (adjusted for age, sex, income, substance dependence, depression and comorbidity) for hospitalisation in subjects with PTSD compared with those without PTSD. Unadjusted mean number of hospitalisations among the PTSD group was 0.43 compared with 0.18 in those without PTSD. No further details are reported on the reason for hospitalisation or length of stay. In the absence of data specific to the UK for people with PTSD, we have assumed, based on the mean values reported in this study, an excess hospitalisation probability of 0.25 per year among people with PTSD. We derived a crude estimate of the average cost of hospitalisation (£2590), based on 2010–11 NHS Reference Costs,[137] by summing the total costs reported for elective and non-elective inpatient HRG data and dividing by the total activity under these headings. On this basis we estimated an additional £7576 for hospitalisations among people with PTSD over the average duration of symptoms of 12.7 years.

The total cost associated with PTSD was £9104 (undiscounted) or £6128 (discounted at 3.5%).

TABLE 47 Unit cost and treatment uptake assumptions used to calculate costs of managing PTSD

Treatment	Unit cost (£)	Proportion (%)	Cost (£)
Watchful waiting	20.20	16.50	3.33
Pharmacological therapy	240.32	30.00	915.62
Combined pharmacological and psychological therapy	240.32	5.00	152.60
Psychological therapy (severe acute cases <1 month)	272.40	14.00	38.14
Psychological therapy (>1 month after traumatic event)	437.39	38.53	168.50
Psychological therapy (severe acute cases >1 month)	437.39	4.62	20.21
Additional, ongoing psychological therapy	181.60	9.97	229.86
Total			1528.26

Patients are assumed to remain on pharmacological therapy for the duration of their PTSD symptoms (12.7 years in the base case). Costs are discounted at 3.5%.

Patients requiring ongoing psychological therapy are assumed to continue treatment for the duration of their PTSD symptoms (12.7 years in the base case). Costs are discounted at 3.5%.

Summary of model inputs The following tables contain a summary of the input parameters in the model, the base-case value and a brief overview of how the data were derived including a source, where relevant. *Table 48* provides a summary of the cost per patient of each depth of anaesthesia technology, including an estimated cost per patient of the depth-monitoring device as well as the cost of consumables (single-use sensors attached to the patient). *Table 49* provides a summary of the baseline cost of anaesthetic drug calculated for standard clinical monitoring in each comparison and the proportionate reduction in consumption associated with depth of anaesthesia monitoring. We have assumed that the reduction in consumption of anaesthetic will be realised only for the general surgical population and not in the population at high risk of awareness, as the raised risk of awareness may be an indication that this group of patients are already at a risk of being underdosed.

Table 50 provides a summary of model inputs related to awareness including the baseline risks for patients considered at high risk of awareness and a general surgical population, the risk reduction associated with depth of anaesthesia monitoring and a list of assumptions underlying the estimation of the cost and outcomes associated with the psychological sequelae of intraoperative awareness.

Table 51 provides a summary of model inputs relating to anaesthetic complications (PONV and POCD in the model), including the baseline risks and the risk reduction associated with depth of anaesthesia monitoring for POCD.

Model results

The model results are presented in separate subsections for BIS, E-Entropy and Narcotrend respectively. Analyses are presented by mode of administration [TIVA and mixed anaesthesia (induction with i.v. anaesthetic and maintenance with inhaled anaesthetic)], with separate analyses reported for patients considered at high risk of awareness and for a general surgical population. No analysis is presented for inhaled GA only. Although trials using this mode of anaesthesia delivery were included in the systematic review of patient outcomes, these did not report any information on anaesthetic drug consumption on which to base a reliable costing.

TABLE 48 Model input parameters–cost per patient of DoA modules

Parameter	Value (£)	Source
BIS		
Cost per patient of depth monitoring device	0.96	Equivalent annual cost for depth monitor (acquisition cost £4350) assuming an effective life of 5 years and using a discount rate of 3.5%. Patient throughput assumed at 1000 per year
Cost per patient of depth monitor sensors	14.50	Manufacturer's price of £362.50 for a box of 25 sensors (for Vista monitor)
E-Entropy		
Cost per patient of depth monitoring device	1.19	Equivalent annual cost for depth monitor (acquisition cost £5352) assuming an effective life of 5 years and using a discount rate of 3.5%. Patient throughput assumed at 1000 per year
Cost per patient of depth monitor sensors	8.68	Manufacturer's price of £217 for a box of 25 sensors
Narcotrend		
Cost per patient of depth monitoring device	2.28	Equivalent annual cost for depth monitor (acquisition cost £10,285, mid-point of range quoted by manufacturer) assuming an effective life of 5 years and using a discount rate of 3.5%. Patient throughput assumed at 1000 per year
Cost per patient of depth monitor sensors	0.56	Average across manufacturer's price of £0.14 per sensor, using three for one-channel recording and five for two-channel recording

TABLE 49 Model input parameters – anaesthetic drug consumption

Parameter	Value (95% CI)	Source
BIS		
Baseline inhaled anaesthetic cost	£11.04	Cost for 1.6 MAC hours (95 minutes) of sevoflurane (concentration of 1.8% and fresh gas flow rate of 4 l/minute). Unit cost of £0.59 per ml, based on price of £148 per 250 ml
Reduction in consumption of inhaled anaesthetic using depth monitor (proportionate reduction compared with standard clinical care)	–0.202 (–0.330 to –0.074)	Mean difference of –0.15 from a (weighted) mean consumption of 0.765 MAC equivalents
Baseline i.v. anaesthetic cost	£20.92	Cost for 2 hours of propofol [at 6.77 mg/kg/hour (from control arms of RCT in meta-analysis) and patient average weight of 77 kg]. Unit cost of £0.0202 per mg
Reduction in consumption of i.v. anaesthetic using depth monitor	–0.193 (–0.272 to –0.113)	Mean difference of –0.130 from a (weighted) mean consumption of 6.73 mg/kg/hour
E-Entropy		
Baseline inhaled anaesthetic cost	£15.93	Cost for 2.3 MAC hours (137 minutes) of sevoflurane (concentration of 1.8% and fresh gas flow rate of 4 l/minute). Unit cost of £0.59 per ml, based on price of £148 per 250 ml
Reduction in consumption of inhaled anaesthetic using depth monitor	–0.286 (–0.492 to 0.079)	Mean difference of –0.04 from patient normalised consumption of 0.14 g/kg/hour (in standard care arm, Aime *et al.*[61])
Baseline i.v. anaesthetic cost	Propofol = £18.85	Ellerkmann *et al.*[62]
	Remifentanil = £4.26	
	Propofol = £14.35	Gruenewald *et al.*[55]
	Remifentanil = £14.94	
Reduction in consumption of i.v. anaesthetic using depth monitor	0.050 (–0.075 to 0.174)	Propofol mean difference of 5 from baseline rate of 101 mg/kg/hour (Ellerkmann *et al.*[62])
	–0.111 (–0.232 to 0.010)	Remifentanil mean difference of –0.01 from baseline rate of 0.09 mg/kg/hour (Ellerkmann *et al.*[62])
	–0.147 (–0.237 to –0.058)	Propofol mean difference of –14 from baseline rate of 95 mg/kg/hour (Gruenewald *et al.*[55])
	0.179 (0.085 to 0.274)	Remifentanil mean difference of 0.07 from baseline rate of 0.39 mg/kg/hour (Gruenewald *et al.*[55])
Narcotrend		
Baseline inhaled anaesthetic cost	£24.09	Cost for 2.1 MAC hours (125 minutes) of desflurane (concentration of 6.6% and fresh gas flow rate of 4 l/minute). Unit cost of £0.30 per ml, based on price of £76 per 250 ml
Reduction in consumption of inhaled anaesthetic using depth monitor	–0.156	Mean difference of –69 mg/minute from 443.6 mg/minute (in standard care arm, Kreuer *et al.*[64])
Baseline i.v. anaesthetic cost	Propofol = £19.39	Cost for 108 minutes of propofol [at 6.81 mg/kg/hour (from control arms of RCT) and patient average weight of 80 kg]. Unit cost of £0.0202 per mg
	Remifentanil = £10.79	Cost for 108 minutes of remifentanil [at 0.120 mg/kg/hour (from control arms of RCT) and patient average weight of 80 kg]. Unit cost of £5.12 per mg

continued

TABLE 49 Model input parameters – anaesthetic drug consumption (*continued*)

Parameter	Value (95% CI)	Source
Reduction in consumption of i.v. anaesthetic using depth monitor	–0.292 (–0.429 to –0.155)	Propofol mean difference of –1.99 from baseline rate of 6.8 mg/kg/hour[60,63]
	–0.054 (–0.158 to 0.050)	Remifentanil mean difference of –0.01 from baseline rate of 0.25 mg/kg/hour[60,63]

TABLE 50 Model input parameters – intraoperative awareness

Parameter	Value (95% CI)	Source
Intraoperative awareness		
Baseline awareness in surgical population at high risk of awareness	0.45% (0.06% to 1.19%)	Pooled estimate from control arms of RCT in high-risk patients
Reduction in awareness using depth monitor		Meta-analysis of RCT in high-risk patients, undertaken as part of this review (see *Assessment of outcomes: Bispectral Index*)
High-risk patients undergoing TIVA (Peto's OR)	0.24 (0.10 to 0.60)	
High-risk patients undergoing anaesthetic induction with i.v. and maintenance with inhaled anaesthetic (Peto's OR)	0.45 (0.25 to 0.81)	
Baseline awareness in general surgical population	0.16% (0.10% to 0.23%)	Pooled estimate from studies reporting incidence of awareness, not specified to be high risk
Reduction in awareness using depth monitor		
General surgical population undergoing TIVA (Peto's OR)	0.24 (0.10 to 0.60)	Meta-analysis of RCT in high-risk patients, undertaken as part of this review (see *Assessment of outcomes: Bispectral Index*).
General surgical population undergoing anaesthetic induction with i.v. and maintenance with inhaled anaesthetic (Peto's OR)	0.45 (0.25 to 0.81)	Effect assumed to be the same as for high-risk patients
Psychological sequelae of intraoperative awareness		
Probability of LPS, given awareness	0.326 (0.195 to 0.480)	Samuelsson *et al.*[115]
Duration of LPS	6 months	Assumption
Unit cost of LPS	0	Assumption
Utility reduction due to LPS	Same as PTSD	Assumption
Probability of PTSD, given awareness	0.177 (0.113 to 0.230)	Samuelsson *et al.*[115] and Lennmarken *et al.*[11]
Duration of PTSD	12.7 years 7.32 years, discounted at 3.5% (8.2 to 21.6 years, 5.6 to 9.6 discounted at 3.5%)	Kessler *et al.*[118]
Unit cost of PTSD	£9104	NICE [consists of £915.62 (60%) pharmacological therapy, £456.71 (30%) psychological therapy and £152.60 (10%) combined pharmacological and psychological therapy]. Excess risk of hospitalisation 25% annually.[136] Average cost of inpatient stay. *NHS Reference Costs 2010–2011*[137]
Utility reduction due to PTSD	0.12	Various

TABLE 51 Model input parameters – postoperative complication (PONV and POCD)

Parameter	Value	Source
PONV		
Baseline PONV	30%	Cohen MM, Duncan PG, DeBoer DP, Tweed WA. The postoperative interview: assessing risk factors for nausea and vomiting. *Anaesth Analg* 1994;**78**:7–16
Reduction in PONV using depth monitor	Not included in base case	Included as a scenario analysis
Unit cost of PONV	£5.39	£5.39 (4 mg of ondansetron)
Utility reduction due to PONV	0	
POCD		
Baseline POCD	Average duration of 29.65 days	ISPOCD study reported POCD in 25.8% (95% CI 23.1% to 28.5%) of patients at 1 week and in 9.9% (95% CI 8.1% to 12.0%) of patients at 3 months after surgery: compared with 3.4% and 2.8%, respectively, in UK controls. At median follow-up of 532 days, 10.4% patients had cognitive dysfunction compared with 10.6% controls (47 non-hospitalised volunteers of similar age). Assume excess of 22.4% at 7 days, reducing to excess of 7.10% at 3 months and excess of 0% at 1.5 years (532/365.25 years) – area under curve = 29.65 days
Reduction in POCD using depth monitor	Average duration of 21.10 days	Chan and colleagues'[47] abstract reported 32.5% (BIS) vs 39.1% (standard clinical monitoring) at 7 days and 8.1% (BIS) vs 12% (standard clinical monitoring) at 3 months OR estimated as 0.75 (at 7 days) and 0.646 (at 3 months) – applied to excess proportions above. Assume average duration of 21.10 days
Unit cost of POCD	0	
Utility reduction due to POCD	0.05	Jonsson *et al.*:[110] difference in utility between an MMSE score >25 (0.69) and an MMSE score between 21 and 25 (0.64). Normal to mild cognitive dysfunction

Bispectral Index compared with standard clinical monitoring

Base case

Total intravenous anaesthesia The costs, QALY and ICER modelled for patients considered at high risk of intraoperative awareness undergoing GA with TIVA, comparing standard clinical monitoring with BIS are presented in *Table 52*.

Bispectral Index monitoring was modelled as being associated with 10.8 cases of awareness, compared with 45 cases for patients receiving standard clinical monitoring, in cohorts of 10,000 patients. This resulted in a reduction of 11.1 cases of LPS (from 14.7 to 3.5), which included a reduction of six cases of PTSD (from 8.0 to 1.9).

The cost of standard clinical monitoring during anaesthesia in high-risk patients was lower than for BIS, with a cost difference of £15.17. The increased cost for BIS monitoring is primarily the result of the sensors attached to the patient (88% of the cost per patient cost) rather than the module. There is no reduction in anaesthetic costs associated with depth of anaesthesia monitoring in this group of patients, although a small amount of the additional cost of depth of anaesthesia monitoring is offset by reduced costs associated with psychological sequelae of awareness (*Table 53*).

The comparatively high cost of sensors for use with BIS suggests that it is unlikely to generate sufficient savings to offset fully the additional costs of depth of anaesthesia monitoring. This analysis suggests that

TABLE 52 Cost-effectiveness of BIS compared with standard clinical monitoring in a population at high risk of awareness undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	24.19		–0.0011		22,339
BIS	39.36	15.17	–0.0005	0.0007	

TABLE 53 Breakdown of total cost for standard clinical monitoring and BIS for patients at high risk of awareness undergoing TIVA

Cost	Standard clinical monitoring (£)	BIS (£)
Depth of anaesthesia monitoring	0.00	16.43
Anaesthetic drugs	20.92	20.92
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	1.66	0.40

the cost-effectiveness of BIS is likely to be highly dependent on the extent to which it delivers improved patient outcomes (such as reduction in episodes of awareness (and the psychological sequelae) or POCD). A threshold analysis showed that, for patients considered at high risk of intraoperative awareness undergoing GA with TIVA, BIS monitoring would be cost-effective (at a threshold of £30,000 per QALY gained) where the OR for awareness (BIS vs standard clinical monitoring) was <0.458.

The costs, QALY and ICER modelled for a general surgical population undergoing GA with TIVA, comparing standard clinical monitoring with monitoring by BIS are presented in *Table 54*.

Although the cost of standard clinical monitoring in this group of patients was slightly lower than for the subgroup of patients at high risk of intraoperative awareness, the incremental cost of BIS monitoring is lower. This is attributable to the potential to offset a reduction in consumption of anaesthetic against the additional costs of depth of anaesthesia monitoring (*Table 55*). Propofol consumption for maintenance of anaesthesia was estimated as being 19.3% lower in the BIS-monitored group, compared with standard clinical monitoring. Given the lower probability of intraoperative awareness in this group of patients, the QALY losses for standard clinical monitoring and BIS monitoring [resulting from psychological sequelae of awareness (LPS and PTSD)] are lower than for the high-risk group. The QALY gain of 0.0003 was lower than in the high-risk group and results in an increased ICER of £34,565 per QALY gained.

Mixed anaesthesia [induction with intravenous anaesthetic (propofol) and maintenance with inhaled anaesthetic (sevoflurane)] The costs, QALY and ICER modelled for patients considered at high risk of intraoperative awareness undergoing GA, comparing standard clinical monitoring with monitoring by BIS, are presented in *Table 56*.

Bispectral Index monitoring was modelled as being associated with 20.3 cases of awareness, compared with 45 cases among patients receiving standard clinical monitoring, in cohorts of 10,000 patients. This resulted in a reduction of 8.1 cases of LPS (from 14.7 to 6.6), which included a reduction of 4.4 cases of PTSD (from 8.0 to 3.6).

The cost of BSI during anaesthesia in high-risk patients was higher than for standard clinical monitoring, with an incremental cost of £15.52. As discussed previously, the majority of the cost increase with BIS

TABLE 54 Cost-effectiveness of BIS compared with standard clinical monitoring in a general surgical population undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	23.13		–0.0007		
BIS	34.10	10.98	–0.0004	0.0003	34,565

TABLE 55 Breakdown of total cost for standard clinical monitoring and BIS for a general surgical population undergoing TIVA

Cost	Standard clinical monitoring (£)	BIS (£)
Depth of anaesthesia monitoring	0.00	15.46
Anaesthetic drugs	20.92	16.88
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	0.59	0.14

TABLE 56 Cost-effectiveness of BIS compared with standard clinical monitoring in a population at high risk of awareness undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	14.31		–0.0011		
BIS	29.83	15.52	–0.0006	0.0005	29,634

monitoring is attributable to the sensors attached to the patient rather than the depth-monitoring module. As with TIVA in high-risk patients, there is no reduction in anaesthetic costs associated with depth of anaesthesia monitoring and limited scope to offset the additional cost of depth of anaesthesia monitoring by reduction in costs associated with psychological sequelae of awareness (*Table 57*).

Bispectral Index monitoring in a general surgical population undergoing mixed anaesthesia was modelled as being associated with 7.2 cases of awareness, compared with 16 cases among patients receiving standard clinical monitoring. This resulted in a reduction of three cases of LPS (from 5.2 to 2.33), which included a reduction of 1.5 cases of PTSD (from 2.8 to 1.3). The costs, QALY and ICER modelled for this population undergoing mixed GA, comparing standard clinical monitoring with monitoring by BIS, are presented in *Table 58*.

Costs of standard clinical monitoring and BIS monitoring in this group of patients are both lower than for the subgroup of patients at high risk of intraoperative awareness. The cost difference is lower, because of the potential to offset a reduction in consumption of anaesthetic against the additional costs of depth of anaesthesia monitoring (*Table 59*). Sevoflurane consumption for maintenance of anaesthesia was estimated as being 20.2% lower in the BIS-monitored group, compared with standard clinical monitoring. Given the lower probability of intraoperative awareness in this group of patients, the QALY losses for standard clinical monitoring and BIS monitoring are lower than for the high-risk group. The effectiveness of BIS monitoring at reducing intraoperative awareness was also assumed to be lower with inhaled anaesthesia (Peto's OR 0.45) compared with TIVA (Peto's OR 0.24). The QALY gain of 0.0003 was lower than in the high-risk group and results in an increased ICER of £49,198 per QALY gained.

TABLE 57 Breakdown of total cost for standard clinical monitoring and BIS in patients at high risk of awareness undergoing mixed anaesthesia

Cost	Standard clinical monitoring (£)	BIS (£)
Depth of anaesthesia monitoring	0.00	16.43
Anaesthetic drugs	11.04	11.04
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	1.66	0.75

TABLE 58 Cost-effectiveness of depth of anaesthesia monitoring with BIS compared with standard clinical monitoring in a general population undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	13.25		–0.0007		
BIS	26.16	12.91	–0.0004	0.0003	49,198

TABLE 59 Breakdown of total cost for standard clinical monitoring and BIS for a general surgical population undergoing mixed anaesthesia

Cost	Standard clinical monitoring (£)	BIS (£)
Depth of anaesthesia monitoring	0.00	15.46
Anaesthetic drugs	11.04	8.81
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	0.59	0.27

Deterministic sensitivity analysis

Total intravenous anaesthesia One-way sensitivity analyses of key parameters were undertaken in both the general surgical population, and the high-risk surgical population undergoing GA with TIVA. The results are shown in *Tables 60* and *61*.

The changes in the probability of awareness in the patients at high risk of intraoperative awareness receiving TIVA resulted in a substantially altered ICER from the base case: £8196 per QALY gained and £84,305 per QALY gained respectively. The ICER was also sensitive to decreased effectiveness of the BIS module, changes in the probability of LPS, the duration of PTSD at 9.6 years, changes in the probability of PTSD, the lower PTSD decrement and the lower unit cost of sensors. Changes in the duration of LPS, or the LPS QoL decrement, had little impact on the ICER.

These results suggest that the ICER for the general surgical population is relatively robust to changes in the duration of LPS, changes in the QoL decrement applied to LPS, and to the probability of patients seeking treatment for PTSD and the duration of PTSD symptoms.

The ICER appears sensitive to the lower probability of awareness, the relative risk of awareness with BIS modules, the decrease in probability of developing LPS, the decreased probability of developing PTSD and changes in the QoL decrement applied to PTSD.

TABLE 60 One-way sensitivity analysis: BIS compared with standard clinical monitoring in patients at high risk of awareness undergoing TIVA

Parameter	Input value	Standard clinical monitoring		BIS		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost	QALY	
Probability awareness	0.0006	22.76	–0.0005	39.02	–0.0003	16.26	0.0002	84,305
	0.0119	26.92	–0.0024	40.03	–0.0008	13.11	0.0016	8196
Operating room awareness with depth of anaesthesia monitor	0.1	24.19	–0.0011	39.13	–0.0004	14.94	0.0008	19,080
	0.6	24.19	–0.0011	39.96	–0.0007	15.77	0.0004	38,193
Duration of LPS (years)	0.25	24.19	–0.0011	39.36	–0.0005	15.17	0.0007	22,854
	1	24.19	–0.0012	39.36	–0.0005	15.17	0.0007	21,375
Probability of LPS[a]	0.195	23.53	–0.0125	39.20	–0.0121	15.67	0.0004	37,905
	0.48	24.98	–0.0302	39.55	–0.0293	14.58	0.0008	17,239
Duration of PTSD (years)	5.6	24.19	–0.0010	39.36	–0.0004	15.17	0.0006	27,364
	9.6	24.19	–0.0014	39.36	–0.0005	15.17	0.0008	17,969
Proportion PTSD[b]	0.345	23.59	–0.0009	39.22	–0.0004	15.63	0.0005	31,289
	0.733	24.78	–0.0014	39.50	–0.0005	14.73	0.0009	17,259
LPS QoL decrement	–0.075	24.19	–0.0011	39.36	–0.0005	15.17	0.0007	22,723
	–0.05	24.19	–0.0011	39.36	–0.0005	15.17	0.0007	22,942
PTSD QoL decrement	–0.134	24.19	–0.0012	39.36	–0.0005	15.17	0.0007	20,473
	–0.068	24.19	–0.0008	39.36	–0.0004	15.17	0.0004	33,770
Probability people with PTSD seek treatment	0	22.54	–0.0011	38.96	–0.0005	16.43	0.0007	24,191
	1	27.41	–0.0011	40.14	–0.0005	12.73	0.0007	18,742
Cost of sensors (£)	10.875	24.19	–0.0011	35.74	–0.0005	11.54	0.0007	17,000
	18.125	24.19	–0.0011	42.99	–0.0005	18.79	0.0007	27,677

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of the results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

Mixed anaesthesia One-way sensitivity analyses of key parameters were undertaken in both the general surgical population and the high-risk surgical population undergoing mixed GA. The results are shown in *Tables 62* and *63*.

The ICER was sensitive to several key parameters in high-risk patients undergoing mixed anaesthesia. The largest variation is seen where the probability of awareness is decreased to 0.0006 and 0.0119, resulting in an ICER of £11,819 and £94,710 per QALY gained respectively. Changes in the relative risk of awareness with the BIS module, probability of developing LPS or PTSD, the duration of PTSD and a decreased PTSD QoL decrement all lead to large variations in the ICER, ranging from £22,610 to £62,482 per QALY gained.

The ICER is again sensitive to several key parameters in a general surgical population undergoing mixed anaesthesia. In this group, an increase in the probability of LPS resulted in the largest variation, to £67,196 per QALY gained. The ICER was again driven by probability of awareness, the relative risk of awareness

TABLE 61 One-way sensitivity analysis: BIS compared with standard clinical monitoring in a general surgical population undergoing TIVA

Parameter	Input value	Standard clinical monitoring		BIS		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost (£)	QALY	
Proportional change in propofol use	–0.272	23.13	–0.0007	32.45	–0.0004	9.33	0.0003	29,362
	–0.113	23.13	–0.0007	35.78	–0.0004	12.65	0.0003	39,835
Probability awareness	0.001	22.91	–0.0006	34.05	–0.0003	11.15	0.0002	45,913
	0.0023	23.38	–0.0008	34.17	–0.0004	10.78	0.0004	26,630
Operating room awareness with depth of anaesthesia monitor	0.1	23.13	–0.0007	34.02	–0.0003	10.90	0.0004	30,741
	0.6	23.13	–0.0007	34.32	–0.0004	11.19	0.0002	50,184
Duration of LPS (years)	0.25	23.13	–0.0007	34.10	–0.0004	10.98	0.0003	35,168
	1	23.13	–0.0007	34.10	–0.0004	10.98	0.0003	33,419
Probability of LPS[a]	0.195	22.89	–0.0122	34.05	–0.0120	11.16	0.0002	50,006
	0.48	23.40	–0.0295	34.17	–0.0292	10.77	0.0004	28,573
Duration of PTSD (years)	5.6	23.13	–0.0006	34.10	–0.0003	10.98	0.0003	40,178
	9.6	23.13	–0.0007	34.10	–0.0004	10.98	0.0004	29,170
Proportion PTSD[b]	0.345	22.91	–0.0006	34.05	–0.0003	11.14	0.0003	43,915
	0.733	23.33	–0.0008	34.15	–0.0004	10.82	0.0004	28,507
LPS QoL decrement	–0.075	23.13	–0.0007	34.10	–0.0004	10.98	0.0003	35,016
	–0.05	23.13	–0.0007	34.10	–0.0004	10.98	0.0003	35,271
PTSD QoL decrement	–0.134	23.13	–0.0007	34.10	–0.0004	10.98	0.0003	32,324
	–0.068	23.13	–0.0006	34.10	–0.0003	10.98	0.0002	46,553
Probability people with PTSD seek treatment	0	22.54	–0.0007	33.96	–0.0004	11.43	0.0003	35,975
	1	24.27	–0.0007	34.38	–0.0004	10.11	0.0003	31,830
Cost of sensors (£)	10.875	23.13	–0.0007	30.48	–0.0004	7.35	0.0003	23,152
	18.125	23.13	–0.0007	37.73	–0.0004	14.60	0.0003	45,979

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of the results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

with the BIS module (increase and decrease), duration and probability of PTSD, and the unit costs of the sensors.

Scenario analysis

Inclusion of anaesthesia-related complication (postoperative nausea and vomiting) The systematic review of patient outcomes did not identify any robust data which reported an estimate of the effect of BIS monitoring on risk of PONV. We developed a scenario analysis using data from the meta-analysis by Liu[105] to investigate the potential impact of including this outcome on the cost-effectiveness results.

TABLE 62 One-way sensitivity analysis: BIS compared with standard clinical monitoring patients at high risk of awareness undergoing mixed GA

Parameter	Input value	Standard clinical monitoring		BIS		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost (£)	QALY	
Probability awareness	0.0006	12.88	–0.0005	29.18	–0.0003	16.31	0.0002	94,710
	0.0119	17.04	–0.0024	31.07	–0.0012	14.03	0.0012	11,819
Operating room awareness with depth of anaesthesia monitor	0.25	14.31	–0.0011	29.50	–0.0005	15.19	0.0007	22,610
	0.81	14.31	–0.0011	30.43	–0.0009	16.11	0.0003	62,482
Duration of LPS (years)	0.25	14.31	–0.0011	29.83	–0.0006	15.52	0.0005	30,274
	1	14.31	–0.0012	29.83	–0.0006	15.52	0.0005	28,432
Probability of LPS[a]	0.195	13.65	–0.0125	29.53	–0.0122	15.88	0.0003	47,890
	0.48	15.10	–0.0302	30.18	–0.0295	15.09	0.0006	23,430
Duration of PTSD (years)	5.6	14.31	–0.0010	29.83	–0.0006	15.52	0.0004	35,798
	9.6	14.31	–0.0014	29.83	–0.0007	15.52	0.0006	24,132
Proportion PTSD[b]	0.345	13.71	–0.0009	29.56	–0.0005	15.85	0.0004	40,248
	0.733	14.90	–0.0014	30.09	–0.0007	15.20	0.0006	23,396
LPS QoL decrement	–0.075	14.31	–0.0011	29.83	–0.0006	15.52	0.0005	30,112
	–0.05	14.31	–0.0011	29.83	–0.0006	15.52	0.0005	30,383
PTSD QoL decrement	–0.134	14.31	–0.0012	29.83	–0.0007	15.52	0.0006	27,301
	–0.068	14.31	–0.0008	29.83	–0.0005	15.52	0.0004	43,413
Probability people with PTSD seek treatment	0	12.66	–0.0011	29.08	–0.0006	16.43	0.0005	31,371
	1	17.53	–0.0011	31.28	–0.0006	13.75	0.0005	26,262
Cost of sensors (£)	10.875	14.31	–0.0011	26.21	–0.0006	11.89	0.0005	22,711
	18.125	14.31	–0.0011	33.46	–0.0006	19.14	0.0005	36,557

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of the results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

For this scenario analysis we used the baseline (control group) risk of PONV as the estimated risk for standard clinical monitoring and applied the OR derived in the meta-analysis (0.77, 95% CI 0.56 to 0.99) and the lower limit of the 95% CI to estimate risk for BIS-monitored patients. We assumed that all treatments (such as prophylaxis against PONV) were the same for each treatment group, and that all patients experiencing PONV were treated using 4 mg ondansetron by intramuscular or slow i.v. injection (unit cost = £5.39; BNF[33]).

Tables 64 and *65* report the results of the scenario analysis for patients at high risk of intraoperative awareness and a general surgical population, respectively, undergoing GA with TIVA. The incremental costs for BIS monitoring are reduced, from the value reported for the base-case analyses (*Tables 52*

TABLE 63 One-way sensitivity analysis: BIS compared with standard clinical monitoring in a general surgical population undergoing mixed GA

Parameter	Input value	Standard clinical care		BIS		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost (£)	QALY	
Proportional change in sevoflurane use	–0.330	13.25	–0.0007	24.74	–0.0004	11.50	0.0003	43,813
	–0.074	13.25	–0.0007	27.57	–0.0004	14.32	0.0003	54,583
Probability awareness	0.001	13.03	–0.0006	26.06	–0.0004	13.03	0.0002	62,569
	0.0023	13.50	–0.0008	26.27	–0.0005	12.77	0.0003	39,224
Operating room awareness with depth of anaesthesia monitor	0.25	13.25	–0.0007	26.04	–0.0004	12.79	0.0003	40,611
	0.81	13.25	–0.0007	26.37	–0.0005	13.12	0.0002	78,178
Duration of LPS (years)	0.25	13.25	–0.0007	26.16	–0.0004	12.91	0.0003	49,948
	1	13.25	–0.0007	26.16	–0.0004	12.91	0.0003	47,764
Probability of LPS[a]	0.195	13.01	–0.0122	26.05	–0.0121	13.04	0.0002	67,196
	0.48	13.52	–0.0295	26.28	–0.0292	12.76	0.0003	41,792
Duration of PTSD (years)	5.6	13.25	–0.0006	26.16	–0.0004	12.91	0.0002	56,055
	9.6	13.25	–0.0007	26.16	–0.0004	12.91	0.0003	42,340
Proportion PTSD[b]	0.345	13.03	–0.0006	26.06	–0.0004	13.03	0.0002	60,266
	0.733	13.45	–0.0008	26.25	–0.0004	12.80	0.0003	41,648
LPS QoL decrement	–0.075	13.25	–0.0007	26.16	–0.0004	12.91	0.0003	49,758
	–0.05	13.25	–0.0007	26.16	–0.0004	12.91	0.0003	50,075
PTSD QoL decrement	–0.134	13.25	–0.0007	26.16	–0.0004	12.91	0.0003	46,382
	–0.068	13.25	–0.0006	26.16	–0.0004	12.91	0.0002	63,521
Probability people with PTSD seek treatment	0	12.66	–0.0007	25.89	–0.0004	13.23	0.0003	50,432
	1	14.39	–0.0007	26.67	–0.0004	12.28	0.0003	46,803
Unit cost of sensors (£)	10.875	13.25	–0.0007	22.53	–0.0004	9.28	0.0003	35,383
	18.125	13.25	–0.0007	29.78	–0.0004	16.53	0.0003	63,013

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of the results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

and *54*), by including an estimate of PONV. However, the change in costs is slight and leaves the ICER largely unchanged.

Tables 66 and *67* report the results of the scenario analysis for patients at high risk of intraoperative awareness and a general surgical population, respectively, undergoing mixed GA. As before, the incremental costs for BIS monitoring are reduced. However, the change in costs is slight and leaves the ICER largely unchanged.

Inclusion of the impact of PONV with BIS monitoring into the base-case analysis is unlikely to substantially affect decisions based on cost-effectiveness criteria.

TABLE 64 Scenario analysis: including an estimated effect of BIS monitoring on incidence of PONV in patients at high risk of awareness undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
OR = 0.77: baseline risk = 0.3, risk with BIS monitoring = 0.248					
Standard clinical monitoring	24.19		–0.0011		
BIS	39.08	14.89	–0.0005	0.0007	21,927
OR = 0.56[a]: baseline risk = 0.3, risk with BIS monitoring = 0.194					
Standard clinical monitoring	24.19		–0.0011		
BIS	38.79	14.60	–0.0005	0.0007	21,494

a Lower limit of 95% CI estimated by Liu.[105]

TABLE 65 Scenario analysis: including an estimated effect of BIS monitoring on incidence of PONV in a general surgical population undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
OR = 0.77: baseline risk = 0.375, risk with BIS monitoring = 0.248					
Standard clinical monitoring	23.13		–0.0007		
BIS	33.82	10.70	–0.0004	0.0003	33,685
OR = 0.56:[a] baseline risk = 0.375, risk with BIS monitoring = 0.194					
Standard clinical monitoring	23.13		–0.0007		
BIS	33.53	10.40	–0.0004	0.0003	32,759

a Lower limit of 95% CI estimated by Liu.[105]

TABLE 66 Scenario analysis: including an estimated effect of BIS monitoring on incidence of PONV in patients at high risk of awareness undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
OR = 0.77: baseline risk = 0.3, risk with BIS monitoring = 0.248					
Standard clinical monitoring	14.31		–0.0011		
BIS	29.55	15.24	–0.0006	0.0005	29,100
OR = 0.56:[a] baseline risk = 0.3, risk with BIS monitoring = 0.194					
Standard clinical monitoring	14.31		–0.0011		
BIS	29.26	14.94	–0.0006	0.0005	28,538

a Lower limit of 95% CI estimated by Liu.[105]

TABLE 67 Scenario analysis: including an estimated effect of BIS monitoring on incidence of PONV in a general surgical population undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
OR = 0.77: baseline risk = 0.3, risk with BIS monitoring = 0.248					
Standard clinical monitoring	13.25		–0.0007		
BIS	25.88	12.63	–0.0004	0.0003	48,132
OR = 0.56:[a] baseline risk = 0.3, risk with BIS monitoring = 0.194					
Standard clinical monitoring	13.25		–0.0007		
BIS	25.58	12.34	–0.0004	0.0003	47,011

a Lower limit of 95% CI estimated by Liu.[105]

Scenario analyses for probability of intraoperative awareness for patients at high risk of intraoperative awareness and for the general surgical population

Our review of published studies of the incidence of intraoperative awareness identified substantial uncertainty over the estimated values. We used pooled values across identified studies in the base-case analysis. However, the value adopted for 'high risk' is lower than that commonly quoted as indicating high risk, and the pooled estimate adopted for a general surgical population excluded two outlying studies (one high and one low extreme value).

For this scenario analysis we replace the base-case estimate for probability of awareness in the high-risk population (0.45%) with a value of 1.0% reported for certain types of surgery (cardiac surgery, caesarean section and trauma surgery).[79,111,112] The effect of this is to approximately double the QALY loss for each group, resulting in a doubling of the QALY gain associated with BIS monitoring, while incremental costs are largely unchanged. The effect of this is to reduce the ICER by about half (*Table 68*).

In the general surgical population, we replaced the base-case estimate for probability of awareness (0.16%) with the incidences reported in the two outlying studies (*Tables 69* and *70*). The results from these two scenarios contrast sharply. At the highest reported incidence of awareness – equivalent to that frequently cited for 'high-risk' populations – the QALY loss for each group increases approximately 2.5-fold, resulting in a three- to fourfold increase in the QALY gain associated with BIS monitoring. The incremental costs are slightly reduced, compared with the base case, and the resulting ICERs are substantially reduced. In the case of the lowest reported probability of awareness, the QALY gain from BIS monitoring is negligible resulting in high-value ICER.

Impact of assumptions on number of patients per device-year

In order to apportion the capital cost of the depth of anaesthesia monitoring modules, we required an estimate of the number of patients in whom the monitor module was used in each year (patients per device-year), throughout its assumed 5-year effective life. The estimate used for the general surgical population was 1000 patients per year (equivalent to four patients per day over 250 working days per year) was based on discussion with clinical experts. This scenario analysis investigates the impact of this assumption on the estimated incremental cost associated with BIS monitoring, compared with standard clinical monitoring, and the resulting effect on the ICER. *Tables 71* and *72* report the incremental cost and ICER for BIS, compared with standard clinical monitoring, at four selected values for the number of patients per device-year: the base-case value of 100 and also for a low value of 10 and a high value of 1500 (six patients per day over 250 working days per year). This suggests that the assumed number of patients per device-year only has a substantial impact on incremental cost (hence on the ICER) at comparatively low volumes.

TABLE 68 Scenario analysis: impact of baseline risk of awareness on cost-effectiveness of BIS monitoring for patients at high risk of awareness

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
TIVA					
Standard clinical monitoring	26.22		–0.0021		
BIS	39.85	13.64	–0.0007	0.0014	10,003
Mixed anaesthesia					
Standard clinical monitoring	16.34		–0.0021		
BIS	30.75	14.41	–0.0010	0.0010	14,168

TABLE 69 Scenario analysis: impact of baseline risk of awareness on cost-effectiveness of BIS monitoring for a general surgical population undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Baseline probability of awareness = 0.99%					
Standard clinical monitoring	26.18		–0.0020		
BIS	34.84	8.66	–0.0007	0.0014	6413
Baseline probability of awareness = 0.007%					
Standard clinical monitoring	22.56		–0.0004		
BIS	33.97	11.41	–0.0003	0.0001	90,014

TABLE 70 Scenario analysis: impact of baseline risk of awareness on cost-effectiveness of BIS monitoring for a general surgical population undergoing mixed GA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Baseline probability of awareness = 0.99%					
Standard clinical monitoring	16.30		–0.0020		
BIS	27.54	11.24	–0.0010	0.0010	11,146
Baseline probability of awareness = 0.007%					
Standard clinical monitoring	12.68		–0.0004		
BIS	25.90	13.22	–0.0003	0.0001	106,347

TABLE 71 Scenario analysis: impact of number of patients per device-year on cost-effectiveness of BIS monitoring in patients undergoing TIVA

Patients per device-year	High-risk patients undergoing TIVA		General surgical population undergoing TIVA	
	Incremental cost (£)	ICER (£/QALY gained)	Incremental cost (£)	ICER (£/QALY gained)
100	22.88	33,689	19.65	61,866
500	15.17	22,339	11.94	37,599
1000	14.21	20,920	10.98	34,565
1500	13.88	20,447	10.66	33,554

Impact of utility decrement for PTSD The QoL decrement applied in the base-case analysis was based on Freed and colleagues'[120] paper on veterans with PTSD. In order to investigate the impact of a sparse evidence base on HRQoL in a group of patients with PTSD, a scenario analysis was undertaken. The utility decrement was adjusted to 0.50 and 0.75 in high-risk and general surgical groups receiving either TIVA or mixed anaesthesia (*Table 73*).

The ICER was sensitive to these alternative scenarios in high-risk patients, both receiving TIVA and mixed anaesthesia. Where the PTSD decrement was increased to –0.5 in TIVA and mixed anaesthesia the ICER reduced to £6431 per QALY gained and £8928 per QALY gained respectively. Where the PTSD decrement was increased further, the ICER decreased again to £4379 and £6116 per QALY gained in the TIVA and mixed anaesthesia groups respectively.

The scenario analyses using alternative PTSD decrements in the general surgical population reflect the results in the high-risk population: there is a substantial reduction in the ICER where these are increased (*Table 74*).

TABLE 72 Scenario analysis: impact of number of patients per device-year on cost-effectiveness of BIS monitoring in patients undergoing mixed anaesthesia

Patients per device-year	High-risk patients undergoing mixed anaesthesia		General surgical population undergoing mixed anaesthesia	
	Incremental cost (£)	ICER (£/QALY gained)	Incremental cost (£)	ICER (£/QALY gained)
100	23.22	44,354	21.58	82,243
500	15.52	29,634	13.87	52,870
1000	14.55	27,794	12.91	49,198
1500	14.23	27,181	12.59	47,974

TABLE 73 Scenario analysis: impact of utility decrement for PTSD on cost-effectiveness of BIS in patients at high risk of awareness undergoing TIVA or mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = –0.5					
Standard clinical monitoring	24.19		–0.0034		
BIS	39.36	15.17	–0.0010	0.0024	6431
Utility decrement for PTSD = –0.75					
Standard clinical monitoring	24.19		–0.0048		
BIS	39.36	15.17	–0.0014	0.0035	4379
Mixed anaesthesia					
Utility decrement for PTSD = –0.5					
Standard clinical monitoring	14.31		–0.0034		
BIS	29.83	15.52	–0.0016	0.0017	8928
Utility decrement for PTSD = –0.75					
Standard clinical monitoring	14.31		–0.0048		
BIS	29.83	15.52	–0.0023	0.0025	6116

E-Entropy compared with standard clinical monitoring

Base case

Total intravenous anaesthesia The costs, QALY and ICER modelled for patients considered at high risk of intraoperative awareness undergoing GA with TIVA, comparing standard clinical monitoring with monitoring by E-Entropy, are presented in *Table 75*.

TABLE 74 Scenario analysis: impact of utility decrement for PTSD cost-effectiveness of BIS in a general surgical population undergoing TIVA or mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = –0.5					
Standard clinical monitoring	23.13		–0.0015		
BIS	34.10	10.98	–0.0005	0.0009	11,994
Utility decrement for PTSD = –0.75					
Standard clinical monitoring	23.13		–0.0020		
BIS	34.10	10.98	–0.0007	0.0013	8390
Mixed anaesthesia					
Utility decrement for PTSD = –0.5					
Standard clinical monitoring	13.25		–0.0015		
BIS	26.16	12.91	–0.0008	0.0007	18,581
Utility decrement for PTSD = –0.75					
Standard clinical monitoring	13.25		–0.0020		
BIS	26.16	12.91	–0.0010	0.0010	13,183

TABLE 75 Cost-effectiveness of E-Entropy compared with standard clinical monitoring in a population at high risk of awareness undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	26.38		–0.0011		
E-Entropy	36.18	9.79	–0.0005	0.0007	14,421

TABLE 76 Breakdown of total cost for standard clinical monitoring and E-Entropy in patients at high risk of awareness undergoing TIVA

Cost	Standard clinical monitoring (£)	E-Entropy (£)
Depth of anaesthesia monitoring	0.00	11.05
Anaesthetic drugs	23.11	23.11
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	1.66	0.40

E-Entropy monitoring was modelled as being associated with 10.8 cases of awareness, compared with 45 cases in patients receiving standard clinical monitoring, in cohorts of 10,000 patients. This resulted in a reduction of 11.1 cases of LPS (from 14.7 to 3.5), which included a reduction of six cases of PTSD (from 8.0 to 1.9).

The cost of standard clinical monitoring during anaesthesia in high-risk patients was lower than for E-Entropy monitoring, with the incremental cost being £9.79. The breakdown of total cost for standard clinical monitoring and E-Entropy is reported in *Table 76*; the costs of anaesthetic drug use outlined in this table apply to the Ellerkmann and colleagues study only.[62] As no reduction in drug costs is expected in the population at high risk of awareness, the cost assumption (for anaesthetic drugs) has no impact on the ICER. The increased cost for E-Entropy monitoring is partially offset by the reduction in costs of patients with PTSD.

As a result of the psychological sequelae of awareness, including LPS, PTSD and POCD, patients in both groups incurred a slight QALY loss. This was lower in the E-Entropy-monitored patients, with a difference of 0.0007 QALY, resulting in an ICER of £14,421 per QALY gained.

In a general surgical population (not just those at high risk of intraoperative awareness) undergoing GA with TIVA, E-Entropy monitoring was modelled as being associated with 3.8 cases of awareness, compared with 16 cases for patients receiving standard clinical monitoring, in cohorts of 10,000 patients. This resulted in a reduction of four cases of LPS (from 5.2 to 1.3), which included a reduction of 2.1 cases of PTSD (from 2.8 to 0.7). The costs, QALY and ICER modelled for this population, comparing standard clinical monitoring with monitoring by E-Entropy are presented in *Table 77* (based on anaesthetic drug consumption from the RCT by Ellerkmann and colleagues[62]) and in *Table 78* (based on anaesthetic drug consumption from the RCT by Gruenewald and colleagues[55]).

Applying the costs of anaesthetic drugs from both the Ellerkmann and colleagues[62] and Gruenewald and colleagues[55] RCTs results in increased costs with E-Entropy. Both RCTs reported slightly lower costs for anaesthetic drug use in the standard clinical monitoring group than with the E-Entropy group. Again, costs for PTSD were slightly lower in the E-Entropy group as a result of lower incidence of awareness (*Table 79*).

The QALY loss incurred by patients undergoing E-Entropy monitoring was slightly less than that of patients in the standard clinical monitoring group, giving an incremental QALY gain of 0.0003. This resulted in an ICER of £31,131 per QALY gained where the anaesthetic consumption from the Ellerkmann and colleagues

TABLE 77 Cost-effectiveness of E-Entropy compared with standard clinical monitoring in a general surgical population undergoing TIVA (drug use based on Ellerkmann *et al.*[62])

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	25.32		–0.0007		
E-Entropy	35.20	9.89	–0.0004	0.0003	31,131

TABLE 78 Cost-effectiveness of E-Entropy compared with standard clinical monitoring in a general surgical population undergoing TIVA (drug use based on Gruenewald *et al.*[55])

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	31.50		–0.0007		
E-Entropy	41.48	9.98	–0.0004	0.0003	31,430

RCT[62] were applied, and £31,430 where anaesthetic consumption from Gruenewald and colleagues[55] were applied.

Mixed anaesthesia [induction with intravenous anaesthetic (propofol and sufentanil) and maintenance with intravenous and inhaled anaesthetic (sufentanil and sevoflurane)] The costs, QALY and ICER modelled for patients considered at high risk of intraoperative awareness undergoing mixed anaesthesia, comparing standard clinical monitoring with monitoring by E-Entropy are presented in *Table 80*.

E-Entropy monitoring was modelled as being associated with 20.3 cases of awareness, compared with 45 cases among patients receiving standard clinical monitoring, in cohorts of 10,000 patients. This resulted in a reduction of 8.1 cases of LPS (from 14.7 to 6.6), which included a reduction of 4.4 cases of PTSD (from 8.0 to 3.6).

The costs of anaesthetic drugs in each group were the same, as shown in the breakdown of total cost in *Table 81*. Sufentanil costs are not included as it is not available in the UK and therefore the costs are not available in the BNF. Given the reduced incidence of awareness, and consequent reduction in cases of

TABLE 79 Breakdown of total cost for standard clinical monitoring and E-Entropy in a general surgical population undergoing TIVA

Cost	Standard clinical monitoring (£)	E-Entropy (£)
Depth of anaesthesia monitoring	0.00	9.87
Anaesthetic drugs		
Gruenewald *et al.*[55]	29.29	29.85
Ellerkmann *et al.*[62]	23.11	23.58
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	0.59	0.14

TABLE 80 Cost-effectiveness of E-Entropy compared with standard clinical monitoring in a population at high risk of awareness undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	19.20		–0.0011		
E-Entropy	29.35	10.14	–0.0006	0.0005	19,367

TABLE 81 Breakdown of total cost for standard clinical monitoring and E-Entropy in a population at high risk of awareness undergoing mixed anaesthesia

Cost	Standard clinical monitoring (£)	E-Entropy (£)
Depth of anaesthesia monitoring	0.00	11.05
Anaesthetic drugs	15.93	15.93
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	1.66	0.75

PTSD, costs for PTSD were lower in the group undergoing E-Entropy monitoring. The incremental cost of E-Entropy monitoring was £10.14.

Again, each group incurred a QALY loss as a result of psychological sequelae such as LPS and PTSD, which resulted in an incremental QALY gain for E-Entropy patients of 0.0005. This yielded an ICER of £19,367 per QALY gained.

In a general surgical population E-Entropy monitoring was modelled as being associated with 7.2 cases of awareness, compared with 16 cases in patients receiving standard clinical monitoring, in cohorts of 10,000 patients. This resulted in a reduction of three cases of LPS (from 5.2 to 2.3), which included a reduction of 1.5 cases of PTSD (from 2.8 to 1.3). The costs, QALYs and ICER modelled for this population undergoing GA with both i.v. and inhaled anaesthetic, comparing standard clinical monitoring with monitoring by E-Entropy, are presented in *Table 82*.

In a general surgical population undergoing mixed anaesthesia with sufentanil and sevoflurane, the costs of E-Entropy monitoring were higher, with an incremental cost of £4.99 (*Table 83*). Costs of anaesthetic drugs were lower in the E-Entropy arm, as were costs associated with PTSD, offsetting a proportion of the additional costs associated with depth of anaesthesia monitoring.

The general surgical population accrued a slightly lower incremental QALY gain of 0.0003, which resulted in an ICER of £19,000 per QALY gained.

Deterministic sensitivity analysis

Total intravenous anaesthesia One-way sensitivity analyses of key parameters were undertaken in both the general surgical population and the high-risk surgical population undergoing general anaesthetic using TIVA. The results for the high-risk surgical population are shown in *Table 84*. Here the anaesthetic drug costs are based on Ellerkmann and colleagues' study.[62] As there is no expected reduction in drug use in this high-risk population, this assumption has no overall impact: anaesthetic drug costs are the same for both standard clinical monitoring and E-Entropy and therefore cancel out in the calculation of incremental cost and in the ICER.

TABLE 82 Cost-effectiveness of E-Entropy compared with standard clinical monitoring in a general surgical population undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	18.14		–0.0007		
E-Entropy	23.12	4.99	–0.0004	0.0003	19,000

TABLE 83 Breakdown of total cost for standard clinical monitoring and E-Entropy in a general surgical population undergoing mixed anaesthesia

Cost	Standard clinical monitoring (£)	E-Entropy (£)
Depth of anaesthesia monitoring	0.00	9.87
Anaesthetic drugs	15.93	11.37
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	0.59	0.27

TABLE 84 One-way sensitivity analysis: E-Entropy compared with standard clinical monitoring in patients at high risk of awareness undergoing TIVA

Parameter	Input value	Standard clinical monitoring		E-Entropy		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost (£)	QALY	
Probability awareness	0.0006	24.95	–0.0005	35.83	–0.0003	10.88	0.0002	56,429
	0.0119	29.11	–0.0024	36.84	–0.0008	7.73	0.0016	4834
Operating room awareness with depth of anaesthesia monitor	0.1	26.38	–0.0011	35.94	–0.0004	9.56	0.0008	12,212
	0.6	26.38	–0.0011	36.77	–0.0007	10.39	0.0004	25,169
Duration of LPS (years)	0.25	26.38	–0.0011	36.18	–0.0005	9.79	0.0007	14,754
	1	26.38	–0.0012	36.18	–0.0005	9.79	0.0007	13,799
Probability of LPS[a]	0.195	25.72	–0.0125	36.02	–0.0121	10.30	0.0004	24,904
	0.48	27.17	–0.0302	36.37	–0.0293	9.20	0.0008	10,880
Duration of PTSD (years)	5.6	26.38	–0.0010	36.18	–0.0004	9.79	0.0006	17,666
	9.6	26.38	–0.0014	36.18	–0.0005	9.79	0.0008	11,601
Proportion PTSD[b]	0.345	25.78	–0.0009	36.03	–0.0004	10.25	0.0005	20,524
	0.733	26.97	–0.0014	36.32	–0.0005	9.35	0.0009	10,958
LPS QoL decrement	–0.075	26.38	–0.0011	36.18	–0.0005	9.79	0.0007	14,669
	–0.05	26.38	–0.0011	36.18	–0.0005	9.79	0.0007	14,811
PTSD QoL decrement	–0.134	26.38	–0.0012	36.18	–0.0005	9.79	0.0007	13,217
	–0.068	26.38	–0.0008	36.18	–0.0004	9.79	0.0004	21,801
Probability people with PTSD seek treatment	0	24.73	–0.0011	35.78	–0.0005	11.05	0.0007	16,274
	1	29.60	–0.0011	36.95	–0.0005	7.35	0.0007	10,825
Unit cost of sensors (£)	6.51	26.38	–0.0011	34.01	–0.0005	7.62	0.0007	11,226
	10.85	26.38	–0.0011	38.35	–0.0005	11.96	0.0007	17,617

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of the results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

The ICER resulting from the one-way sensitivity analysis in a high-risk population receiving TIVA ranged from £4834 to £56,429 per QALY gained. The ICER was insensitive to decreases in the LPS QoL decrement, and to the unit costs of sensors, but a decrease in the PTSD decrement pushed the ICER up to £21,801 per QALY gained from the base case of £14,421. The ICER appears driven by changes in the effectiveness of the E-Entropy module: where the relative risk of awareness is increased to 0.6, the ICER increases to £25,169 per QALY gained. Similarly, the ICER was very sensitive to changes in the probability of awareness. A decrease in this probability to 0.0006 increases the ICER substantially to £56,429 per QALY gained. Conversely, an increase in this probability to 0.0119 decreased the ICER to £4834 per QALY gained.

The results for the one-way sensitivity analyses in the general surgical population are shown in *Table 85* (anaesthetic drug costs based on usage reported by Ellerkmann and colleagues[62]) and *Table 86* (anaesthetic drug costs based on usage reported by Gruenewald and colleagues[55]).

TABLE 85 One-way sensitivity analysis: E-Entropy compared with standard clinical monitoring in a general surgical population undergoing TIVA (drug use based on Ellerkmann *et al.*[62])

Parameter	Input value	Standard clinical monitoring		E-Entropy		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost (£)	QALY	
Proportional change in propofol use	–0.075	25.32	–0.0007	32.85	–0.0004	7.53	0.0003	23,712
	0.174	25.32	–0.0007	37.54	–0.0004	12.22	0.0003	38,490
Proportional change in remifentanil	–0.232	25.32	–0.0007	34.69	–0.0004	9.37	0.0003	29,508
	0.010	25.32	–0.0007	35.72	–0.0004	10.40	0.0003	32,754
Probability awareness	0.0010	25.10	–0.0006	35.15	–0.0003	10.06	0.0002	41,419
	0.0023	25.57	–0.0008	35.27	–0.0004	9.69	0.0004	23,936
Operating room awareness with depth of anaesthesia monitor	0.1	25.32	–0.0007	35.12	–0.0003	9.80	0.0004	27,663
	0.6	25.32	–0.0007	35.42	–0.0004	10.10	0.0002	45,292
Duration of LPS (years)	0.25	25.32	–0.0007	35.20	–0.0004	9.89	0.0003	31,674
	1	25.32	–0.0007	35.20	–0.0004	9.89	0.0003	30,099
Probability of LPS[a]	0.195	25.08	–0.0122	35.15	–0.0120	10.07	0.0002	45,117
	0.48	25.59	–0.0295	35.27	–0.0292	9.68	0.0004	25,678
Duration of PTSD (years)	5.6	25.32	–0.0006	35.20	–0.0003	9.89	0.0003	36,186
	9.6	25.32	–0.0007	35.20	–0.0004	9.89	0.0004	26,271
Proportion PTSD[b]	0.345	25.10	–0.0006	35.15	–0.0003	10.05	0.0003	39,615
	0.733	25.52	–0.0008	35.25	–0.0004	9.73	0.0004	25,633
LPS QoL decrement	–0.075	25.32	–0.0007	35.20	–0.0004	9.89	0.0003	31,536
	–0.05	25.32	–0.0007	35.20	–0.0004	9.89	0.0003	31,766
PTSD QoL decrement	–0.134	25.32	–0.0007	35.20	–0.0004	9.89	0.0003	29,112
	–0.068	25.32	–0.0006	35.20	–0.0003	9.89	0.0002	41,927
Probability people with PTSD seek treatment	0	24.73	–0.0007	35.06	–0.0004	10.34	0.0003	32,540
	1	26.46	–0.0007	35.48	–0.0004	9.02	0.0003	28,395
Unit cost of sensors (£)	6.51	25.32	–0.0007	33.03	–0.0004	7.72	0.0003	24,298
	10.85	25.32	–0.0007	37.37	–0.0004	12.06	0.0003	37,963

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of the results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

The one-way sensitivity analysis results in the general surgical population undergoing TIVA, and using costs applied from Ellerkmann and colleagues[62] (*Table 85*) reflects those in the high-risk population. Again, the results are generally insensitive to changes in the duration of LPS, and the LPS QoL decrement. The greatest changes in ICERs were again generated as a result of changes in the probability of awareness (£23,936 per QALY gained, and £41,419 per QALY gained), a reduction in effectiveness of the E-Entropy module (£45,292 per QALY gained,) and the probability of LPS and a reduction in the PTSD QoL decrement.

Again, the one-way sensitivity analysis in the general surgical population receiving TIVA and applying costs from Gruenewald and colleagues[55] (*Table 86*) reflect the results in the high-risk group. Whereas the ICER

TABLE 86 One-way sensitivity analysis: E-Entropy compared with standard clinical monitoring in a general surgical population undergoing TIVA (drug use based on Gruenewald *et al.*[55])

Parameter	Input value	Standard clinical monitoring		Entropy		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost (£)	QALY	
Proportional change in propofol use	–0.237	31.50	–0.0007	40.19	–0.0004	8.69	0.0003	27,364
	–0.058	31.50	–0.0007	42.76	–0.0004	11.26	0.0003	35,452
Proportional change in remifentanil	0.085	31.50	–0.0007	40.07	–0.0004	8.58	0.0003	27,009
	0.274	31.50	–0.0007	42.90	–0.0004	11.40	0.0003	35,899
Probability awareness	0.001	31.28	–0.0006	41.43	–0.0003	10.15	0.0002	41,811
	0.0023	31.75	–0.0008	41.54	–0.0004	9.79	0.0004	24,171
Operating room awareness with depth of anaesthesia monitor	0.1	31.50	–0.0007	41.40	–0.0003	9.90	0.0004	27,932
	0.6	31.50	–0.0007	41.69	–0.0004	10.19	0.0002	45,719
Duration of LPS (years)	0.25	31.50	–0.0007	41.48	–0.0004	9.98	0.0003	31,979
	1	31.50	–0.0007	41.48	–0.0004	9.98	0.0003	30,388
Probability of LPS[a]	0.195	31.26	–0.0122	41.42	–0.0120	10.16	0.0002	45,544
	0.48	31.77	–0.0295	41.55	–0.0292	9.77	0.0004	25,931
Duration of PTSD (years)	5.6	31.50	–0.0006	41.48	–0.0003	9.98	0.0003	36,534
	9.6	31.50	–0.0007	41.48	–0.0004	9.98	0.0004	26,524
Proportion PTSD[b]	0.345	31.28	–0.0006	41.43	–0.0003	10.15	0.0003	39,990
	0.733	31.70	–0.0008	41.53	–0.0004	9.82	0.0004	25,884
LPS QoL decrement	–0.075	31.50	–0.0007	41.48	–0.0004	9.98	0.0003	31,840
	–0.05	31.50	–0.0007	41.48	–0.0004	9.98	0.0003	32,072
PTSD QoL decrement	–0.134	31.50	–0.0007	41.48	–0.0004	9.98	0.0003	29,393
	–0.068	31.50	–0.0006	41.48	–0.0003	9.98	0.0002	42,331
Probability people with PTSD seek treatment	0	30.91	–0.0007	41.34	–0.0004	10.43	0.0003	32,840
	1	32.64	–0.0007	41.75	–0.0004	9.11	0.0003	28,695
Unit cost of sensors (£)	6.51	31.50	–0.0007	39.31	–0.0004	7.81	0.0003	24,598
	10.85	31.50	–0.0007	43.65	–0.0004	12.15	0.0003	38,263

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of the results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

appears relatively insensitive to the changes in LPS QoL and LPS duration, the key parameters driving the results are a reduction in the probability of awareness, an increase in the relative risk of awareness with the E-Entropy module, a reduction in the probability of LPS and a reduction in the PTSD decrement applied.

Mixed anaesthesia One-way sensitivity analyses of key parameters were undertaken in both the general surgical population and the high-risk surgical population undergoing general anaesthetic using mixed anaesthesia [induction with i.v. anaesthetic (remifentanil) and maintenance with i.v. and inhaled anaesthetic (remifentanil and sevoflurane)]. The results are shown in *Tables 87* and *88*.

TABLE 87 One-way sensitivity analysis: E-Entropy compared with standard clinical monitoring in patients at high risk of awareness undergoing mixed anaesthesia

Parameter	Input value	Standard clinical monitoring		E-Entropy		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost (£)	QALY	
Probability awareness	0.0006	17.77	–0.0005	28.70	–0.0003	10.93	0.0002	63,483
	0.0119	21.93	–0.0024	30.58	–0.0012	8.65	0.0012	7290
Operating room awareness with depth of anaesthesia monitor	0.25	19.20	–0.0011	29.01	–0.0005	9.81	0.0007	14,605
	0.81	19.20	–0.0011	29.94	–0.0009	10.74	0.0003	41,635
Duration of LPS (years)	0.25	19.20	–0.0011	29.35	–0.0006	10.14	0.0005	19,785
	1	19.20	–0.0012	29.35	–0.0006	10.14	0.0005	18,582
Probability of LPS[a]	0.195	18.54	–0.0125	29.04	–0.0122	10.51	0.0003	31,680
	0.48	19.99	–0.0302	29.70	–0.0295	9.71	0.0006	15,082
Duration of PTSD (years)	5.6	19.20	–0.0010	29.35	–0.0006	10.14	0.0004	23,395
	9.6	19.20	–0.0014	29.35	–0.0007	10.14	0.0006	15,771
Proportion PTSD	0.345	18.60	–0.0009	29.07	–0.0005	10.47	0.0004	26,595
	0.733	19.79	–0.0014	29.61	–0.0007	9.82	0.0006	15,119
LPS QoL decrement	–0.075	19.20	–0.0011	29.35	–0.0006	10.14	0.0005	19,679
	–0.05	19.20	–0.0011	29.35	–0.0006	10.14	0.0005	19,857
PTSD QoL decrement	–0.134	19.20	–0.0012	29.35	–0.0007	10.14	0.0006	17,843
	–0.068	19.20	–0.0008	29.35	–0.0005	10.14	0.0004	28,372
Probability people with PTSD seek treatment	0	17.55	–0.0011	28.60	–0.0006	11.05	0.0005	21,104
	1	22.42	–0.0011	30.80	–0.0006	8.38	0.0005	15,995
Unit cost of sensors (£)	6.51	19.20	–0.0011	27.18	–0.0006	7.97	0.0005	15,223
	10.85	19.20	–0.0011	31.52	–0.0006	12.31	0.0005	23,511

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of the results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

The results of the one-way sensitivity analysis in high-risk patients receiving mixed anaesthesia reflect those in patients receiving TIVA. The ICER in a high-risk surgical group receiving mixed anaesthesia is very sensitive to both increase and decrease in the probability of awareness (*Table 87*), resulting in ICER of £7290 per QALY gained and £63,483 per QALY gained respectively. The ICER was also sensitive to increase in the relative risk of awareness with the E-Entropy module, giving £41,635 per QALY gained. Again, the ICER was sensitive to changes in the probability of LPS, a decrease in the probability of PTSD, and a decrease in the PTSD QoL decrement, while being insensitive to the LPS decrement and duration.

In the general surgical population the largest variation in the ICER from the base case of £19,000 per QALY gained was driven by proportional decreases in sevoflurane, resulting in ICER of £6494 per QALY gained and £31,567 per QALY gained. The remaining results reflect the sensitivity in other patient groups undergoing TIVA and mixed anaesthesia, but to a lesser extent. The decrease and increase in probability of awareness yielded ICERs of £14,881 per QALY gained and £24,521 per QALY gained respectively. Again, the ICER is sensitive to a decrease in the effectiveness of the E-Entropy module, which results in an ICER of

TABLE 88 One-way sensitivity analysis: E-Entropy compared with standard clinical monitoring in a general surgical population undergoing mixed anaesthesia

Parameter	Input value	Standard clinical monitoring		Entropy		Incremental		ICER (£/QALY gained)
		Cost (£)	QALYs	Cost (£)	QALYs	Cost (£)	QALYs	
Proportional change in sevoflurane	–0.492	18.14	–0.0007	19.84	–0.0004	1.70	0.0003	6494
	–0.079	18.14	–0.0007	26.42	–0.0004	8.28	0.0003	31,567
Probability awareness	0.001	17.92	–0.0006	23.02	–0.0004	5.11	0.0002	24,521
	0.0023	18.39	–0.0008	23.24	–0.0005	4.84	0.0003	14,881
Operating room awareness with depth of anaesthesia monitor	0.25	18.14	–0.0007	23.00	–0.0004	4.87	0.0003	15,454
	0.81	18.14	–0.0007	23.33	–0.0005	5.20	0.0002	30,967
Duration of LPS (years)	0.25	18.14	–0.0007	23.12	–0.0004	4.99	0.0003	19,290
	1	18.14	–0.0007	23.12	–0.0004	4.99	0.0003	18,446
Probability of LPS[a]	0.195	17.90	–0.0122	23.02	–0.0121	5.12	0.0002	26,362
	0.48	18.41	–0.0295	23.25	–0.0292	4.83	0.0003	15,833
Duration of PTSD (years)	5.6	18.14	–0.0006	23.12	–0.0004	4.99	0.0002	21,648
	9.6	18.14	–0.0007	23.12	–0.0004	4.99	0.0003	16,351
Proportion PTSD[b]	0.345	17.92	–0.0006	23.03	–0.0004	5.10	0.0002	23,609
	0.733	18.34	–0.0008	23.22	–0.0004	4.87	0.0003	15,856
LPS QoL decrement	–0.075	18.14	–0.0007	23.12	–0.0004	4.99	0.0003	19,216
	–0.05	18.14	–0.0007	23.12	–0.0004	4.99	0.0003	19,339
PTSD QoL decrement	–0.134	18.14	–0.0007	23.12	–0.0004	4.99	0.0003	17,912
	–0.068	18.14	–0.0006	23.12	–0.0004	4.99	0.0002	24,531
Probability people with PTSD seek treatment	0	17.55	–0.0007	22.86	–0.0004	5.31	0.0003	20,234
	1	19.28	–0.0007	23.64	–0.0004	4.36	0.0003	16,604
Unit cost of sensors (£)	6.51	18.14	–0.0007	20.95	–0.0004	2.82	0.0003	10,730
	10.85	18.14	–0.0007	25.29	–0.0004	7.16	0.0003	27,270

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of the results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

£30,967 per QALY gained. Changes in the probability of LPS, of PTSD, a reduction in the QoL decrement for PTSD and the changes in the unit costs of sensors appear to drive the results in this group of patients.

Scenario analysis

Inclusion of anaesthesia-related complication (postoperative nausea and vomiting) The systematic review of patient outcomes did not identify any robust data that reported an estimate of the effect of E-Entropy monitoring on risk of PONV. We developed a scenario analysis using data from a meta-analysis by Liu,[105] on the effectiveness of BIS on a range of outcomes including PONV, to investigate the potential impact of including this outcome on the cost-effectiveness results.

For this scenario analysis we assumed a baseline PONV risk of 30%,[102–104] for standard clinical monitoring and applied the OR derived in the meta-analysis (0.77, 95% CI 0.56 to 0.99) to estimate risk for E-Entropy monitored patients. We assumed that all treatments (such as prophylaxis against PONV) were the same for each treatment group, and that all patients experiencing PONV were treated using 4 mg ondansetron by intramuscular or slow i.v. injection (unit cost = £5.39, BNF[33]).

Tables 89 and *90* report the results of this scenario analysis for high-risk patients and general surgical patients, respectively, undergoing GA with TIVA.

TABLE 89 Scenario analysis: including an estimated effect of E-Entropy monitoring on the incidence of PONV in patients at high risk of awareness undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
OR = 0.77: baseline risk = 0.3, risk with E-Entropy monitoring = 0.248					
Standard clinical monitoring	26.38		–0.0011		
E-Entropy	35.60	9.51	–0.0005	0.0007	14,010
OR = 0.56:[a] baseline risk = 0.3, risk with E-Entropy monitoring = 0.194					
Standard clinical monitoring	26.38		–0.0011		
E-Entropy	35.60	9.22	–0.0005	0.0007	13,576

a Lower limit of 95% CI estimated by Liu.[105]

TABLE 90 Scenario analysis: including an estimated effect of E-Entropy monitoring on the incidence of PONV in a general surgical population undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
***Anaesthetic drug consumption based on Ellerkmann* et al.[62]**					
OR = 0.77: baseline risk = 0.3, risk with E-Entropy monitoring = 0.248					
Standard clinical monitoring	25.32		–0.0007		
E-Entropy	34.92	9.61	–0.0004	0.0003	30,250
OR = 0.56: baseline risk = 0.3, risk with E-Entropy monitoring = 0.194					
Standard clinical monitoring	25.32		–0.0007		
E-Entropy	34.63	9.31	–0.0004	0.0003	29,324
***Anaesthetic drug consumption based on Gruenewald* et al.[55]**					
OR = 0.77: baseline risk = 0.3, risk with E-Entropy monitoring = 0.248					
Standard clinical monitoring	31.50		–0.0007		
E-Entropy	41.20	9.70	–0.0004	0.0003	30,550
OR = 0.56: baseline risk = 0.3, risk with E-Entropy monitoring = 0.194					
Standard clinical monitoring	31.50		–0.0007		
E-Entropy	40.90	9.41	–0.0004	0.0003	29,624

a Lower limit of 95% CI estimated by Liu.[105]

The base-case ICER of £14,421 per QALY gained was insensitive to both changes in OR of PONV with E-Entropy monitoring. An OR of 0.77 applied to the baseline risk resulted in an ICER of £14,010 per QALY gained, whereas an OR of 0.56 resulted in an ICER of £13,576 per QALY gained.

Again, changes in the OR of PONV as a result of E-Entropy monitoring make little difference to the ICER in a general surgical population undergoing TIVA. The base-case ICER of £31,131 applying Ellerkmann and colleagues' anaesthetic consumption estimates, became £29,324 and £30,250 per QALY gained with ORs applied to the baseline risk of 0.56 and 0.77 respectively. Applying Gruenewald and colleagues' anaesthetic consumption estimates resulted in ICERs of £30,550 per QALY gained (OR 0.77) and £29,624 per QALY gained (OR 0.56).

Tables 91 and *92* report the results of this scenario analysis for patients at high risk and for patients at average risk of intraoperative awareness, respectively, undergoing GA with mixed anaesthesia (induction with i.v. anaesthetic and maintenance with i.v. and inhaled anaesthetic).

Where the OR for PONV was changed to 0.77 and 0.56 in a high-risk population receiving mixed anaesthesia, the ICER reduced slightly, but was generally insensitive to the changes, which resulted in ICER of £18,833 and £18,271 per QALY gained respectively.

TABLE 91 Scenario analysis: including an estimated effect of E-Entropy monitoring on incidence of PONV in patients at high risk of awareness undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
OR = 0.77: baseline risk = 0.3, risk with E-Entropy monitoring = 0.248					
Standard clinical monitoring	19.20		–0.0011		
E-Entropy	29.07	9.86	–0.0006	0.0005	18,833
OR = 0.56: baseline risk = 0.3, risk with E-Entropy monitoring = 0.194					
Standard clinical monitoring	19.20		–0.0011		
E-Entropy	28.77	9.57	–0.0006	0.0005	18,271

a Lower limit of 95% CI estimated by Liu.[105]

TABLE 92 Scenario analysis: including an estimated effect of E-Entropy monitoring on incidence of PONV in a general surgical population undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALYs	ICER (£/QALY gained)
OR = 0.77: baseline risk = 0.3, risk with E-Entropy monitoring = 0.248					
Standard clinical monitoring	18.14		–0.0007		
E-Entropy	22.84	4.71	–0.0004	0.0003	17,934
OR = 0.56: baseline risk = 0.3, risk with E-Entropy monitoring = 0.194					
Standard clinical monitoring	18.14		–0.0007		
E-Entropy	22.55	4.41	–0.0004	0.0003	16,813

a Lower limit of 95% CI estimated by Liu.[105]

The changes in OR for PONV to 0.77 and 0.56 again resulted in a slightly larger reduction in the ICER in this scenario (in a general surgical population receiving mixed anaesthesia), to £17,934 per QALY gained and £16,813 per QALY gained respectively.

Scenario analyses for probability of intraoperative awareness for patients at high risk of intraoperative awareness and for the general surgical population Our review of published studies of the incidence of intraoperative awareness identified substantial uncertainty over the estimated values. We used pooled values across identified studies in the base-case analysis. However, the value adopted for 'high risk' is lower than the 1% incidence cited in the publication reporting one of the included trials[44] (based on incidences reported by Phillips and colleagues,[138] Ranta and colleagues[112] and Myles and colleagues[79]), and the pooled estimate adopted for a general surgical population excluded two outlying studies (one high and one low extreme value).

For this scenario analysis we replace the base-case estimate for probability of awareness in high-risk population (0.45%) with the higher value of 1% (*Table 93*). The effect of this is to reduce the ICER to £6059 per QALY gained for TIVA and to £8882 for mixed anaesthesia.

For the general surgical population, we replaced the base-case estimate for probability of awareness (0.16%) with the extreme high and low values reported in the literature (0.99% and 0.007%, *Tables 94* and *95*).

The ICER was sensitive to changes in the probability of awareness, where the outlying values were adopted. In each case (where anaesthetic consumption estimates were applied from either Ellerkmann and colleagues[62] or Gruenewald and colleagues[55]), these range from approximately £5600 per QALY gained to approximately £80,000 per QALY gained respectively.

In threshold analyses we found that depth of anaesthesia monitoring with E-Entropy for patients undergoing GA with TIVA was cost-effective if the probability of awareness was $>$ 0.192–0.194%, at a willingness-to-pay threshold of £30,000 per QALY gained. Depth of anaesthesia monitoring with E-Entropy was cost-effective if the probability of awareness was $>$ 0.315–0.318%, at a willingness-to-pay threshold of £20,000 per QALY gained. We report a range of values for the probability of awareness, as the exact values depend on which study the anaesthetic drug consumption is based (Ellerkmann and colleagues[62] or Gruenewald and colleagues[55]).

The ICER is sensitive to a scenario where the outlying probabilities of awareness are applied in a general population undergoing mixed anaesthesia. Where the lower probability of 0.007 is applied, the ICER increases to £42,599 per QALY gained. Where the probability is set at 0.99%, the ICER decreases considerably to £3286.

In threshold analyses we found that depth of anaesthesia monitoring with E-Entropy for patients undergoing mixed GA was cost-effective if the probability of awareness was $>$ 0.098%, at a willingness-to-pay threshold of £30,000 per QALY gained. The required probability, at a willingness-to-pay threshold of £20,000 per QALY gained, is 0.196%.

Impact of assumptions on number of patients per device-year In order to apportion the capital cost of the depth of anaesthesia monitoring modules we required an estimate of the number of patients/cases in which the monitor module was used in each year (patients per device-year), throughout its assumed 5-year effective life. The estimate used for the general surgical population was 1000 patients per year (equivalent to four patients per day over 250 working days per year), based on discussion with clinical experts. This scenario analysis investigates the impact of this assumption on the estimated incremental cost associated with E-Entropy monitoring, compared with standard clinical monitoring, and the resulting effect on the ICER. *Table 96* reports the incremental cost and ICER for E-Entropy compared with standard clinical monitoring at four selected values for the number of patients per device-year: the base-case value of 500

TABLE 93 Scenario analysis: impact of baseline risk of awareness on cost-effectiveness of E-Entropy monitoring for patients at high risk of awareness

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
TIVA					
Standard clinical monitoring	28.41		–0.0021		
E-Entropy	36.67	8.26	–0.0007	0.0014	6059
Mixed anaesthesia					
Standard clinical monitoring	21.23		–0.0021		
E-Entropy	30.26	9.03	–0.0010	0.0010	8882

TABLE 94 Scenario analysis: impact of baseline risk of awareness on cost-effectiveness of E-Entropy monitoring in a general surgical population undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
***Anaesthetic drug consumption based on Ellerkmann* et al.[62]**					
Baseline probability of awareness = 0.99%					
Standard clinical monitoring	28.37		–0.0020		
E-Entropy	35.94	7.57	–0.0007	0.0014	5605
Baseline probability of awareness = 0.007%					
Standard clinical monitoring	24.75		–0.0004		
E-Entropy	35.07	10.32	–0.0003	0.0001	81,406
***Anaesthetic drug consumption based on Gruenewald* et al.[55]**					
Baseline probability of awareness = 0.99%					
Standard clinical monitoring	34.55		–0.0020		
E-Entropy	42.22	7.67	–0.0007	0.0014	5676
Baseline probability of awareness = 0.007%					
Standard clinical monitoring	30.93		–0.0004		
E-Entropy	41.34	10.41	–0.0003	0.0001	82,157

TABLE 95 Scenario analysis: impact of baseline risk of awareness on cost-effectiveness of E-Entropy monitoring in a general surgical population undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Baseline probability of awareness = 0.99%					
Standard clinical monitoring	21.19		–0.0020		
E-Entropy	24.51	3.31	–0.0010	0.0010	3286
Baseline probability of awareness = 0.007%					
Standard clinical monitoring	17.57		–0.0004		
E-Entropy	22.87	5.30	–0.0003	0.0001	42,599

TABLE 96 Scenario analysis: impact of number of patients per device-year on cost-effectiveness of E-Entropy monitoring in a general surgical population

Patients per device-year	Standard clinical monitoring (£)	E-Entropy (£)	Incremental cost (£)	ICER (£/QALY gained
TIVA				
100	25.32	45.87	20.56	64,720
500	25.32	36.39	11.07	34,863
1000	25.32	35.20	9.89	31,131
1500	25.32	34.81	9.49	29,887
Mixed anaesthesia				
100	18.14	33.79	15.65	59,657
500	18.14	24.31	6.17	23,517
1000	18.14	23.12	4.99	19,000
1500	18.14	22.73	4.59	17,494

and also for a low value of 10 and high values of 1000 (four patients per day over 250 working days per year) and 1500 (six patients per day over 250 working days per year). This suggests that the assumed number of patients per device-year only has a substantial impact on incremental cost (hence on the ICER) at very low volumes.

Impact of alternative assumptions on the utility decrement for PTSD The QoL decrement applied in the base case was based on Freed and colleagues'[120] paper on veterans with PTSD. In order to investigate the impact of a sparse evidence base on HRQoL in a group of patients with PTSD, a scenario analyses was undertaken. The utility decrement was adjusted to 0.50 and 0.75 in high-risk and general surgical groups receiving either TIVA or mixed anaesthesia (*Tables 97* and *98*).

The ICER was sensitive to these alternative scenarios in high-risk patients, both receiving TIVA and mixed anaesthesia. Where the PTSD decrement was increased to 0.5 in TIVA and mixed anaesthesia, the ICER reduced to £4152 per QALY gained and £5835 per QALY gained respectively. Where the PTSD decrement was increased further, the ICER decreased again to £2827 and £3997 per QALY gained in the TIVA and mixed anaesthesia groups respectively.

The scenario analyses using alternative PTSD decrements in the general population reflect the results in the high-risk population: there is a substantial reduction in the ICER where these are decreased.

Narcotrend compared with standard clinical monitoring

Base case

Total intravenous anaesthesia The costs, QALY and ICER modelled for patients considered at high risk of intraoperative awareness undergoing GA with TIVA, comparing standard clinical monitoring with depth of anaesthesia monitoring by Narcotrend are presented in *Table 99*.

Narcotrend monitoring was modelled as being associated with 10.8 cases of awareness, compared with 45 cases among patients receiving standard clinical monitoring, in a cohort of 10,000 patients. This results in a reduction of 11.1 cases of LPS (from 14.7 to 3.5), which includes a reduction of six cases of PTSD (from 8.0 to 1.9).

TABLE 97 Scenario analysis: impact of utility decrement for PTSD on cost-effectiveness of E-Entropy in patients at high risk of awareness undergoing TIVA or mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = 0.50					
Standard clinical monitoring	26.38		–0.0034		
E-Entropy	36.18	9.79	–0.0010	0.0024	4152
Utility decrement for PTSD = 0.75					
Standard clinical monitoring	26.38		–0.0048		
E-Entropy	36.18	9.79	–0.0014	0.0035	2827
Mixed anaesthesia					
Utility decrement for PTSD = 0.50					
Standard clinical monitoring	19.20		–0.0034		
E-Entropy	29.35	10.14	–0.0016	0.0017	5835
Utility decrement for PTSD = 0.75					
Standard clinical monitoring	19.20		–0.0048		
E-Entropy	29.35	10.14	–0.0023	0.0025	3997

TABLE 98 Scenario analysis: impact of utility decrement for PTSD on cost-effectiveness of E-Entropy in a general surgical population undergoing TIVA or mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = 0.50					
Standard clinical monitoring	25.32		–0.0015		
E-Entropy	35.20	9.98	–0.0005	0.0009	10,803
Utility decrement for PTSD = 0.75					
Standard clinical monitoring	25.32		–0.0020		
E-Entropy	35.20	9.89	–0.0007	0.00013	7556
Mixed anaesthesia					
Utility decrement for PTSD = 0.50					
Standard clinical monitoring	18.14		–0.0015		
E-Entropy	23.12	4.99	–0.0008	0.0007	7176
Utility decrement for PTSD = 0.75					
Standard clinical monitoring	18.14		–0.0020		
E-Entropy	23.12	4.99	–0.0010	0.0010	5091

The cost of standard clinical monitoring during anaesthesia in high-risk patients was lower than for Narcotrend depth of anaesthesia monitoring, with the incremental cost being £3.86. The increased cost for Narcotrend monitoring is largely the result of the additional costs of the depth monitor (80% of the per patient cost) rather than the sensors attached to the patients (20% of the per patient cost). There is no reduction in anaesthetic costs associated with depth of anaesthesia monitoring, for this group of patients, although some of the additional cost of depth of anaesthesia monitoring is offset by reduced costs associated with psychological sequelae of awareness (*Table 100*).

Patients in both groups incurred a slight QALY loss, resulting from psychological sequelae of awareness (LPS and PTSD) and from POCD in older patients. This was lower in the Narcotrend-monitored patients, with a difference of 0.0007 QALY, resulting in an ICER of £5681 per QALY gained.

The costs, QALY and ICER modelled for a general surgical population (not just those at high risk of intraoperative awareness) undergoing GA with TIVA, comparing standard clinical monitoring with depth of anaesthesia monitoring by Narcotrend are presented in *Table 101*.

In the general surgical population, Narcotrend monitoring was modelled as being associated with 3.8 cases of awareness, compared with 16 cases in patients receiving standard clinical monitoring, in cohorts of 10,000 patients. This results in a reduction of four cases of LPS (from 5.2 to 1.3), which includes a reduction of 2.1 cases of PTSD (from 2.8 to 0.7).

In this patient population, depth of anaesthesia monitoring with Narcotrend is associated with lower costs than for standard clinical monitoring (see *Table 101*). This results from reduction in the use of anaesthetic drugs (and to a lesser extent with lower PTSD-related costs, because of the lower incidence of awareness), which offset the additional costs associated with depth of anaesthesia monitoring (*Table 102*).

Given the lower probability of intraoperative awareness in this group of patients, the QALY losses for both standard clinical monitoring and Narcotrend monitoring, resulting from psychological sequelae of awareness (LPS and PTSD), are lower than for the high-risk group. The QALY loss arising from the LPS and PTSD following awareness and from POCD are lower for patients monitored with Narcotrend compared with those receiving standard clinical monitoring. As better outcomes are modelled as being achieved at lower costs, Narcotrend dominates standard clinical monitoring for this population.

TABLE 99 Cost-effectiveness of Narcotrend compared with standard clinical monitoring in a population at high risk of awareness undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	33.45		–0.0011		
Narcotrend	37.31	3.86	–0.0005	0.0007	5681

TABLE 100 Breakdown of total cost for standard clinical monitoring and Narcotrend for patients at high risk of awareness undergoing TIVA

Cost	Standard clinical monitoring (£)	Narcotrend (£)
Depth of anaesthesia monitoring	0.00	5.12
Anaesthetic drugs	30.18	30.18
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	1.66	0.40

Mixed anaesthesia [induction with intravenous anaesthetic (remifentanil) and maintenance with intravenous and inhaled anaesthetic (remifentanil and desflurane)] The costs, QALY and ICER modelled for patients considered at high risk of intraoperative awareness undergoing mixed anaesthesia, comparing standard clinical monitoring with depth of anaesthesia monitoring by Narcotrend are presented in *Table 103*.

Narcotrend monitoring is modelled as being associated with 20.3 cases of awareness, compared with 45 cases among patients receiving standard clinical monitoring, in cohorts of 10,000 patients. This results in a reduction of 8.1 cases of LPS (from 14.7 to 6.6), which includes a reduction of 4.4 cases of PTSD (from 8.0 to 3.6).

In a high-risk population receiving mixed anaesthesia, Narcotrend monitoring resulted in an incremental cost of £4.21. The increased costs in the Narcotrend group are associated with the depth of anaesthesia monitoring costs. Anaesthetic drug costs are the same in each group, but again the monitoring costs incurred by the Narcotrend group are, to an extent, offset by reduced costs associated with PTSD (see *Table 104*).

The reduced QALY loss in high-risk patients undergoing monitoring with Narcotrend compared with patients undergoing standard monitoring occurred as a result of the lower probability of awareness in this group, with a difference of 0.0005 QALY. This resulted in an ICER of £8033 per QALY gained.

TABLE 101 Cost-effectiveness of Narcotrend compared with standard clinical monitoring in a general surgical population undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	32.39		–0.0007		
Narcotrend	28.53	–3.85	–0.0004	0.0003	Narcotrend dominates

TABLE 102 Breakdown of total cost for standard clinical monitoring and Narcotrend in a general surgical population undergoing TIVA

Cost	Standard clinical monitoring (£)	Narcotrend (£)
Depth of anaesthesia monitoring	0.00	2.84
Anaesthetic drugs	30.18	23.94
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	0.59	0.14

TABLE 103 Cost-effectiveness of Narcotrend compared with standard clinical monitoring in a high-risk population undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALYs	ICER (£/QALY gained)
Standard clinical monitoring	38.99		–0.0011		
Narcotrend	43.20	4.21	–0.0006	0.0005	8033

The costs, QALY and ICER modelled for a general surgical population (not just those at high risk of intraoperative awareness) undergoing mixed GA, comparing standard clinical monitoring with depth of anaesthesia monitoring by Narcotrend are presented in *Table 105*.

Narcotrend monitoring was modelled as being associated with 7.2 cases of awareness, compared with 16 cases among patients receiving standard clinical monitoring, in cohorts of 10,000 patients. This results in a reduction of three cases of LPS (from 5.2 to 2.3), which includes a reduction of 1.5 cases of PTSD (from 2.8 to 1.3).

Narcotrend monitoring is associated with lower costs than for standard clinical monitoring in this patient population (*Table 106*). This arises from the relatively small additional cost of depth of anaesthesia monitoring with Narcotrend (the sensors are available at a low cost, whereas the capital cost of the monitor is spread across a relatively large patient throughput) and from savings because of a reduction in the use of anaesthetic drugs (and to a lesser extent with lower PTSD-related costs, because of the lower incidence of awareness).

As better outcomes are modelled as being achieved at lower costs, Narcotrend dominates standard clinical monitoring for this population.

TABLE 104 Breakdown of total cost for standard clinical monitoring and Narcotrend in patients at high risk of awareness undergoing mixed anaesthesia

Cost	Standard clinical monitoring (£)	Narcotrend (£)
Depth of anaesthesia monitoring	0.00	5.12
Anaesthetic drugs	35.72	35.72
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	1.66	0.75

TABLE 105 Cost-effectiveness of Narcotrend compared with standard clinical monitoring in a general surgical population undergoing mixed anaesthesia

Intervention	Cost	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Standard clinical monitoring	37.93		–0.0007		
Narcotrend	36.18	–1.74	–0.0004	0.0003	Narcotrend dominates

TABLE 106 Breakdown of total cost for standard clinical monitoring and Narcotrend for a general surgical population undergoing mixed anaesthesia

Cost	Standard clinical monitoring (£)	Narcotrend
Depth of anaesthesia monitoring	0.00	2.84
Anaesthetic drugs	35.72	31.46
PONV	1.62	1.62
POCD	0.00	0.00
PTSD	0.59	0.27

Deterministic sensitivity analysis

Total intravenous anaesthesia One-way sensitivity analyses of key parameters were undertaken in both the general surgical population and the high-risk surgical population undergoing general anaesthetic using TIVA. The results are shown in *Tables 107* and *108*.

The one-way sensitivity analysis of key parameters in the high-risk surgical group receiving TIVA resulted in ICER ranging from £1123 to £25,656 per QALY gained. However, the ICER appears robust to the majority of changes in parameters in this group. The ICER also increases where the probability of awareness, of LPS, and the PTSD decrements are reduced, and the relative risk of awareness increases.

The one-way sensitivity analysis of key parameters demonstrated that the ICER in the general surgical population is robust where these parameters are varied. In each case Narcotrend dominates standard

TABLE 107 One-way sensitivity analysis: Narcotrend compared with standard clinical monitoring in patients at high risk of awareness undergoing TIVA

Parameter	Input value	Standard clinical monitoring		Narcotrend		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost (£)	QALY	
Probability awareness	0.0006	32.02	–0.0005	36.97	–0.0003	4.95	0.0002	25,656
	0.0119	36.18	–0.0024	37.97	–0.0008	1.80	0.0016	1123
Operating room awareness with depth of anaesthesia monitor	0.1	33.45	–0.0011	37.08	–0.0004	3.63	0.0008	4631
	0.6	33.45	–0.0011	37.91	–0.0007	4.45	0.0004	10,792
Duration of LPS (years)	0.25	33.45	–0.0011	37.31	–0.0005	3.86	0.0007	5812
	1	33.45	–0.0012	37.31	–0.0005	3.86	0.0007	5436
Probability of LPS[a]	0.195	32.79	–0.0125	37.15	–0.0121	4.36	0.0004	10,552
	0.48	34.24	–0.0302	37.50	–0.0293	3.26	0.0008	3861
Duration of PTSD (years)	5.6	33.45	–0.0010	37.31	–0.0004	3.86	0.0006	6959
	9.6	33.45	–0.0014	37.31	–0.0005	3.86	0.0008	4570
Proportion PTSD[b]	0.345	32.85	–0.0009	37.17	–0.0004	4.32	0.0005	8640
	0.733	34.04	–0.0014	37.45	–0.0005	3.41	0.0009	4002
LPS QoL decrement	–0.075	33.45	–0.0011	37.31	–0.0005	3.86	0.0007	5779
	–0.05	33.45	–0.0011	37.31	–0.0005	3.86	0.0007	5835
PTSD QoL decrement	–0.134	33.45	–0.0012	37.31	–0.0005	3.86	0.0007	5207
	–0.068	33.45	–0.0008	37.31	–0.0004	3.86	0.0004	8589
Probability people with PTSD seek treatment	0	31.80	–0.0011	36.91	–0.0005	5.12	0.0007	7534
	1	36.67	–0.0011	38.09	–0.0005	1.42	0.0007	2085
Unit cost of sensors (£)	0.42	33.45	–0.0011	37.17	–0.0005	3.72	0.0007	5475
	0.70	33.45	–0.0011	37.45	–0.0005	4.00	0.0007	5887

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as a proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

TABLE 108 One-way sensitivity analysis: Narcotrend compared with standard clinical monitoring in a general surgical population undergoing TIVA

Parameter	Input value	Standard clinical monitoring		Narcotrend		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost (£)	QALY	
Proportional change in propofol use	–0.429	32.39	–0.0007	24.65	–0.0004	–7.73	0.0003	Narcotrend dominates
	–0.0155	32.39	–0.0007	31.19	–0.0004	–1.20	0.0003	
Proportional change in remifentanil	–0.158	32.39	–0.0007	27.41	–0.0004	–4.98	0.0003	
	0.050	32.39	–0.0007	29.65	–0.0004	–2.73	0.0003	
Probability awareness	0.001	32.17	–0.0006	28.48	–0.0003	–3.69	0.0002	
	0.0023	32.64	–0.0008	28.59	–0.0004	–4.05	0.0004	
Operating room awareness with depth of anaesthesia monitor	0.1	32.39	–0.0007	28.45	–0.0003	–3.94	0.0004	
	0.6	32.39	–0.0007	28.74	–0.0004	–3.64	0.0002	
Duration of LPS (years)	0.25	32.39	–0.0007	28.53	–0.0004	–3.85	0.0003	
	1	32.39	–0.0007	28.53	–0.0004	–3.85	0.0003	
Probability of LPS[a]	0.195	32.15	–0.0122	28.48	–0.0120	–3.67	0.0002	
	0.48	32.66	–0.0295	28.60	–0.0292	–4.07	0.0004	
Duration of PTSD (years)	5.6	32.39	–0.0006	28.53	–0.0003	–3.85	0.0003	
	9.6	32.39	–0.0007	28.53	–0.0004	–3.85	0.0004	
Proportion PTSD[b]	0.345	32.17	–0.0006	28.48	–0.0003	–3.69	0.0003	
	0.733	32.59	–0.0008	28.58	–0.0004	–4.01	0.0004	
LPS QoL decrement	–0.075	32.39	–0.0007	28.53	–0.0004	–3.85	0.0003	
	–0.05	32.39	–0.0007	28.53	–0.0004	–3.85	0.0003	
PTSD QoL decrement	–0.134	32.39	–0.0007	28.53	–0.0004	–3.85	0.0003	
	–0.068	32.39	–0.0006	28.53	–0.0003	–3.85	0.0002	
Probability people with PTSD seek treatment	0	31.80	–0.0007	28.39	–0.0004	–3.41	0.0003	
	1	33.53	–0.0007	28.81	–0.0004	–4.72	0.0003	
Unit cost of sensors (£)	0.42	32.39	–0.0007	28.39	–0.0004	–3.99	0.0003	
	0.70	32.39	–0.0007	28.67	–0.0004	–3.71	0.0003	

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as a proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

clinical monitoring in the general surgical population receiving TIVA, by generating improved outcome at reduced cost.

Mixed anaesthesia One-way sensitivity analyses of key parameters were undertaken in both the general surgical population and the high-risk surgical population undergoing general anaesthetic using mixed anaesthesia [induction with i.v. anaesthetic (remifentanil) and maintenance with i.v. and inhaled anaesthetic (remifentanil and desflurane)]. The results are shown in *Tables 109* and *110*.

TABLE 109 One-way sensitivity analysis: Narcotrend compared with standard clinical monitoring in patients at high risk of awareness undergoing mixed anaesthesia

Parameter	Input value	Standard clinical monitoring		Narcotrend		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost (£)	QALY	
Probability awareness	0.0006	37.56	–0.0005	42.55	–0.0003	4.99	0.0002	29,010
	0.0119	41.72	–0.0024	44.44	–0.0012	2.72	0.0012	2290
Operating room awareness with depth of anaesthesia monitor	0.25	38.99	–0.0011	42.87	–0.0005	3.87	0.0007	5769
	0.81	38.99	–0.0011	43.80	–0.0009	4.80	0.0003	18,621
Duration of LPS (years)	0.25	38.99	–0.0011	43.20	–0.0006	4.21	0.0005	8206
	1	38.99	–0.0012	43.20	–0.0006	4.21	0.0005	7707
Probability of LPS[a]	0.195	38.33	–0.0125	42.90	–0.0122	4.57	0.0003	13,785
	0.48	39.78	–0.0302	43.55	–0.0295	3.78	0.0006	5865
Duration of PTSD (years)	5.6	38.99	–0.0010	43.20	–0.0006	4.21	0.0004	9704
	9.6	38.99	–0.0014	43.20	–0.0007	4.21	0.0006	6542
Proportion PTSD[b]	0.345	38.39	–0.0009	42.93	–0.0005	4.54	0.0004	11,522
	0.733	39.58	–0.0014	43.46	–0.0007	3.89	0.0006	5982
LPS QoL decrement	–0.075	38.99	–0.0011	43.20	–0.0006	4.21	0.0005	8162
	–0.05	38.99	–0.0011	43.20	–0.0006	4.21	0.0005	8236
PTSD QoL decrement	–0.134	38.99	–0.0012	43.20	–0.0007	4.21	0.0006	7401
	–0.068	38.99	–0.0008	43.20	–0.0005	4.21	0.0004	11,768
Probability people with PTSD seek treatment	0	37.34	–0.0011	42.45	–0.0006	5.12	0.0005	9770
	1	42.21	–0.0011	44.65	–0.0006	2.44	0.0005	4661
Unit cost of sensors (£)	0.42	38.99	–0.0011	43.06	–0.0006	4.07	0.0005	7766
	0.70	38.99	–0.0011	43.34	–0.0006	4.35	0.0005	8300

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as a proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

The results of the one way-sensitivity analysis in high-risk patients undergoing mixed anaesthesia range from £2290 to £29,010 per QALY gained. The ICER appears least sensitive to changes in the LPS decrement and most affected by the changes in probability of awareness to 0.0119 and 0.006, resulting in the lowest and highest ICERs of £2290 and £29,010 per QALY gained respectively. The results are also sensitive to the estimated effect of monitoring on the incidence of awareness, the proportion of patients with LPS who develop PTSD and to the size of utility decrement for PTSD.

The one-way sensitivity analysis suggests that the results in the general surgical population are generally robust to variation in key input parameters. The exception is the proportional change in use of desflurane. The upper limit of the 95% CI is close to zero, indicating only limited savings in cost of anaesthetic gas to offset against the cost of Narcotrend monitoring, resulting in a positive incremental cost.

TABLE 110 One-way sensitivity analysis: Narcotrend compared with standard clinical monitoring in a general surgical population undergoing mixed anaesthesia

Parameter	Input value	Standard clinical monitoring		Narcotrend		Incremental		ICER (£/QALY gained)
		Cost (£)	QALY	Cost (£)	QALY	Cost	QALY	
Proportional change in desflurane	–0.256	37.93	–0.0007	33.77	–0.0004	–4.15	0.0003	Narcotrend dominates
	–0.056	37.93	–0.0007	38.59	–0.0004	0.66	0.0003	2534
Proportional change in remifentanil	–0.168	37.93	–0.0007	34.73	–0.0004	–3.20	0.0003	Narcotrend dominates
	0.081	37.93	–0.0007	37.62	–0.0004	–0.30	0.0003	
Probability awareness	0.001	37.71	–0.0006	36.08	–0.0004	–1.62	0.0002	
	0.0023	38.18	–0.0008	36.30	–0.0005	–1.89	0.0003	
Operating room awareness with depth of anaesthesia monitor	0.25	37.93	–0.0007	36.06	–0.0004	–1.86	0.0003	
	0.81	37.93	–0.0007	36.39	–0.0005	–1.53	0.0002	
Duration of LPS (years)	0.25	37.93	–0.0007	36.18	–0.0004	–1.74	0.0003	
	1	37.93	–0.0007	36.18	–0.0004	–1.74	0.0003	
Probability of LPS[a]	0.195	37.69	–0.0122	36.08	–0.0121	–1.61	0.0002	
	0.48	38.20	–0.0295	36.31	–0.0292	–1.90	0.0003	
Duration of PTSD (years)	5.6	37.93	–0.0006	36.18	–0.0004	–1.74	0.0002	
	9.6	37.93	–0.0007	36.18	–0.0004	–1.74	0.0003	
Proportion PTSD[b]	0.345	37.71	–0.0006	36.09	–0.0004	–1.63	0.0002	
	0.733	38.13	–0.0008	36.28	–0.0004	–1.86	0.0003	
LPS QoL decrement	–0.075	37.93	–0.0007	36.18	–0.0004	–1.74	0.0003	
	–0.05	37.93	–0.0007	36.18	–0.0004	–1.74	0.0003	
PTSD QoL decrement	–0.134	37.93	–0.0007	36.18	–0.0004	–1.74	0.0003	
	–0.068	37.93	–0.0006	36.18	–0.0004	–1.74	0.0002	
Probability people with PTSD seek treatment	0	37.34	–0.0007	35.92	–0.0004	–1.42	0.0003	
	1	39.07	–0.0007	36.70	–0.0004	–2.37	0.0003	
Unit cost of sensors (£)	0	37.93	–0.0007	36.04	–0.0004	–1.88	0.0003	
	1	37.93	–0.0007	36.32	–0.0004	–1.60	0.0003	

a Changing the probability of LPS also changes the probability of PTSD, as PTSD is calculated as a proportion of the probability of LPS (i.e. people with PTSD are a subgroup of people with LPS). Sensitivity of results to changes in this parameter is therefore a combination of the effect of LPS (overall) and PTSD combined.

b Varying the proportion with PTSD within the population of LPS.

Scenario analysis

Inclusion of anaesthesia-related complication (postoperative nausea and vomiting) The systematic review of patient outcomes did not identify any robust data that reported an estimate of the effect of Narcotrend monitoring on risk of PONV. We developed a scenario analysis using data from a meta-analysis by Liu,[105] on the effectiveness of BIS on a range of outcomes including PONV, to investigate the potential impact of including this outcome on the cost-effectiveness results.

For this scenario analysis we assumed a baseline PONV risk of 30%,[102–104] for standard clinical monitoring and applied the OR derived in the meta-analysis (0.77, 95% CI 0.56 to 0.99) to estimate risk for Narcotrend-monitored patients. We assumed that all treatments (such as prophylaxis against PONV) were the same for each treatment group, and that all patients experiencing PONV were treated using 4 mg ondansetron by intramuscular or slow i.v. injection (unit cost = £5.39; BNF[33]).

Tables 111 and *112* report the results of this scenario analysis for high-risk patients and general surgical patients, respectively, undergoing GA with TIVA.

Variation in the OR of PONV applied in the model does not have an impact on the ICER, either in the case of the high-risk population (*Table 111*) or in the general surgical population (*Table 112*) undergoing TIVA.

Tables 113 and *114* report the results of this scenario analysis for patients at high risk and for patients at average risk of intraoperative awareness, respectively, undergoing GA with mixed anaesthesia (induction with i.v. anaesthetic and maintenance with i.v. and inhaled anaesthetic).

Where the variations in the OR of PONV are applied to the high-risk patients undergoing mixed anaesthesia there is a slight reduction in the ICER. An OR of 0.77 results in an ICER of £7499 per QALY gained and an OR of 0.56 yields an ICER of £6937 per QALY gained in this group.

In the case of the general risk group receiving mixed anaesthesia, the ICER is robust to the variation in risk of PONV, and Narcotrend continues to dominate.

TABLE 111 Scenario analysis: including an estimated effect of Narcotrend monitoring on the incidence of PONV in patients at high risk of awareness undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
OR = 0.77: baseline risk = 0.3, risk with Narcotrend monitoring = 0.248					
Standard clinical care	33.45		–0.0011		
Narcotrend	37.03	3.58	–0.0005	0.0007	5270
OR = 0.56: baseline risk = 0.3, risk with Narcotrend monitoring = 0.194					
Standard clinical care	33.45		–0.0011		
Narcotrend	36.74	3.28	–0.0005	0.0007	4836

TABLE 112 Scenario analysis: including an estimated effect of Narcotrend monitoring on the incidence of PONV in a general surgical population undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
OR = 0.77: baseline risk = 0.3, risk with Narcotrend monitoring = 0.248					
Standard clinical care	32.39		–0.0007		
Narcotrend	28.25	–4.13	–0.0004	0.0003	Narcotrend dominates
OR = 0.56: baseline risk = 0.3, risk with Narcotrend monitoring = 0.194					
Standard clinical care	32.39		–0.0007		
Narcotrend	27.96	–4.13	–0.0004	0.0003	Narcotrend dominates

TABLE 113 Scenario analysis: including an estimated effect of Narcotrend on the incidence of PONV in patients at high risk of awareness undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
OR = 0.77: baseline risk = 0.3, risk with Narcotrend monitoring = 0.248					
Standard clinical care	38.99		–0.0011		
Narcotrend	42.92	3.93	–0.0006	0.0005	7499
OR = 0.56: baseline risk = 0.3, risk with Narcotrend monitoring = 0.194					
Standard clinical care	38.99		–0.0011		
Narcotrend	42.63	3.63	–0.0006	0.0005	6937

TABLE 114 Scenario analysis: including an estimated effect of Narcotrend on the incidence of PONV in a general surgical population undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
OR = 0.77: baseline risk = 0.3, risk with Narcotrend monitoring = 0.248					
Standard clinical care	37.93	–2.02	–0.0007	0.0003	Narcotrend dominates
Narcotrend	35.90		–0.0004		
OR = 0.56: baseline risk = 0.3, risk with Narcotrend monitoring = 0.194					
Standard clinical care	37.93	–2.32	–0.0007	0.0003	Narcotrend dominates
Narcotrend	35.61		–0.0004		

Scenario analyses for probability of intraoperative awareness for patients at high risk of intraoperative awareness and for the general surgical population

Our review of published studies of the incidence of intraoperative awareness identified substantial uncertainty over the estimated values. We used pooled values across identified studies in the base-case analysis. However, the value adopted for 'high risk' is lower than the 1% incidence cited in the publication reporting one of the included trials[44] (based on incidences reported by Phillips and colleagues,[138] Ranta and colleagues[112] and Myles and colleagues[79]), and the pooled estimate adopted for a general surgical population excluded two outlying studies (one high and one low extreme value).

For this scenario analysis we replace the base-case estimate for probability of awareness in high-risk population (0.45%) with the higher value of 1% (*Table 115*).

The ICERs decrease substantially in the high-risk population receiving either TIVA or mixed anaesthesia, where the probability of awareness is set to 1%, from £8033 to £3047 per QALY gained in the group receiving mixed, and from £5681 to £1705 in the group receiving TIVA.

In the general surgical population, we replace the base-case estimate for probability of awareness (0.16%) with the extreme high and low values reported in the literature (0.99% and 0.007%, *Tables 116* and *117*).

Where the outlying probabilities are applied the ICER is robust and Narcotrend continues to dominate in TIVA and mixed anaesthesia patients.

Impact of assumptions on number of patients per device-year

In order to apportion the capital cost of the depth of anaesthesia monitoring modules, we required an estimate of the number of patients/cases in

TABLE 115 Scenario analysis: impact of baseline risk of awareness on cost-effectiveness of Narcotrend monitoring for patients at high risk of awareness

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
TIVA					
Standard clinical care	35.48		–0.0021		
Narcotrend	37.80	2.32	–0.0007	0.0014	1705
Mixed anaesthesia					
Standard clinical care	41.02		–0.0021		
Narcotrend	44.12	3.10	–0.0010	0.0010	3047

TABLE 116 Scenario analysis: impact of baseline risk of awareness on cost-effectiveness of Narcotrend monitoring for a general surgical population, undergoing TIVA

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Baseline probability of awareness = 0.99%					
Standard clinical care	35.44		–0.0020		Narcotrend dominates
Narcotrend	29.27	–6.17	–0.0007	–0.0014	
Baseline probability of awareness = 0.007%					
Standard clinical care	31.82		–0.0004		Narcotrend dominates
Narcotrend	28.40	–3.43	–0.0003	0.0001	

TABLE 117 Scenario analysis: impact of baseline risk of awareness on cost-effectiveness of Narcotrend monitoring for a general surgical population, undergoing mixed anaesthesia

Intervention	Cost per patient (£)	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
Baseline probability of awareness = 0.99%					
Standard clinical care	40.98		–0.0020		
Narcotrend	37.57	–3.42	–0.0010	0.0010	Narcotrend dominates
Baseline probability of awareness = 0.007%					
Standard clinical care	37.36		–0.0004		
Narcotrend	35.93	–1.43	–0.0003	0.0001	Narcotrend dominates

which the monitor module was used in each year (patients per device-year), throughout its assumed 5-year effective life. The estimate used for the general surgical population was 1000 patients per year (equivalent to four patients per day over 250 working days per year), which was based on discussion with clinical experts. This scenario analysis investigates the impact of this assumption on the estimated incremental cost associated with Narcotrend monitoring, compared with standard clinical monitoring, and the resulting effect on the ICER. *Table 118* reports the incremental cost and ICER for Narcotrend compared with standard clinical monitoring, at four selected values for the number of patients per device-year: the base-case value of 1000 and also for a low value of 10, intermediate value of 500 and a high value of 1500 (six

patients per day over 250 working days per year). This suggests that the assumed number of patients per device-year only has a substantial impact on incremental cost (hence on the ICER) at very low throughput.

Impact of alternative assumptions on the utility decrement for post-traumatic stress disorder The QoL decrement applied in the base case was based on Freed and colleagues'[120] paper on veterans with PTSD. In order to investigate the impact of a sparse evidence base on HRQoL in a group of patients with PTSD, a scenario analysis was undertaken. The utility decrement was adjusted to 0.50 and 0.75 in high-risk and general surgical groups receiving either TIVA (*Table 119*) or mixed anaesthesia (*Table 120*).

The ICER is substantially reduced in the high-risk surgical population where higher decrements for PTSD QoL are applied (see *Table 119*). These are reduced to £1636 and £1114 per QALY gained for a 0.5 and 0.75 decrement, respectively, in the group undergoing TIVA. The ICER is reduced to £2420 and £1658 for a 0.5 and 0.75 decrement in the group undergoing mixed anaesthesia.

Where the alternative values for PTSD decrement are applied for the general surgical population in both the TIVA and mixed anaesthesia groups, Narcotrend continues to dominate (see *Table 120*).

Cost-effectiveness summary

We have presented modelled cost-effectiveness analyses for BIS, E-Entropy and Narcotrend compared with standard clinical monitoring, for two modes of anaesthetic administration. There is substantial uncertainty associated with the analysis, given the weakness of the evidence base for the majority of outcomes included in the model. No robust evidence was identified on the effectiveness of E-Entropy or Narcotrend in avoiding intraoperative awareness or POCD and, in the absence of such evidence, we have assumed that the effect estimates derived for BIS can be applied. However, even in the case of BIS the evidence base is currently severely lacking. There is also limited evidence on the baseline incidence of anaesthetic complications included in the model. There is more evidence on the benefit in terms of reduced anaesthetic drug consumption, although for some technologies the evidence is inconclusive.

Overall the economic evaluation indicates that, for general surgical patients, some of the additional costs of depth of anaesthesia monitoring may be offset by reduction in consumption of anaesthetic drugs. However, the size of these savings may not fully offset the additional cost. Given the comparative rarity of awareness, cost-savings through the avoidance of PTSD are unlikely to offset the additional costs. However, avoidance of the psychological sequelae of awareness yields gains in outcome that, depending

TABLE 118 Scenario analysis: impact of number of patients per device-year on cost-effectiveness of Narcotrend monitoring in general surgical patients

Patients per device-year	Standard clinical monitoring (£)	Narcotrend (£)	Incremental cost (£)	ICER (£/QALY gained
TIVA				
100	32.39	49.03	16.65	52,414
500	32.39	30.81	–1.58	Narcotrend dominates
1000	32.39	28.53	–3.85	Narcotrend dominates
1500	32.39	27.7	–4.61	Narcotrend dominates
Mixed anaesthesia				
100	37.93	26.68	18.76	71,484
500	37.93	38.46	0.53	2035
1000	37.93	36.18	–1.74	Narcotrend dominates
1500	37.93	35.42	–2.50	Narcotrend dominates

TABLE 119 Scenario analysis: impact of utility decrement for PTSD on cost-effectiveness of Narcotrend in patients at high risk of awareness undergoing TIVA or mixed anaesthesia

Intervention	Cost	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = 0.50					
Standard clinical care	33.45		–0.0034		
Narcotrend	37.31	3.86	–0.0010	0.0024	1636
Utility decrement for PTSD = 0.75					
Standard clinical care	33.45		–0.0048		
Narcotrend	37.31	3.86	–0.0014	0.0035	1114
Mixed anaesthesia					
Utility decrement for PTSD = 0.50					
Standard clinical care	38.99		–0.0034		
Narcotrend	43.20	4.21	–0.0016	0.0017	2420
Utility decrement for PTSD = 0.75					
Standard clinical care	38.99		–0.0048		
Narcotrend	43.20	4.21	–0.0023	0.0025	1658

TABLE 120 Scenario analysis: impact of utility decrement for PTSD on cost-effectiveness of Narcotrend in general surgical population undergoing TIVA or mixed anaesthesia

Intervention	Cost	Incremental cost (£)	QALY	Incremental QALY	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = 0.50					
Standard clinical care	32.39		–0.0015		Narcotrend dominates
Narcotrend	28.53	–3.85	–0.0005	0.0009	
Utility decrement for PTSD = 0.75					
Standard clinical care	32.39		–0.0020		Narcotrend dominates
Narcotrend	28.53	–3.85	–0.0007	0.0013	
Mixed anaesthesia					
Utility decrement for PTSD = 0.50					
Standard clinical care	37.93		–0.0015		Narcotrend dominates
Narcotrend	36.18	–1.74	–0.0008	0.0007	
Utility decrement for PTSD = 0.75					
Standard clinical care	37.93		–0.0020		Narcotrend dominates
Narcotrend	36.18	–1.74	–0.0010	0.0010	

on the utility losses associated with these conditions, may be acceptable in cost-effectiveness terms. The economic analysis suggests that, other than at comparatively low patient volumes, the acquisition cost of the DoA modules may be less significant in determining cost-effectiveness than the cost of consumables – in particular the sensors attached to the patient. Other key determinants of the cost-effectiveness of depth of anaesthesia monitoring appear to be the baseline risk of awareness and, unsurprisingly, the effect size in terms of avoiding awareness.

Chapter 4 Assessment of factors relevant to the NHS and other parties

Few of the trials included in this report reported whether or not anaesthetists had received training in use of the EEG devices. In their evidence submissions to NICE, the manufacturers of the three EEG devices assessed suggested varying lengths of training necessary, from 30 minutes' instruction in placing of the sensors for the E-Entropy module, to a whole day of lecture and training for Narcotrend in the operating theatre. The manufacturer of BIS suggests no additional training is necessary, but that a modest amount of additional training further enhances safe and effective use. Expert clinical opinion suggests that it is relatively straightforward to learn how to attach sensors and interpret the device values, but also that some training may be of benefit. In terms of cost implications, training would be provided for free by the manufacturer in the operating theatre, and/or anaesthetists would be able to access education materials including online multimedia courses. The main cost would therefore be for the operating theatre and the anaesthetist's time. Once a device has been installed and any initial training given, anaesthetists would need a period of time to become accustomed to using the device in practice.

The long-term impact of intraoperative awareness can have a profound impact on the health and well-being of patients. Psychological symptoms[7] such as disturbed sleep, phobias, depression, anxiety and PTSD may limit daily activities including their ability to work, resulting in periods of sickness absence and with consequent financial implications for employers. In extreme cases patients may have to cease working altogether and therefore their financial livelihood will be significantly impaired, and they may become reliant on welfare services. There may also be knock-on effects on patients' families and friends, for example, to provide social, emotional and practical support. Strain may be placed on marriages and partnerships, leading to separation in more severe cases. Patients may seek treatment for their symptoms which will involve primary and community care services (e.g. to provide counselling and/or medication) and in some cases secondary care (e.g. psychiatric supervision).

Chapter 5 Discussion

Statement of principal findings

Systematic review of patient outcomes

The eligible evidence base for BIS-guided anaesthesia (11 RCTs, plus 31 RCTs included in the Cochrane BIS review[34]) is larger than that for entropy-guided or Narcotrend-guided anaesthesia (seven and four RCTs respectively). A notable feature of the primary studies within each of the BIS, E-Entropy and Narcotrend technologies is that very few RCTs were methodologically similar to one another, which in most cases precluded the pooling of outcomes across studies.

Explicit intraoperative awareness

The effect estimate for intraoperative awareness in the Cochrane BIS review[34] was updated using data from two recent large RCTs. One of these was the BAG-RECALL RCT by Avidan and colleagues,[44] which compared BIS monitoring with monitoring of end-tidal anaesthetic agent concentration. The trial, which took place across three centres in the USA and Canada, randomised at least 3020 patients per study group, and patients received only inhaled GA. The RCT by Zhang and colleagues[40] also recruited large numbers of patients (around 5000) but was conducted in China across 13 centres, and patients received TIVA, rather than inhaled anaesthesia. BIS-guided TIVA was compared against routine TIVA (no further details given). Both trials were statistically powered to detect explicit intraoperative awareness in patients considered to be at higher risk. The trials reported contrasting findings, with Avidan and colleagues[44] noting a higher but non-statistically significant incidence of definite awareness in BIS-monitored patients, and Zhang and colleagues[40] finding a statistically significantly lower incidence of confirmed awareness in patients monitored with BIS.

When both of these trials were added to the Cochrane meta-analysis the pooled Peto's OR remained statistically significant 0.45 (95% CI 0.25 to 0.81), favouring BIS, though with significant heterogeneity. We classified the trials into subgroups based on the type of GA used (inhaled only; mixed inhaled and i.v.; total i.v.). The pooled Peto's OR for the subgroups of mixed inhaled and i.v. GA, and TIVA were both consistent with the overall pooled OR (i.e. statistically significant in favour of BIS). In contrast, the pooled estimate for the trials of inhaled GA, including the BAG-RECALL RCT[44] and another large RCT (the B-Unaware trial[27]), favoured standard clinical monitoring although the confidence intervals overlapped with 1 indicating potential advantage to both BIS and to standard clinical monitoring. Importantly the BAG-RECALL RCT[44] was designed to overcome some of the methodological limitations of the B-Unaware trial,[27] such as use of a larger sample of patients, more than one centre, and use of only major risk factors for awareness. It is not fully clear why the results of this trial were contrary to expectation. Notably these trials compared a structured BIS protocol with a structured ETAC protocol, comprising target anaesthetic concentration-linked audible alarms, staff education and checklists. The comparators reported in the other trials in the meta-analysis did not report use of structured protocols.

The remaining trials that reported intraoperative awareness either assessed this as a main outcome (one RCT on BIS[49]) or as a secondary outcome (three RCTs on BIS,[48,51,62] six of the seven RCTs on entropy,[54,55,57,58,61,62] and all four of the RCTs on Narcotrend[59,60,63,64]). Although the RCT by Kerssens and colleagues[49] specified that intraoperative awareness was the main outcome, the authors reported that the study was not powered statistically for this outcome. None of the remaining studies reported whether or not it was powered statistically for detecting a clinically meaningful difference in intraoperative awareness. In these RCTs the sample sizes ranged from 10 to 160 patients per study group, which most likely would be insufficient for detecting clinically meaningful differences in intraoperative awareness, given the low incidence of this event (see *Table 1* in *Incidence of intraoperative awareness*). Only two of these RCTs reported cases of intraoperative awareness, both in adult populations, but did not test differences between

the study groups statistically. Kerssens and colleagues[49] reported that incidence rates in BIS-guided and standard clinical monitoring groups were 2.9% (2/67) and 1.6% (1/61) respectively. Gruenewald and colleagues[55] reported that incidence rates of intraoperative awareness in entropy-guided and standard clinical practice groups were 0% (0/37) and 2.9% (1/35) respectively. These incidence rates are relatively high compared with those estimated from much larger studies (see *Table 1*), although in the Gruenewald study awareness was experienced by only one patient.[55]

The case of awareness reported by Gruenewald and colleagues[55] might have happened outside of the period of GA, as patients were asked if they had any memory or awareness during different stages of their procedure, including in the ward, induction room, during surgery or extubation, or in the recovery room.[55] The reason for the relatively high incidence of awareness observed in the Kerssens study[49] is not clear. Although Kerssens and colleagues[49] did not specify that their patients were at risk of awareness, the patients did appear to be relatively old (early 60s), possibly overweight or obese, and half of them had notable illness (ASA physical score grade III). The awareness assessment conducted by Kerssens and colleagues[49] involved asking patients five questions that were very similar to those of the Brice interview. Both of these RCTs[49,55] stated that their outcome assessors were blinded to the study group. Assessment of awareness in these RCTs took place 6 hours[49] or 24 hours[55] after surgery, without any longer-term follow-up. In fact, only the large trial by Avidan and colleagues[44] conducted follow-up assessments longer than 3 days after surgery (30 days after extubation); all other trials that assessed intraoperative awareness conducted follow-up assessments only 1 day or less post surgery,[54,57,59,60] 3 days post surgery[48,58,61–64] or did not state when follow-up occurred.[51] As occurrences of intraoperative awareness may take time to develop (see *Incidence of intraoperative awareness*), these follow-up periods may have been too short for detecting all cases of awareness.

Weighing up the strengths and limitations of the studies, an appropriate conclusion would be that, in patients considered to be at increased risk of awareness, BIS monitoring is associated with a reduced likelihood of explicit intraoperative awareness. However, this may not be applicable where inhaled GA is solely used. There is no evidence that EEG device-titrated anaesthesia significantly affects incidence of explicit intraoperative awareness in surgical patients not considered to be at increased risk, primarily as trials large enough to detect awareness have not been conducted.

Implicit intraoperative awareness

Implicit awareness (i.e. awareness that the patient does not necessarily recall experiencing) was reported only in one BIS trial, as a secondary outcome.[49] The assessment involved presenting patients audibly with words during anaesthesia then conducting specialist word recall tests after recovery from anaesthesia. The results showed that only patients in the BIS group selected target words more often than distractor words, and that patients in the BIS group selected target words more often than in the standard clinical monitoring group. Although appearing to indicate implicit intraoperative awareness, these findings would only have clinical relevance if the patients were followed up and found to have related clinical sequelae. Such follow-up has not been done and, in general, the possible longer-term implications for patients of implicit intraoperative awareness are not well understood.

Sequelae and long-term consequences of intraoperative awareness

None of the trials reported longer-term detrimental impacts of awareness such as PTSD. The BAG RECALL trial by Avidan and colleagues[44] reported patient distress and sequelae associated with awareness as a post hoc secondary outcome, based on the Michigan Awareness Classification Instrument, in which distress related to intraoperative awareness includes reports of fear, anxiety, suffocation, a sense of doom or a sense of impending death. Avidan and colleagues[44] found a higher percentage of distress in the BIS-monitored group (0.28% compared with 0.04%), but no statistically significant difference between the groups. No other trials included in the systematic review assessed patients' distress, anxiety or depression.

Anaesthetic consumption

The RCTs that reported anaesthesia consumption as an outcome can be summarised in various ways, as they differed in their populations (adults or children) anaesthesia (volatile or i.v.), sample sizes, and the methods used to measure anaesthetic consumption. The specific details of the outcomes summarised in the table can be obtained from *Table 9* (BIS), *Table 17* (E-Entropy) and *Table 24* (Narcotrend) in *Results of systematic review of patient outcomes* of this report.

Anaesthetic consumption was a statistically powered outcome in four RCTs: for sevoflurane in adults,[61] propofol in adults,[62] sevoflurane in children[54] and propofol in children.[46] The outcomes were powered to detect either a 20% reduction in anaesthetic consumption[46,54,62] or a 50% reduction.[61]A further RCT on adults specified sevoflurane as the main outcome but the outcome was not powered statistically.[58] The statistically powered RCT reported significant reductions of sevoflurane consumption under entropy-guided anaesthesia relative to standard clinical monitoring (i.e. favouring the E-Entropy group) in both adults[61] and children,[54] but no difference in propofol consumption between BIS, E-Entropy and standard clinical monitoring groups in adults.[62] However, the last trial[62] has high risk of bias because of an imbalance in the patient attrition between the study groups (see *Quantity and quality of research available*). The one trial that was powered to detect clinically relevant differences in propofol consumption in children[46] did not report a statistical comparison between the study groups, but in this trial, by Bhardwaj and colleagues,[46] the propofol consumption rate was higher in the BIS-guided than the standard clinical monitoring group (see *Table 9*). Overall, the findings from the statistically powered RCT indicate that E-Entropy-guided and BIS-guided anaesthesia reduce the consumption of sevoflurane but not propofol in both adults and children, although it should be noted that the methods used to assess anaesthesia consumption differed between the studies. None of the trials of Narcotrend were statistically powered to detect differences in anaesthetic consumption.

The remaining trials were not specifically powered to detect differences in anaesthetic consumption but their findings for sevoflurane consumption are similar to those of the powered trials. Three RCTs that assessed sevoflurane consumption in adults found that consumption was significantly lower in the BIS-guided group[45,49] or E-Entropy-guided group[58] than under standard clinical monitoring. Two RCTs that assessed sevoflurane in children also found consumption to be lower in the BIS group[51] or E-Entropy group.[54] In contrast with the statistically powered trials, most of the trials that assessed consumption of propofol as a secondary outcome, which were all on adult populations, reported significant differences in consumption in favour of the EEG-guided anaesthesia group. These differences were reported for E-Entropy-guided anaesthesia[55,139] and Narcotrend-guided anaesthesia,[59,60,63] whereas one RCT on BIS-guided anaesthesia reported a reduced propofol consumption in the BIS group but without an indication of statistical significance.[47]

Two RCTs assessed the consumption of other anaesthetics as secondary outcomes. These were desflurane consumption in adults[64] and isoflurane consumption in children.[56] These trials found that EEG-guided anaesthesia significantly reduced consumption, either using Narcotrend monitoring in adults[64] or E-Entropy monitoring in children.[56]

It was possible to update effect estimates for anaesthetic consumption in the Cochrane review[34] for volatile anaesthesia (sevoflurane) using data from a RCT by Kerssens and colleagues,[49] and for TIVA (propofol) using data from a RCT by Ellerkmann and colleagues.[62] For both types of anaesthesia, the updated effect estimate (mean difference) remained statistically significantly different from zero and in favour of the BIS group. However, heterogeneity was statistically significant even when using a random-effects model.

Time to recovery from anaesthesia

Recovery from anaesthesia was assessed in several different ways. The most frequent measurements reported were time to eye opening (11 RCTs) and time to extubation (11 RCTs).

Other recovery outcomes that were assessed included time to arrival in the PACU (five RCTs); duration of stay in the PACU (two RCTs); time to discharge from the PACU (five RCTs); time to response to commands (three RCTs) time to recovery of orientation (three RCTs); time to first movement response (two RCTs); time to recovery based on recovery scores (two RCTs); time to spontaneous breathing (one RCT); time to laryngeal mask airway removal (one RCT); and time to phonation (one RCT). 'PACU stay' was an outcome in the Cochrane review[34] but does not appear to distinguish between PACU admissions, stay and discharge times. For this reason the Cochrane review meta-analysis was not updated with data from the RCTs identified in the current review.

Time to eye opening

Four of the 11 RCTs that assessed this outcome were powered statistically to detect a difference between the study groups of 1.5 minutes,[64] 3 minutes[55,63] or 5 minutes.[56] Two of these powered trials detected a statistically significant difference in time to eye opening[56,63] and two did not.[55,64] Among the remaining seven RCTs[46,48,51,54,57,61,62] that were not specifically powered for this outcome, one[57] detected a significant difference between the study groups in time to eye opening and six[46,48,51,54,61,62] did not. In the three RCTs[56,57,63] that reported significant effects, the time to eye opening was consistently shorter in the EEG group than the standard clinical monitoring group. The significant reductions did not show any clear pattern with regard to whether the population (adults/children), EEG device used (BIS, E-Entropy, Narcotrend) or type of anaesthesia (volatile, total i.v. or mixed) could be explanatory variables. It is unclear whether or not these differences would impact on the comparability of the findings (and they do not appear to have been considered in the Cochrane review[34]). The statistically significant reductions in time to eye opening ranged from 2.72 to 5.9 minutes. It is not possible to draw any firm conclusions about the clinical significance of these reductions (e.g. their implications for health services) because the majority of the RCTs did not detect significant reductions in time to eye opening; one of the four trials that did report a significant effect is at high risk of bias because of the authors' conflict of interests[57] (see *Quantity and quality of research available*); and the pooled effect estimate from the Cochrane review,[34] although statistically significant, has high heterogeneity in the random-effects model used.

Time to extubation

One of the 11 RCTs that assessed this outcome was powered statistically to detect a specific difference (of 3 minutes) between the study groups, but did not detect a significant effect of Narcotrend monitoring on time to extubation.[60] Among the remaining 10 RCTs, six reported a significant reduction in the time to extubation, which, in all cases, favoured the EEG group relative to standard clinical monitoring. The reductions in time to extubation in these six trials ranged from 1.4 minutes to 6 minutes, with the largest reductions being for Narcotrend-guided total i.v. anaesthesia in adults (6 minutes),[63] BIS-guided volatile anaesthesia in children (5 minutes)[53] and BIS-guided volatile anaesthesia in adults (4.2 minutes).[45]

In general, the same cautions in interpreting these results apply as noted above for the time to eye opening. Taking these limitations into consideration, there appears to be an overall favourable effect of EEG-guided anaesthetic monitoring on time to extubation but no clear pattern that would identify possible explanatory variables (such as the importance of population, EEG monitor or type of anaesthesia). It is unlikely that a saving of 6 minutes (the best achieved) in the time to extubation would have importance for patients or for service provision, given that it represents $< 10\%$ of the total time patients were undergoing surgical procedures.

Outcomes related to postanaesthesia care unit stay

None of the RCTs that assessed outcomes related to PACU stay was specifically powered statistically to detect differences in these outcomes.

All five RCTs that reported the time to arrival at the PACU found that the arrival time was significantly shorter under EEG-guided anaesthesia than following standard clinical monitoring.[48,56,57,63,64] Together, these RCTs represented both adults and children, different types of anaesthesia, and different EEG

monitoring devices. The time savings ranged from 1.4 minutes to 5.8 minutes, with the largest differences being for Narcotrend-guided TIVA in adults (5.8 minutes),[63] BIS-guided mixed anaesthesia in adults (4.7 minutes)[48] and E-Entropy-guided mixed anaesthesia in children (4.0 minutes).[56] A difficulty in comparing these studies is that the starting point for measuring the time of arrival at the PACU was variable and sometimes unclear.

The two RCTs that reported the duration of stay in the PACU both examined BIS-guided volatile anaesthesia in children and both reported significant reductions in the duration of stay in the BIS-guided anaesthesia group compared with standard clinical monitoring.[52,53] In these RCTs the time savings in PACU stay ranged from 16 minutes[53] to 26 minutes.[52] These RCTs, which were both by Messieha and colleagues,[52,53] were similar and studied children undergoing complete dental rehabilitation. A notable difference is that in one RCT the target BIS value was 55–65,[53] whereas in the other RCT the target BIS value was 65–70.[52] Although the higher BIS values in the latter trial would represent lighter depth of anaesthesia, both of these trials supplemented their BIS-guided anaesthesia with monitoring of clinical signs, which makes it difficult to determine whether or not the differences between the trials in PACU stay relate directly to the use of different target BIS values.

Three of the five RCTs that reported time to PACU discharge found significant differences between EEG-guided anaesthesia and standard clinical monitoring.[45,48,52] These trials were all on BIS-guided anaesthesia, and included volatile anaesthesia in adults,[45] mixed anaesthesia in adults[48] or volatile anaesthesia in children.[52] In all cases the time to discharge was shorter in the BIS-guided group, with the time saved ranging from 6.7 minutes to 30 minutes. The trials that reported the longest time savings, of 30 minutes[52] and 24.7 minutes,[48] both measured time to discharge from the end of GA. These reductions in discharge times are relatively large compared with the total durations of surgery in these trials, which were approximately 91 minutes (adults)[48] and 139 minutes (children),[52] suggesting possible benefits for patient throughput or PACU bed occupancy, as well as indicating improved clinical recovery of patients.

As noted above, the 'PACU stay' outcome in the Cochrane review[34] seems to combine different aspects of time to PACU arrival, stay and/or discharge so may be difficult to interpret precisely. The outcome is consistent with the overall results of the individual RCTs included in the current systematic review, which indicate that EEG-guided anaesthesia reduces time to PACU admission, stay and discharge. However, although the pooled effect estimate in the Cochrane review is statistically significant, it has high statistical heterogeneity in the random-effects model used.

Time to response to commands

One RCT was powered statistically to detect a 20% difference in the time to response to verbal commands.[57] This trial and a further RCT[59] reported statistically significant reductions in time to response in E-Entropy-guided anaesthesia[57] and Narcotrend-guided anaesthesia[59] compared with standard clinical practice. Both these trials were on adults receiving TIVA. The third RCT, on children receiving TIVA, did not provide quantitative data but stated that the study groups were comparable.[46] The reductions in time to response to commands were 4.1 minutes (median) for time to hand squeezing on command (start time not reported)[57] and 4.6 minutes (mean) for time from end of anaesthetic to eye opening on command (also referred to as 'arousal time').[59]

Time to recovery of orientation

The three RCTs measuring this outcome all reported statistically significant reductions in time to orientation in E-Entropy-guided[54,57] or Narcotrend-guided[59] anaesthesia compared with standard clinical practice. The reported time savings were 4.8 minutes (median) in E-Entropy-guided TIVA in adults,[57] 5.1 minutes (mean) in E-Entropy-guided volatile anaesthesia in children[54] and 5.6 minutes (mean) in Narcotrend-guided TIVA in adults.[59] However, these RCTs were not specifically powered for this outcome; none of them defined orientation, and only one defined the time period to orientation [stated as the time between opening eyes on command and (undefined) orientation[59]].

Time to first movement response

Both of the RCTs measuring this outcome examined BIS-guided volatile anaesthesia, in adults[45] or children.[51] The latter RCT was powered statistically to detect a 30% reduction in the time to first movement response. Both the trials reported statistically significant reductions in time to first movement in the BIS-guided anaesthesia group compared with standard clinical monitoring. The mean time savings were 2.8 minutes[45] and 2.5 minutes.[51]

Time to achieve specified recovery scores

Both of the RCTs measuring this outcome evaluated E-Entropy-guided anaesthesia in children who received either volatile anaesthetic (sevoflurane)[54] or mixed anaesthetic (comprising propofol or sevoflurane for induction and isoflurane for maintenance).[56] One trial defined time to complete recovery as the time to reach a score of ≥9 on a modified Aldrete scale.[54] In the other trial time to recovery was defined as the time to reach a score of 6 on a modified Steward scale.[56] Time to recovery was significantly shorter, by a mean of 4.5 minutes, in the E-Entropy-guided than the standard clinical practice group in one trial (Aldrete score),[54] but did not differ significantly in the other trial.[56]

Time to spontaneous breathing

This RCT[57] evaluated BIS-guided TIVA in adults and found a significantly shorter time to spontaneous breathing in the E-Entropy-guided than the standard clinical practice group. The median time difference was 2.33 minutes. Limitations to interpretation are: the RCT was not powered specifically for this outcome; the time to spontaneous breathing was not formally defined.

Time to laryngeal mask airway removal and time to phonation

This RCT[51] evaluated BIS-guided volatile anaesthesia in children. The times from the last surgical suture to removal of the laryngeal mask airway and to phonation did not differ significantly between the BIS and standard clinical practice groups. A potential limitation to interpretation is that this trial was not specifically powered to detect differences in these outcomes.

Adverse effects of anaesthesia

Few of the trials reported anaesthesia-related adverse effects outcomes. The most frequently reported adverse outcomes were PONV (four RCTs), postoperative pain (two RCTs), POCD in elderly patients (one RCT) and emergence delirium in children (one RCT). These adverse effects are particularly relevant to situations in which overdosing of anaesthesia occurs. They were all reported as secondary outcomes (i.e. they were not specifically powered statistically) in the RCTs.

Postoperative nausea and vomiting

The four RCTs reporting this outcome evaluated BIS-guided volatile anaesthesia in children,[51] entropy-guided TIVA in adults[55] and Narcotrend-guided TIVA in adults.[59,60] In two trials PONV occurred but did not differ significantly in frequency between standard clinical monitoring and the BIS group[51] or E-Entropy group.[55] In the third trial no cases of PONV occurred in either the Narcotrend or standard monitoring practice groups.[59] The remaining RCT reported PONV scores based on a VAS (no details provided) rather than frequency of occurrence, and found significantly higher (better) scores (indicating less frequent PONV) in the Narcotrend group compared with standard clinical practice.[60] However, this difference was significant only 10 minutes after the end of surgery and not at 30 or 90 minutes post surgery.

Postoperative pain

The two RCTs that assessed postoperative pain evaluated E-Entropy-guided anaesthesia, either in adults under TIVA[55] or in children under mixed anaesthesia.[56] Pain was assessed as a score on a 0–10 scale[55] or using the CHEOPS.[56] Pain scores were significantly lower in the E-Entropy group than standard clinical monitoring for the adult population.[55] In the paediatric population, the CHEOPS scores were significantly lower in the E-Entropy group at 60, 90 and 120 minutes after arrival in the PACU but not at 30 minutes after arrival.[56]

Postoperative cognitive dysfunction

The RCT that assessed this outcome evaluated BIS-guided i.v. anaesthesia in elderly patients.[47] At 1 week post surgery, the incidence of POCD was 32.5% in the E-Entropy group and 39.1% in the standard clinical monitoring group. At 3 months post surgery the incidences were 8.1% and 12.0% respectively. Only the 3-month results were statistically significant. Interpretation is limited because the RCT is reported only in a conference abstract, which provides very limited information about the study.

Emergence delirium

The RCT that assessed this outcome was a study of BIS-guided volatile anaesthesia in children.[51] In this trial, emergence delirium was assessed using the PAED Instrument. The highest PAED scores recorded during the first 30 minutes after awakening were compared between the study groups and did not differ significantly.

Economic evaluation

Systematic review of published economic evaluations

Systematic searches identified 134 potentially relevant references. Studies were eligible for inclusion if they were full economic evaluations, including an assessment of any depth of anaesthesia monitoring device, conducted in patients receiving general anaesthetic for surgery. One study met all of the a priori inclusion criteria. This was a cost-effectiveness study reporting outcomes as cost of preventing an episode of awareness in all patients[97] using data drawn from a prospective study by Ekman and colleagues[98] and from the RCT reported by Myles and colleagues[79] and Avidan and colleagues.[27] The analysis was limited only to the cost of the BIS and sensors to be attached to the patient, whereas outcomes were limited to cases of awareness. Based on an estimated incidence of awareness of 0.04% with BIS and 0.18% with standard clinical monitoring the cost-effectiveness of depth of anaesthesia monitoring was estimated as US$4410 per case avoided. The authors of the study concluded that the use of BIS monitoring was unlikely to be cost-effective. However, the results and conclusions should be viewed with caution because of weaknesses in methodology and poor reporting quality.

De novo economic evaluation

We developed a decision-analytic model to assess the cost-effectiveness of depth of anaesthesia monitoring compared with standard clinical monitoring. The model incorporated evidence on outcomes from the systematic review of patient outcomes (change in anaesthetic drug consumption, change in incidence of awareness and POCD) combined with data identified through targeted searches (incidence of long-term psychological sequelae of awareness, duration and cost of PTSD, QOL impact of LPS and PTSD, duration of POCD). Outcomes in the model are expressed as QALY. The model evaluates costs from the perspective of the NHS and Personal Social Services. Costs are expressed in UK sterling (pounds, £) at a 2011 price base. Cost-effectiveness was assessed using ICER for each technology, compared with standard clinical monitoring. Separate analyses are presented for each of the included technologies, compared with standard clinical monitoring – the included technologies are not compared with each other as this was not within the scope of the appraisal issued by NICE.

Bispectral Index compared with standard clinical monitoring

We presented a base-case analysis for two modes of anaesthetic administration [TIVA and mixed anaesthesia (induction with i.v. anaesthesia and maintenance with inhaled anaesthesia or a combination of inhaled and i.v. anaesthetic)] and for two patient populations (those considered at high risk of intraoperative awareness and a general surgical population, at average risk of intraoperative awareness).

For patients undergoing GA with TIVA, we used the OR of awareness with BIS monitoring (0.24), compared with standard clinical monitoring, reported in the meta-analysis in our systematic review of patient outcomes (see *Results of systematic review of patient outcomes*) and baseline awareness risks identified and pooled in this assessment (0.45% in patients at high risk of intraoperative awareness and

0.16% for a general surgical population, at average risk of intraoperative awareness) to estimate the risk reduction for awareness and its psychological sequelae associated with BIS monitoring. All of the trials included in the meta-analysis were conducted in patients at high risk of awareness. In the absence of any evidence on the effectiveness of BIS on the incidence of awareness in the general surgical population, we applied the same OR reported in the meta-analysis to both groups of patients.

Anaesthetic drug costs were based on reported consumption in trials included in the meta-analysis reported in the systematic review of patient outcomes (see *Results of systematic review of patient outcomes*). None of the trials included in the meta-analysis of drug consumption were conducted in patients at high risk of awareness, as these did not report anaesthetic drug consumption. In the model we assumed that the clinical characteristics of high-risk patients mean that anaesthetists will be particularly cautious regarding the dose of anaesthetic drugs and that the higher risk of awareness is associated with a tendency to underdose patients. As a result, we assumed that the potential reduction in anaesthetic dose, through the use of depth of anaesthesia monitoring, would not apply in this group of patients.

In cohorts of 10,000 patients, at high risk of intraoperative awarenes GA with TIVA, BIS monitoring was modelled as being associated with 10.8 cases of awareness, compared with 45 in patients receiving standard clinical monitoring. This resulted in a reduction of 11 cases of LPS (from 14.7 to 3.5), which included a reduction of six cases of PTSD (from 8.0 to 1.9). The modelled cost per patient was higher with BIS monitoring than for standard clinical monitoring, although some of the additional cost was offset by reduced costs associated with psychological sequelae of awareness. The majority of the additional cost of BIS monitoring was attributable to the sensors attached to the patient (88% of additional cost, per patient). By reducing the incidence of awareness and longer-term effects of POCD, BIS monitoring was associated with improved outcomes. The ICER, for BIS compared with standard clinical monitoring in this population was £22,339. Deterministic sensitivity analyses indicated that the ICER was sensitive to the baseline incidence of awareness, effectiveness of BIS in reducing awareness, probability of LPS, QoL decrement for PTSD, the proportion of people with LPS who have PTSD and the cost of sensors. Scenario analyses were undertaken to address the question of variables omitted from the base case and to explore the impact of key baseline assumptions. These indicate that the cost-effectiveness results were largely insensitive to including an effect of BIS on PONV and to assumptions regarding patient throughput (except at comparatively low volumes, below 500 cases per year per module), whereas they were highly sensitive to assumptions regarding the baseline risk of awareness and the QoL decrement for PTSD.

For the population of general surgical patients undergoing GA with TIVA, BIS monitoring was modelled as being associated with 3.8 cases (per 10,000 patients) of awareness, compared with 16 in patients receiving standard clinical monitoring. This resulted in a reduction of four cases of LPS (from 5.2 to 1.3), which included a reduction of two cases of PTSD (from 2.8 to 0.7). Although the modelled cost per patient was higher with BIS than with standard clinical monitoring, a larger proportion was offset by reductions in other costs (primarily anaesthetic drug costs) than was the case for patients at high risk of intraoperative awareness (where no saving in anaesthetic drug costs was included). As with the analysis for high-risk patients, the majority of the additional cost of BIS monitoring was attributable to the sensors attached to the patient, rather than the monitor module itself. Given the lower baseline risk of awareness in this population, the QALY gain with BIS monitoring was lower (0.0003) than for high-risk patients. This resulted in a higher ICER (£34,565) despite the lower incremental cost estimated for this population, arising from reduced anaesthetic consumption. Deterministic sensitivity analyses indicated that the ICER was sensitive to the same input parameters as for the population at high risk of awareness. In the majority of cases the ICER remained above £30,000 per QALY gained – the most favourable ICER was associated with a reduction in the cost of sensors. Conclusions from the scenario analyses were similar to those undertaken for high-risk patients. In particular, more favourable ICERs were associated with a higher baseline incidence of awareness and with a higher utility decrement for PTSD.

For patients undergoing mixed GA (induction with i.v. and maintenance including inhaled anaesthetic), we used the pooled OR of awareness with BIS monitoring, compared with standard clinical monitoring,

calculated in the meta-analysis reported in the systematic review of patient outcomes (0.45) and baseline awareness risks identified and pooled in this review to estimate the risk reduction for awareness and its psychological sequelae associated with BIS monitoring.

The baseline estimates of awareness, LPS and PTSD were the same as for high-risk patients undergoing TIVA (45, 14.7 and 8 per 10,000 patients respectively). However, given that the OR of awareness with BIS monitoring was higher in this analysis, the estimated reduction in LPS and PTSD was lower. In this patient population BIS monitoring was associated with 20.3 cases of awareness, 6.6 cases of LPS, including 3.6 cases of PTSD. BIS monitoring had higher costs and improved outcomes compared with standard clinical monitoring. However, the QALY gain (0.0005) was lower than for patients undergoing TIVA. The ICER, for BIS compared with standard clinical monitoring in this population was £29,634. Deterministic sensitivity analyses indicated that the ICER was sensitive to the baseline incidence of awareness, effectiveness of BIS in reducing awareness, probability of LPS, QoL decrement for PTSD, the proportion of people with LPS who have PTSD and the cost of sensors. The highest incidence of awareness (1.19%), largest effect size (0.25, lower 95% confidence limit for OR of awareness with BIS vs standard clinical care), longest duration of LPS (1 year), highest probability of LPS (0.48), longest duration of PTSD (21.6 years), highest probability of PTSD (0.239), greatest utility reduction associated with PTSD (–0.134), largest proportion of PTSD patients being treated (100%) and lowest cost of sensors (£10.875, 75% of base-case value) tested in the sensitivity analysis resulted in ICERs below £30,000 per QALY gained, although the majority remained above £20,000 per QALY gained. Conclusions from the scenario analyses were similar to those for high-risk patients undergoing TIVA.

The baseline estimates of awareness, LPS and PTSD in the population of general surgical patients undergoing mixed GA were the same as for TIVA (16, 5.2 and 2.8 per 10,000 patients respectively), whereas BIS monitoring in this patient population was modelled as being associated with 7.2, 2.3 and 1.3 cases respectively. Although a proportion of the higher cost associated with BIS monitoring was offset by reduction in anaesthetic consumption, the cost-saving for inhaled anaesthesia was lower than for TIVA. As a result the incremental cost was greater (£12.91 compared with £10.98). Given the lower baseline risk of awareness in this population, the QALY gain with BIS monitoring was lower (0.0003) than for high-risk patients, resulting in a higher ICER (£49,198). Deterministic sensitivity analyses indicated that the ICER was sensitive to the same input parameters as for the population at high risk of awareness. However, in all cases the ICER remained above conventional thresholds – the most favourable ICER was associated with a reduction in the cost of sensors. Conclusions from the scenario analyses were also similar to those undertaken for high-risk patients.

E-Entropy compared with standard clinical monitoring

A base-case analysis was presented for two modes of anaesthetic administration [TIVA and mixed anaesthesia (induction with i.v. anaesthesia and maintenance with inhaled anaesthesia or a combination of inhaled and i.v. anaesthetic)] and for two patient populations (those considered at high risk of intraoperative awareness and a general surgical population, at average risk of intraoperative awareness).

Insufficient evidence was identified to estimate the effectiveness of depth of anaesthesia monitoring with E-Entropy on the incidence of intraoperative awareness or on POCD. In the absence of evidence specific to E-Entropy we have applied the effectiveness estimates derived for BIS, described above. This meant that the modelled clinical effectiveness of E-Entropy was identical to that reported for BIS – this is an untested assumption and must be considered a weakness in the evidence base for E-Entropy. Anaesthetic drug costs were based on consumption reported in the included trials, and were valued using current unit costs.

In patients considered at high risk of awareness undergoing GA with TIVA, the modelled cost per patient with E-Entropy monitoring was higher than with standard clinical monitoring, although some of the additional cost was offset by reduced cost associated with psychological sequelae of awareness. The additional cost of E-Entropy monitoring was approximately two-thirds that of BIS monitoring, with the majority being attributable to the sensors attached to the patient (80% of additional cost per patient).

E-Entropy monitoring was associated with improved outcomes, based on applying clinical effectiveness evidence reported for BIS. The ICER for E-Entropy compared with standard clinical monitoring in this population was £14,421. Deterministic sensitivity analyses indicated that the ICER was sensitive to the baseline incidence of awareness, effectiveness of E-Entropy in reducing awareness, probability of LPS, QoL decrement for PTSD, the proportion of people with LPS who have PTSD and the cost of sensors. The least favourable ICERs were for low baseline incidence of awareness, lower effectiveness on incidence of awareness and a lower probability of patients with awareness developing LPS. Scenario analyses, undertaken to consider variables omitted from the base case and to explore the impact of key baseline assumptions, indicated that the cost-effectiveness results were highly sensitive to assumptions regarding the baseline risk of awareness and the QoL decrement for PTSD, whereas they were largely insensitive to including an effect of E-Entropy on PONV and to assumptions regarding patient throughput (except at comparatively low volumes, $<$ 500 cases per year per module).

In the population of general surgical patients undergoing GA with TIVA, E-Entropy monitoring had a higher cost per patient than standard clinical monitoring. Anaesthetic drug costs derived from two clinical trials were modelled separately, as we considered them unsuitable for pooling, given substantial differences in the patient populations (one trial in orthopaedic surgery and the other in elective gynaecological laparoscopy). Neither of the trials showed an overall reduction in anaesthetic drug consumption and as a result there was no reduction in anaesthetic drug costs to offset the additional costs of E-Entropy monitoring. As with the analysis for high-risk patients, the majority of the additional cost of monitoring was attributable to the sensors attached to the patient. Given the lower baseline risk of awareness in this population, the QALY gain was lower than for high-risk patients, which resulted in a higher ICER (£31,131–31,430). Deterministic sensitivity analyses indicated that the ICER was sensitive to the baseline incidence of awareness, effectiveness in reducing awareness, probability of LPS, QoL decrement for PTSD, the proportion of people with LPS who have PTSD and the cost of sensors. The lower limit of anaesthetic drug consumption, the highest incidence of awareness, largest effect size, greatest probability of LPS, longest duration of PTSD, greatest probability of PTSD and lowest cost of sensors tested in the sensitivity analysis resulted in an ICER of $<$ £30,000 per QALY gained, although they remained above £20,000 per QALY gained. Conclusions from the scenario analyses were similar to those undertaken for high-risk patients.

As noted above, in the absence of evidence specific to E-Entropy, we have applied the effectiveness estimates derived for BIS in this analysis. For patients undergoing mixed GA (induction with i.v. and maintenance including inhaled anaesthetic), the pooled OR of awareness with BIS monitoring, compared with standard clinical monitoring, (0.45) was higher than for TIVA, resulting in a smaller reduction in cases of awareness, LPS and PTSD.

In patients considered at high risk of awareness undergoing mixed GA, E-Entropy monitoring had higher costs and improved outcomes compared with standard clinical monitoring. However, the QALY gain (0.0005) was lower than for patients undergoing TIVA. The ICER for E-Entropy compared with standard clinical monitoring in this population was £19,367. Deterministic sensitivity analyses indicated that the ICER was sensitive to the baseline incidence of awareness, effectiveness in reducing awareness, probability of LPS, QoL decrement for PTSD, the proportion of people with LPS who have PTSD and the cost of sensors. The least favourable ICERs were found with a low incidence of awareness (lower limit of 95% CI), lesser effect size (upper limit of 95% CI for OR of awareness with monitoring vs standard clinical care) and greater probability of LPS (0.48). The majority of the ICERs remained at $<$ £20,000 per QALY gained. Conclusions from the scenario analyses were similar to those undertaken for high-risk patients undergoing TIVA.

In the population of general surgical patients undergoing mixed GA, E-Entropy monitoring had higher costs than standard clinical monitoring. In contrast with the analysis for TIVA, the clinical trial used to estimate inhaled anaesthetic drug consumption reported a substantial decrease (29%), which resulted in approximately half of the additional cost of E-Entropy monitoring being offset by a reduction in

anaesthetic drug costs. Despite the lower baseline risk of awareness, which resulted in a lower QALY gain with E-Entropy monitoring than for high-risk patients, the lower incremental cost resulted in an equivalent ICER (£19,000). Deterministic sensitivity analyses indicated that the ICER was sensitive to the same input parameters as for the population at high risk of awareness. The least favourable ICERs were found with a low reduction in anaesthetic drug consumption (lower limit of 95% CI) and lesser effect size (upper limit of 95% CI for OR of awareness with monitoring vs standard clinical care). The majority of the ICERs remained below £20,000 per QALY gained. Conclusions from the scenario analyses were also similar to those undertaken for high-risk patients.

Narcotrend compared with standard clinical monitoring

We presented a base-case analysis for two modes of anaesthetic administration [TIVA and mixed anaesthesia (induction with i.v. anaesthesia and maintenance with inhaled anaesthesia or a combination of inhaled and i.v. anaesthetic)] and for two patient populations (those considered at high risk of intraoperative awareness and a general surgical population, at average risk of intraoperative awareness).

Anaesthetic drug costs were based on consumption reported in the included trials, and were valued using current unit costs. Insufficient evidence was identified to estimate the effectiveness of depth of anaesthesia monitoring with Narcotrend on the incidence of intraoperative awareness or on POCD. In the absence of evidence specific to Narcotrend, we have applied the effectiveness estimates derived for BIS, described above. This means that the modelled clinical effectiveness of Narcotrend is identical to that reported for BIS – this is an untested assumption and must be considered a weakness in the evidence base for Narcotrend.

In patients considered at high risk of awareness undergoing GA with TIVA, the modelled cost per patient with Narcotrend monitoring was higher than with standard clinical monitoring, although some of the additional cost was offset by reduced cost associated with psychological sequelae of awareness. The additional cost of Narcotrend monitoring was approximately half that of E-Entropy monitoring, and approximately one-quarter that of BIS – primarily because of differences in the cost of the sensors attached to the patient. In contrast with BIS and E-Entropy, the majority of the additional cost of Narcotrend monitoring was attributable to the monitor (90% of additional cost per patient) rather than the sensors. Narcotrend monitoring was associated with improved outcomes, based on applying clinical effectiveness evidence reported for BIS. The ICER for Narcotrend compared with standard clinical monitoring in this population was £5681. Deterministic sensitivity analyses indicated that the ICER was sensitive to the baseline incidence of awareness, effectiveness in reducing awareness, probability of LPS, QoL decrement for PTSD, the proportion of people with LPS who have PTSD and the cost of sensors. The least favourable ICER was for low baseline incidence of awareness. Scenario analyses, undertaken to consider variables omitted from the base case and to explore the impact of key baseline assumptions, indicated that the cost-effectiveness results were highly sensitive to assumptions regarding the baseline risk of awareness and the QoL decrement for PTSD, whereas they were largely insensitive to including an effect of E-Entropy on PONV.

In the general surgical population undergoing GA with TIVA Narcotrend monitoring had a lower cost per patient than standard clinical monitoring. The additional cost of monitoring was reduced to £2.84 per patient (£2.28 per patient for the monitor and £0.56 for the sensors attached to the patient). This was more than offset by reduction in anaesthetic drug consumption. Given the lower baseline risk of awareness in this population, the QALY gain was lower than for high-risk patients. However, given that Narcotrend was associated with improved outcomes and reduced costs it dominated standard clinical monitoring. Narcotrend remained dominant in all the deterministic sensitivity analyses. Conclusions from the scenario analyses were similar to those undertaken for high-risk patients.

As noted above, in the absence of evidence specific to Narcotrend, we have applied the effectiveness estimates derived for BIS in this analysis. For patients undergoing mixed GA (induction with i.v. and maintenance including inhaled anaesthetic), the pooled OR of awareness with BIS monitoring compared

with standard clinical monitoring (0.45) is higher than for TIVA, resulting in a smaller reduction in cases of awareness, LPS and PTSD.

In patients considered at high risk of awareness undergoing mixed GA, Narcotrend monitoring had higher costs and improved outcomes compared with standard clinical monitoring, although the QALY gain (0.0005) was lower than for patients undergoing TIVA. The ICER for Narcotrend compared with standard clinical monitoring in this population was £8033. Deterministic sensitivity analyses indicated that the ICER was sensitive to the baseline incidence of awareness, effectiveness in reducing awareness, probability of LPS, QoL decrement for PTSD, the proportion of people with LPS who have PTSD and the cost of sensors. The least favourable ICERs were found with a low incidence of awareness (lower limit of 95% CI) and lesser effect size (upper limit of 95% CI for OR of awareness with monitoring vs standard clinical care). Conclusions from the scenario analyses were similar to those undertaken for high-risk patients undergoing TIVA.

In the population of general surgical patients undergoing mixed GA, Narcotrend monitoring had higher costs than standard clinical monitoring. Although the proportionate reduction in consumption of inhaled anaesthetic (desflurane) was lower than the reduction in i.v. anaesthetic (propofol) for TIVA, the reduction in cost of anaesthetic (£4.26) was sufficient to offset the additional cost of Narcotrend monitoring (£2.84). Given the lower baseline risk of awareness in this population, the QALY gain was lower than for high-risk patients. However, as Narcotrend was associated with improved outcomes and reduced costs, it dominated standard clinical monitoring. Narcotrend remained dominant in the majority of deterministic sensitivity analyses. At the upper limit of the 95% CI for proportional change in desflurane use, the reduction in cost of anaesthetic was insufficient to offset the additional cost of Narcotrend monitoring and the resulting ICER was £2534. Conclusions from the scenario analyses were similar to those undertaken for high-risk patients.

Strengths and limitations of the assessment

The current evidence synthesis followed an accepted standard procedure for conducting a systematic review of the evidence, based on a published protocol, so as to minimise bias and, where possible, provide the most precise estimates of effects for relevant outcomes. The work was carried out by a team experienced in health technology appraisal, independent of any vested interest.

We know of only two other relevant systematic reviews in this topic area, both of which focused on the effects of BIS-guided depth of anaesthetic monitoring. A systematic review and meta-analysis reported by Liu (2004)[105] investigated the use of BIS-guided anaesthetic delivery in ambulatory anaesthesia. Eleven RCTs were included and BIS-guided anaesthesia was found to significantly reduce anaesthetic consumption, PONV and time spent in the recovery room (PACU). However, the benefits did not reduce the time spent in the ambulatory surgery unit overall. More recently, a more comprehensive Cochrane systematic review and meta-analysis of the use of BIS monitoring to improve anaesthetic delivery and postoperative recovery, not limited to ambulatory anaesthesia, was conducted by Punjasawadwong and colleagues.[34] As noted above, our current systematic review complements this Cochrane review for BIS studies and, where possible, we have updated meta-analyses in the Cochrane review using data from new RCTs that we have identified. For pragmatic reasons (to keep the work manageable with the available resources), we did not duplicate the searches of the Cochrane review or re-extract data for those RCTs already included in it, but instead systematically sought and appraised new RCTs about BIS-guided anaesthesia that have been published since the search dates of the Cochrane review. The Cochrane review was limited to RCTs on adults but, as specified in the protocol, we have included in our systematic review RCTs on children as well as adults. In practice, we found new evidence to update the Cochrane review meta-analysis for three outcomes (intraoperative awareness, consumption of volatile anaesthetic and consumption of i.v. anaesthetic), although for anaesthetic consumption the precision of the existing effect estimates was not necessarily improved because of significant statistical heterogeneity. A disadvantage of our pragmatic approach is

that we have not presented full details of those BIS trials included in the Cochrane review, although these can be ascertained from the Cochrane review itself. Although Cochrane reviews are generally conducted to high standards, there appear to be some limitations in the publication by Punjasawadwong and colleagues,[34] which we have noted above when interpreting specific outcomes. For instance, a meta-analysis relating to 'PACU stay' appears to have combined several outcomes concerning the time to PACU arrival, stay and discharge, which would be more informative if analysed separately.

As no systematic reviews of E-Entropy-guided anaesthesia or Narcotrend-guided anaesthesia appear to have been published, for these technologies we conducted more extensive searches to locate all relevant RCTs, on both adults and children, which were then screened for relevance and, where they met the inclusion criteria, were subjected to full systematic review. The current work represents the most comprehensive systematic review of BIS-, E-Entropy- and Narcotrend-guided anaesthesia that has been conducted to date.

A notable limitation to our assessment of patient outcomes is that the quality of reporting in the primary studies was often limited, which gives rise to numerous uncertainties in the interpretation of the primary evidence (see *Uncertainties*). As discussed above, the studies were diverse in their methodological characteristics, which limited opportunities to pool their data in meta-analyses. The primary studies also predominantly reported secondary outcomes, which were often based on relatively small sample sizes, with unknown statistical validity.

We undertook a comprehensive search for studies that would be potentially relevant to the assessment of cost-effectiveness, by identifying full economic evaluations of any depth of anaesthesia monitoring device compared with standard clinical monitoring. One published study was identified, which reported ICER as the incremental cost of BIS monitoring per case of intraoperative recall avoided. We did not identify any published economic evaluations that reported outcomes in terms of QALY or similar units, nor did we identify any studies that explicitly compared the additional costs of depth of anaesthesia monitoring with potential savings in anaesthetic drug use. We developed a de novo decision-analytic model to provide an assessment of the cost-effectiveness of depth of anaesthesia monitoring, compared with standard clinical monitoring, incorporating patient outcomes (in terms of avoided cases of PTSD and POCD) as QALY and anaesthetic drug use. The model provided a means to synthesise data from the systematic review of patient outcomes (in terms of the effectiveness of depth of anaesthesia monitoring on intraoperative awareness, POCD and anaesthetic drug use). This was supplemented by information identified using targeted searches (on the baseline incidence of intraoperative awareness in high-risk and general surgical populations, proportion of patients who experience POCD, the proportion of patients experiencing intraoperative awareness who develop long-term psychological illness, the duration, cost and the QoL impact of those conditions).

In the model, GA exposes patients to a risk of intraoperative awareness that is defined either as high or average (the latter corresponding to the risk of awareness in the general surgical population), and to POCD, which have consequences for QoL. In patients experiencing long-term psychological illness as a consequence of an awareness episode, there are also associated health-care costs. Other costs considered in the model are costs of anaesthetic drugs, as well as the cost of the depth of anaesthesia monitors. Cost-effectiveness was assessed by estimating ICER for each mode of anaesthesia and each technology. We undertook a range of sensitivity analyses and scenario analysis to identify the key determinants of the cost-effectiveness results as well as the impact of key assumptions and of variables missing from the analysis.

Evidence to populate the model was limited. In particular no evidence on the effectiveness of E-Entropy and Narcotrend on the incidence of intraoperative awareness was identified. In the case of BIS, where such evidence was identified it was limited to patients considered at high risk of awareness. We have assumed in the model that the effectiveness evidence in high-risk patients can be applied to the general surgical population (at average risk of awareness) and that the effectiveness evidence for BIS can be applied to both E-Entropy and Narcotrend. These are untested assumptions and must be considered a weakness in

the cost-effectiveness evidence base. Whereas more evidence is available on the baseline risk of awareness, there was considerable inconsistency in the estimated incidence in studies identified in our targeted searches. As a result we used pooled values (excluding outliers) in the base-case analyses, with the outlying values adopted in scenario analyses.

Evidence on the effectiveness of any depth of anaesthesia monitoring on POCD is also limited to BIS, with the only published study being available only in abstract form. As with the evidence on effectiveness with respect to awareness, we have assumed that evidence for BIS can also be applied to E-Entropy and Narcotrend – again this is an untested assumption. Evidence on the incidence of POCD was also limited and is subject to considerable uncertainty (primarily concerning the extent to which pre-existing, but unrecognised, cognitive dysfunction may be incorrectly identified as a postoperative complication). The best evidence we could identify that reported POCD in patients who had been assessed preoperatively compared with a matched group of non-surgical controls is over 10 years old, and it is not clear whether or not this will reflect incidence of POCD in current practice.

Although we were able to identify some evidence on the incidence of PTSD in patients who experienced awareness during GA, we did not identify any studies reporting overall QoL impact, health state utilities or mean duration of symptoms in PTSD sufferers with awareness as the trigger. The evidence base for people with PTSD relates to a range of trauma exposures (including military service and other wartime exposures, natural disaster, domestic abuse), and it is not clear whether or not this can be applied directly to people who have developed psychological illness following intraoperative awareness.

We adopted a modelling approach that did not explicitly identify patients exposed to overdose or underdose of anaesthetics, although this may allow a clearer assessment of the potential benefits of depth of anaesthesia monitoring. Intraoperative awareness may be identified as being particularly closely associated with anaesthetic underdose, whereas POCD and PONV may be more closely associated with overdose. The potential for savings in terms of anaesthetic drug use may also primarily arise in this latter group. Although it may have been preferable to adopt this more explicit structure, we did not identify data to support this approach. We have therefore adopted a more simple model structure, although we have implicitly incorporated some of these assumptions into our model.

The scope of the appraisal issued by NICE required the three depth of anaesthesia monitoring devices to be compared with standard clinical monitoring, rather than with each other. A direct comparison of the cost-effectiveness of the three technologies was therefore not conducted. However, such a comparison would not be feasible as there is limited direct trial evidence comparing the technologies with each other, and an indirect comparison would not be possible because of the lack of outcome data on intraoperative awareness for E-Entropy and Narcotrend. Indeed, because of the lack of awareness data, one of the assumptions we have had to make is that all three technologies would be similar in terms of preventing awareness, based on the data available for BIS.

Uncertainties

One of the biggest uncertainties in the evidence base assessed in this report is the impact of EEG monitoring on intraoperative awareness, and other significant adverse effects such as POCD. The lack of outcome data from RCTs on awareness was particularly the case for E-Entropy and Narcotrend. Likewise, the only RCT data available for POCD was for BIS and was available only in a conference abstract. In situations where evidence for specific outcomes from RCTs is lacking, it is pragmatic to use data from other types of study design, including non-experimental studies (e.g. cohort studies). However, we did not identify any such studies of BIS, E-Entropy and Narcotrend in our literature searches that reported on intraoperative awareness or POCD.

The nature of standard clinical monitoring varied across the included trials, with some trials giving more information than others. For example, in one study[59] it is reported that 'in the clinical group, the depth of anaesthesia was *primarily* evaluated by clinical indices including heart rate, blood pressure and body movement' (our emphasis) so it is not known what other methods may have been used. Also patients in the EEG arm of some of the trials were potentially assessed on the basis of standard clinical monitoring with the EEG reading used as an adjunct to other physiological parameters in assessing the effects of anaesthetic agents; however, this was not always explicitly stated in the trials. The BAG-RECALL trial by Avidan and colleagues[44] used 'structured protocols' to remind anaesthesiologists that patients may be aware, but not necessarily to prescribe changes in anaesthetic. As patients can have their anaesthesia adjusted on the basis of standard clinical monitoring or EEG monitoring or both, the effect is not solely as a result of the technology being considered (BIS, E-Entropy and Narcotrend) in the intervention arm of most of the studies.

Details of the technologies used in the trials are also often limited and confusing. It is not always clear or specified as to which monitor has been used or which version of the software has been used. There also seems to be some confusion between monitor version and software version in the reporting of the trials, and also between the terms 'monitor' and 'module'. For example, the trials of Narcotrend report Narcotrend Monitor version 2.0 AF,[60] Narcotrend monitor (software version 2.0 AF)[63,64] and Narcotrend monitor (MonitorTechnik, Germany).[59] This also happens in the studies reported in Cochrane review of BIS.[34] Anaesthesia monitors assess a range of parameters such as EEG, ECG, respiration, temperature, anaesthetic gases, and can be used for viewing and processing information (e.g. Datex-Ohmeda S/5 monitor); an EEG monitor with BIS, monitors the state of the brain by data acquisition of EEG and BIS is the processed EEG variable. However, device manufacturers also use the terms monitor and module interchangeably. This is probably because some monitors incorporate processing modules. For example, the A-2000 EEG monitor with BIS was upgraded to the A-3000 EEG monitor, which incorporated a BIS module and is known as a BIS monitor.

It therefore appears that the technologies considered are continually evolving, and different versions of the software have been used to interpret EEG readings in the different trials. It is not clear exactly what alterations have been made to the algorithms and how these influence the trial results as the algorithms are proprietary and not completely published. However, it is suggested by the manufacturer of BIS that update versions of the module have focused on artefact detection and removal, rather than fundamental changes to the algorithm. In the Narcotrend industry submission to NICE, one trial showed that Narcotrend does not differentiate reliably between conscious and unconscious patients. The reason given to explain these results is that both these studies were carried out using older versions of the algorithm and that the studies had methodological flaws. Whatever the reasons are for these results, this does emphasise the potential lack of consistency between the different versions and need for care when interpreting results from studies using different software versions.

There is also inconsistency in EEG values used in the trials, both overall and at different time points during surgery, making comparison across trials difficult. In the BIS trials there was notable variation in target values from 40 to 70. E-Entropy values during the maintenance phase of anaesthesia ranged from 35 to 60 for response entropy and 40–65 for state entropy, but in some trials higher values were permitted near the end of surgery, and the response entropy–state entropy difference was also used as a target value in some trials. Narcotrend values ranged from D_0 to C_1, D_2 to E_0 adjusted to D_0 to D_1 and D_2 to E_0, which means that the level of anaesthesia varied across trials within the same technology.

Outcomes were also defined differently in the different studies, which may affect results. For example, the starting point for the recovery process can be the last stitch performed during surgery or the end of application of dressings.

Other issues to consider when interpreting results are investigator bias (subtle unconscious or conscious influence of investigator on results which can overestimate results) and 'learning contamination bias' (the unintended improvement of standard clinical monitoring occurring with the introduction of a new monitoring device which can reduce the difference in results). Not many of the included studies discussed these aspects or reported experience of the anaesthetist. Ellerkmann and colleagues[62] used experienced anaesthesiologists and suggest that results may have been different had they used less experienced staff. Kreuer and colleagues[63] discount learning contamination bias in the standard clinical monitoring group of their trials as the anaesthesiologist was also experienced in use of Narcotrend/BIS.

Additional factors for consideration include inter-individual variability and sex differences in response to anaesthesia which complicate interpretation of results. For example in one trial, with comparable amounts of propofol, women in the standard clinical monitoring group had significantly shorter recovery times than men; in EEG-monitored groups (BIS and Narcotrend) propofol consumption was lower for men.[63] Also effects differ between i.v. anaesthesia and volatile anaesthetics and also depend on the specific drug used. For example, more rapid recovery can be expected with desflurane/remifentanil (which is washed out quicker) compared with propofol, so comparisons across trials using different anaesthetic agents are not valid. In addition, as anaesthesia is the interaction between hypnosis and sedation, the relative proportion of the drugs used to achieve these elements of anaesthesia may have an impact on EEG monitoring. Also, different approaches were used in the trials to manage inadequate anaesthesia, such as narcotics (fentanyl, sufentanil, alfentanil), which could impact on results.

Taking into account the above issues such as the methodology of the trials, the lack of clarity of reporting, the differences in patient characteristics and differences in technologies and anaesthesia used, brings into question the overall generalisability of the results and makes interpretation of results problematic, especially as some of the observed differences are minimal and may not be judged as clinically significant.

Chapter 6 Conclusions

In general, BIS, E-Entropy and Narcotrend technologies for monitoring the depth of anaesthesia are associated with reductions in general anaesthetic consumption and decreased anaesthetic recovery times, compared with monitoring of clinical signs alone. However, these reductions may be considered clinically modest. The available evidence on the impact of the technologies on reducing the likelihood of intraoperative awareness is limited. Overall, BIS was associated with a statistically significant reduction in intraoperative awareness in patients classified as at higher risk, although there is uncertainty in effect estimates because of significant heterogeneity. Caution is advised because of uncertainties about the risk of bias of many of the included trials, and because many outcome measures were not statistically powered.

The cost-effectiveness of depth of anaesthesia monitoring appears to be highly dependent on the incidence of awareness, the HRQoL impact of psychological sequelae of awareness, the probability of developing psychological illness following awareness as well as the effectiveness of depth of anaesthesia monitoring in reducing awareness. Cost-savings, resulting from reduced use of anaesthetic drugs may offset some of the additional cost of depth of anaesthesia monitoring. The cost of sensors attached to the patient appears to be a key factor in the additional cost of depth of anaesthesia monitoring.

Implications for service provision

The main implications for service provision will be the installation of the EEG module, any training required, and follow-up module maintenance. Module installation is unlikely to be particularly disruptive, although a separate compatible monitor may also be required, depending on which module is being introduced. As discussed earlier, training in use of the modules is not likely to be extensive.

Suggested research priorities

The following research recommendations are listed in order of perceived priority.

1. There is a lack of RCTs of E-Entropy-guided and Narcotrend-guided anaesthesia monitoring to detect explicit intraoperative awareness, specifically in high-risk patients. Given that incidence of awareness will be higher in this group, it may be more feasible to mount a trial than in the general surgical population, notwithstanding an adequate a priori statistical power calculation (although see below). Future trials should incorporate adequate length of follow-up to detect delayed cases of awareness. Cases of awareness may emerge after the first postoperative week, but in nearly all of the currently available RCTs of BIS, E-Entropy and Narcotrend, intraoperative awareness was assessed only within 1–3 days post surgery. It should be noted that in the RCTs we reviewed, the timing of follow-up was not always clearly specified and/or it was not clear to which outcomes the specified follow-up periods applied. Clear reporting of these crucial aspects of the RCT should be strongly encouraged. Future RCTs should also evaluate the effects of anaesthesia overdosing, including short-term effects such as nausea and vomiting, as well as longer-term impact on cognitive function.
2. There were no trials of the use of Narcotrend in children, and only two paediatric studies of E-Entropy in our systematic review. Future evaluation of these technologies would be warranted in these groups.
3. Further evidence on the incidence of intraoperative awareness is needed. The Royal College of Anaesthetists ran the National Audit Project (NAP) 5 to estimate the incidence of awareness in all UK hospitals (1 June 2012 until 31 May 2013). This may provide useful data for future economic modelling of depth of anaesthesia technologies in the UK.
4. Our literature searches identified three ongoing RCTs that would meet the inclusion criteria of our systematic review (see *Appendix 12*), all of which are investigating anaesthesia depth titrated

according to BIS values. A further recent RCT (accepted for publication), which is similar to our inclusion criteria, is the Michigan Awareness Control Study, comparing BIS-guided and MAC-guided electronic alerts for the prevention of awareness under GA.[140] The target sample size was 15,000 patients in each group (aged >18 years) at both low and high risk for awareness, and a total of 21,601 patients were enrolled at the time of interim analysis. The primary outcome measure was intraoperative awareness, with explicit recall measured at 28–30 days post anaesthesia. Modified ITT interim analysis found no statistically significant difference between BIS- and MAC-guided alerts in incidence of definite awareness, and the trial was therefore terminated because of futility. Post hoc power analysis showed that around 30,000 patients in each group would be required to detect a difference between the two interventions.[141] This calls into question the feasibility of future RCTs of depth of anaesthesia monitoring, particularly in the general surgical population.

Acknowledgements

We would like to thank members of our advisory group panel who provided expert advice and comments on the protocol and a draft of this report:

Professor Brian Pollard, Professor of Anaesthesia, Manchester Royal Infirmary, UK. (Professor Pollard has in the past received funding from Aspect Medical, the manufacturers of the BIS monitor, to undertake some research using the BIS monitor.)

Marta Soares, Research Fellow, Centre for Health Economics, University of York, UK.

Dr Gerlinde Mandersloot, Consultant in Intensive Care and Anaesthesia, The Royal London Hospital Whitechapel, UK (Representing ICU Steps – The Intensive Care Patient Support Charity).

We are also grateful to Karen Welch, Information Scientist, SHTAC, University of Southampton, for generating and running the literature searches, and Andy Clegg, Professor of Health Services Research, SHTAC, University of Southampton, for reviewing a draft of this report.

Contributions of authors

J Shepherd (Principal Research Fellow) developed the research protocol, assisted in the development of the search strategy, assessed studies for inclusion, extracted data from, and quality assessed, included studies, synthesised evidence, drafted and edited the final report, and project managed the study.

J Jones (Principal Research Fellow) developed the research protocol, assessed studies for inclusion, extracted data from, and quality assessed, included studies, synthesised evidence, developed the economic evaluation, and drafted the final report.

GK Frampton (Research Fellow) assessed studies for inclusion, extracted data from, and quality assessed, included studies, synthesised evidence, and drafted the final report.

L Baxter (Research Fellow) assessed studies for inclusion, extracted data from, and quality assessed, included studies, synthesised evidence, assisted in developing the economic evaluation and drafted the final report.

J Bryant (Principal Research Fellow) developed the research protocol, contributed to the background section, extracted data from, and quality assessed, included studies, synthesised evidence, drafted and edited the final report.

K Cooper (Senior Research Fellow) assessed studies for inclusion, extracted data from, and quality assessed, included studies, synthesised evidence, and developed the economic evaluation.

References

1. Jones JG, Konieczko K. Hearing and memory in anaesthetised patients. *Br Med J* 1986;**292**:1291. http://dx.doi.org/10.1136/bmj.292.6531.1291
2. Kerssens C, Klein J, Bonke B. Awareness: Monitoring versus remembering what happened. *Anesthesiology* 2003;**99**:570–5. http://dx.doi.org/10.1097/00000542-200309000-00011
3. Enlund M, Hassan HG. Intraoperative awareness: detected by the structured Brice interview? *Acta Anaesthesiol Scand* 2002;**46**:345–9. http://dx.doi.org/10.1034/j.1399-6576.2002.460402.x
4. Russell IF. Comparison of wakefulness with two anaesthetic regimens. *Br J Anaesth* 1986;**58**:965–8. http://dx.doi.org/10.1093/bja/58.9.965
5. Moerman N, Bonke B, Oosting J. Awareness and recall during general anaesthesia. *Anesthesiology* 1993;**79**:454–64. http://dx.doi.org/10.1097/00000542-199309000-00007
6. Ranta S, Laurila R, Saario J, Ali-Melkkila T, Hynynen M. Awareness with recall during general anaesthetia: incidence and risk factors. *Anesth Analg* 1998;**86**:1084–9.
7. Ghoneim MM, Block RI, Haffarnan M, Mathews MJ. Awareness during anaesthesia: risk factors, causes and sequelae: A review of reported cases in the literature. *Anesth Analg* 2009;**108**:527–35. http://dx.doi.org/10.1213/ane.0b013e318193c634
8. Leslie K, Chan MTV, Myles PS, Forbes A, McCulloch TJ. Posttraumatic stress disorder in aware patients from the B-Aware trial. *Anesth Analg* 2010;**110**:823–8. http://dx.doi.org/10.1213/ANE.0b013e3181b8b6ca
9. Sandin RH, Enlund G, Samuelsson P, Lennmarken C. Awareness during anaesthesia: a propsective case study. *Lancet* 2000;**355**:707–11. http://dx.doi.org/10.1016/S0140-6736(99)11010-9
10. Osterman JE, Hopper J, Heran WJ, Keane TM, van der Volk BA. Awareness under anaesthesia and the development of posttraumatic stress disorder. *Gen Hosp Psychiatry* 2001;**23**:198–204. http://dx.doi.org/10.1016/S0163-8343(01)00142-6
11. Lennmarken C, Bildfors K, Enlund G, Samuelsson P, Sandin R. Victims of awareness. *Acta Anaesthesiol Scand* 2002;**46**:229–31. http://dx.doi.org/10.1034/j.1399-6576.2002.t01-1-460301.x
12. Myles PS, Williams DL, Hendrata M, Anderson H, Weeks AM. Patient satisfaction after anaesthesia and surgery: results of a prospective survey of 10,811 patients. *Br J Anaesth* 2000;**84**:6–10. http://dx.doi.org/10.1093/oxfordjournals.bja.a013383
13. Sebel PS, Bowdle TA, Ghoneim MM, Rampil IJ, Padilla RE, Gan TJ, *et al.* The incidence of awareness during anaesthesia: a multicenter United States study. *Anesth Analg* 2004;**99**:833–9. http://dx.doi.org/10.1213/01.ANE.0000130261.90896.6C
14. Pollard RJ, Coyle JP, Gilbert RL, Beck JR. Intraoperative awarenes in a regional medical system. A review of 3 years' data. *Anesthesiology* 2007;**106**:269–74. http://dx.doi.org/10.1097/00000542-200702000-00014
15. Xu L, Wu A-S, Yue Y. The incidence of intra-operative awareness during general anaesthesia in China: a multi-center observational study. *Acta Anaesthesiol Scand* 2009;**53**:873–82. http://dx.doi.org/10.1111/j.1399-6576.2009.02016.x
16. Leslie K, Davidson AJ. Awareness during anesthesia: a problem without solutions? *Minerva Anestesiol* 2010;**76**:624–8.
17. Ronald A, Abdul Aziz KA, George Day T, Scott M. In patients undergoing cardiac surgery, thoracic epidural analgesia combined with general anaesthesia results in faster recovery and

fewer complications but does not affect length of hospital stay. *Interact Cardiovasc Thorac Surg* 2006;**5**:207–16. http://dx.doi.org/10.1510/icvts.2005.125054

18. Errando CL, Sigl JC, Robles M, Calabuig E, Garcia J, Arocas F, *et al.* Awareness with recall during general anaesthesia: a propsective observational evaluation of 4001 patients. *Br J Anaesth* 2008;**101**:178–85. http://dx.doi.org/10.1093/bja/aen144
19. Domino KB, Posner KL, Caplan RA, Cheney FW. Awareness during anaesthesia. A closed claims analysis. *Anesthesiology* 1999;**90**:1053–61. http://dx.doi.org/10.1097/00000542-199904000-00019
20. Hardman JG, Aitkenhead AR. Awareness during anesthesia. *Contin Educ Anaesth Crit Care Pain* 2005;**5**:183–6. http://dx.doi.org/10.1093/bjaceaccp/mki049
21. Thomas G, Morgan M. Monitoring neuromuscular blockade and depth of anaesthesia. *Anaesth Intensive Care Med* 2011;**12**:271–4. http://dx.doi.org/10.1016/j.mpaic.2011.03.008
22. Orser BA, Mazer D, Baker AJ. Awareness during anesthesia. *Can Med Assoc J* 2008;**178**:185–8. http://dx.doi.org/10.1503/cmaj.071761
23. Bruhn J, Myles PS, Sneyd R, Struys MMRF. Depth of anaesthesia monitoring: What's available, what's validated and what's next? *Br J Anaesth* 2006;**97**:85–94. http://dx.doi.org/10.1093/bja/ael120
24. Brice DD, Hetherington RR, Utting JE. A simple study of awareness and dreaming during anaesthesia. *Br J Anaesth* 1970;**42**:535–41. http://dx.doi.org/10.1093/bja/42.6.535
25. Wennervita J, Ranta SO, Hynynen M. Awareness and recall in outpatient anesthesia. *Anesth Analg* 2002;**95**:72–7. http://dx.doi.org/10.1097/00000539-200207000-00013
26. Avidan MS, Palanca BJ, Glick D, Jacobsohn E, Villafranca A, O'Connor M, *et al.* Protocol for the BAG-RECALL clinical trial: a prospective, multi-center, randomized, controlled trial to determine whether a bispectral index-guided protocol is superior to an anesthesia gas-guided protocol in reducing intraoperative awareness with explicit recall in high risk surgical patients. *BMC Anesthesiol* 2009;**9**:8TN.
27. Avidan MS, Zhang L, Burnside BA, Finkel KJ, Searleman AC, Selvidge JA, *et al.* Anesthesia awareness and the bispectral index. *N Engl J Med* 2008;**358**:1097–108. http://dx.doi.org/10.1056/NEJMoa0707361
28. Eger EI, Sonner JM. Response to Enlund M: TIVA, awareness and the Brice interview. *Anesth Analg* 2006;**102**:965. http://dx.doi.org/10.1213/01.ANE.0000190857.67309.1E
29. Lubke GH, Kerssens C, Gershon RY, Sebel PS. Memory formation during general anesthesia for emergency cesarean sections. *Anesthesiology* 2000;**92**:1029–34. http://dx.doi.org/10.1097/00000542-200004000-00020
30. Hudson AE, Hemmings HC. Are anaesthetics toxic to the brain? *Br J Anaesth* 2011;**107**:30–7. http://dx.doi.org/10.1093/bja/aer122
31. The Association of Anaesthetists of Great Britain and Ireland. *Recommendations for standards of monitoring during anaesthesia and recovery.* 4th edn. London: The Association of Anaesthetists of Great Britain and Ireland; 2007.
32. Bromhead HJ, Jones NA. The use of anaesthetic rooms for induction of anaesthesia: a postal survey of current practice and attitudes in Great Britain and Northern Ireland. *Anaesthesia* 2002;**57**:850–4. http://dx.doi.org/10.1046/j.1365-2044.2002.02779.x
33. British Medical Association and Royal Pharmaceutical Society of Great Britain. *British national formulary No. 62*, September 2011. London: BMA and RPS; 2011.

34. Punjasawadwong Y, Phongchiewboon A, Bunchungmongkol N. Bispectral index for improving anaesthetic delivery and postoperative recovery. *Cochrane Database Syst Rev* 2007;**4**:CD003843.

35. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, *et al.* The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;**343**:d5928. http://dx.doi.org/10.1136/bmj.d5928

36. Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 2005.

37. Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. Good practice guidelines for decision-analytic modelling in health technology assessment: A review and consolidation of quality assessment. *Pharmacoeconomics* 2006;**24**:355–71. http://dx.doi.org/10.2165/00019053-200624040-00006

38. National Institute for Health and Care Excellence (NICE). *Guide to the methods of technology appraisal*. 2008. 4 July 2011.

39. Treasury HM. The Green Book. *Appraisal and Evaluation in Central Government.* URL: www.hm-treasury.gov.uk/data_greenbook_index.htm. 2010. 1 April 2009.

40. Zhang C, Xu L, Ma Y-Q, Sun Y-X, Li Y-H, Zhang L, *et al.* Bispectral index monitoring prevent awareness during total intravenous anesthesia: A prospective, randomized, double-blinded, multi-center controlled trial. *Chin Med J* 2011;**124**:3664–9.

41. Sahni N, Anand LK, Gombar KK, Gombar S. Effect of intraoperative depth of anesthesia on postoperative pain and analgesic requirement: A randomized prospective observer blinded study. *J Anaesthesiol Clin Pharmacol* 2011;**27**:500–5. http://dx.doi.org/10.4103/0970-9185.86595

42. Sakaguchi M, Higuchi H, Maeda S, Miyawaki T. Dental sedation for patients with intellectual disability: A prospective study of manual control versus Bispectral Index-guided target-controlled infusion of propofol. *J Clin Anesth* 2011;**23**:636–42. http://dx.doi.org/10.1016/j.jclinane.2011.04.008

43. Hampshire P, Guha A, Welters I, Murphy J, Poole L. Bispectral index monitoring reduces sedative and vasopressor requirements during percutaneous tracheostomy. *Crit Care* 2011;**15**:S123–4. http://dx.doi.org/10.1186/cc9768

44. Avidan MS, Jacobsohn E, Glick D, Burnside BA, Zhang L, Villafranca A, *et al.* Prevention of intraoperative awareness in a high-risk surgical population. *New Engl J Med* 2011;**365**:591–600. http://dx.doi.org/10.1056/NEJMoa1100403

45. Bannister CF, Brosius KK, Sigl JC, Meyer BJ, Sebel PS. The effect of bispectral index monitoring on anesthetic use and recovery in children anesthetized with sevoflurane in nitrous oxide. *Anesth Analg* 2001;**92**:877–81. http://dx.doi.org/10.1097/00000539-200104000-00015

46. Bhardwaj N, Yaddanapudi S. A randomized trial of propofol consumption and recovery profile with BIS-guided anesthesia compared to standard practice in children. *Paediatr Anaesth* 2010;**20**:160–7. http://dx.doi.org/10.1111/j.1460-9592.2009.03240.x

47. Chan MTV, Cheng B, Gin T, Lee T, Koo E. Cognitive dysfunction after anesthesia: a randomized controlled trial. *J Neurosurg Anesthesiol* 2010;**22**:408–9.

48. Kamal NM, Omar SH, Radwan KG, Youssef A. Bispectral index monitoring tailors clinical anesthetic delivery and reduces anesthetic drug consumption. *J Med Sci* 2009;**9**:10–16. 48. Kamal NM, Omar SH, Radwan KG, Youssef A. Bispectral index monitoring tailors clinical anesthetic delivery and reduces anesthetic drug consumption. *J Med Sci* 2009;**9**:10–16. http://dx.doi.org/10.3923/jms.2009.10.16

49. Kerssens C, Gaither JR, Sebel PS. Preserved memory function during bispectral index-guided anesthesia with sevoflurane for major orthopedic surgery. *Anesthesiology* 2009;**111**:518–24. http://dx.doi.org/10.1097/ALN.0b013e3181b05f0b

50. Leslie K, Myles PS, Forbes A, Chan MT. The effect of bispectral index monitoring on long-term survival in the B-aware trial. *Anesth Analg* 2010;**110**:816–22. http://dx.doi.org/10.1213/ANE.0b013e3181c3bfb2

51. Liao WW, Wang JJ, Wu GJ, Kuo CD. The effect of cerebral monitoring on recovery after sevoflurane anesthesia in ambulatory setting in children: a comparison among bispectral index, A-line autoregressive index, and standard practice. *J Chin Med Assoc* 2011;**74**:28–36. http://dx.doi.org/10.1016/j.jcma.2011.01.004

52. Messieha ZS, Ananda RC, Hoffman WE, Punwani IC, Koenig HM. Bispectral Index System (BIS) monitoring reduces time to discharge in children requiring intramuscular sedation and general anesthesia for outpatient dental rehabilitation. *Pediatr Dent* 2004;**26**:256–60.

53. Messieha ZS, Ananda RC, Hoffman WE, Punwani IC, Koenig HM. Bispectral index system (BIS) monitoring reduces time to extubation and discharge in children requiring oral presedation and general anesthesia for outpatient dental rehabilitation. *Pediatric Dent* 2005;**27**:500–4.

54. Choi SR, Lim YH, Lee SC, Lee JH, Chung CJ. Spectral entropy monitoring allowed lower sevoflurane concentration and faster recovery in children. *Acta Anaesthesiol Scand* 2010;**54**:859–62. http://dx.doi.org/10.1111/j.1399-6576.2010.02212.x

55. Gruenewald M, Zhou J, Schloemerkemper N, Meybohm P, Weiler N, Tonner PH, *et al.* M-Entropy guidance vs standard practice during propofol–remifentanil anaesthesia: a randomised controlled trial. *Anaesthesia* 2007;**62**:1224–9. http://dx.doi.org/10.1111/j.1365-2044.2007.05252.x

56. Talawar P, Chhabra A, Trikha A, Arora MK, Chandralekha. Entropy monitoring decreases isoflurane concentration and recovery time in pediatric day care surgery: a randomized controlled trial. *Paediatr Anaesth* 2010;**20**:1105–10. http://dx.doi.org/10.1111/j.1460-9592.2010.03441.x

57. Vakkuri A, Yli-Hankala A, Sandin R, Mustola S, Hoymork S, Nyblom S, *et al.* Spectral entropy monitoring is associated with reduced propofol use and faster emergence in propofol-nitrous oxide-alfentanil anesthesia. *Anesthesiology* 2005;**103**:274–9. http://dx.doi.org/10.1097/00000542-200508000-00010

58. Wu SC, Wang PC, Liao WT, Shih TH, Chang KA, Lin KC, *et al.* Use of spectral entropy monitoring in reducing the quantity of sevoflurane as sole inhalational anesthetic and in decreasing the need for antihypertensive drugs in total knee replacement surgery. *Acta Anaesthesiol Taiwan* 2008;**46**:106–11. http://dx.doi.org/10.1016/S1875-4597(08)60003-X

59. Lai R-C, Lu Y-L, Huang W, Xu M-X, Lai J-L, Xie J-D, *et al.* Application of a narcotrend-assisted monitor for depth of anesthesia on microwave coagulation for the patients with liver cancer during total intravenous anesthesia with propofol and fentanyl. *Chin J Cancer* 2010;**29**:110–13. http://dx.doi.org/10.5732/cjc.009.10244

60. Rundshagen I, Hardt T, Cortina K, Pragst F, Fritzsche T, Spies C. Narcotrend-assisted propofol/remifentanil anaesthesia vs clinical practice: does it make a difference? *Br J Anaesth* 2007;**99**:686–93. http://dx.doi.org/10.1093/bja/aem231

61. Aime I, Verroust N, Masson-Lefoll C, Taylor G, Laloe PA, Liu N, *et al.* Does monitoring bispectral index or spectral entropy reduce sevoflurane use? *Anesth Analg* 2006;**103**:1469–77. http://dx.doi.org/10.1213/01.ane.0000246838.93153.23

62. Ellerkmann RK, Soehle M, Riese G, Zinserling J, Wirz S, Hoeft A, *et al.* The Entropy Module and Bispectral Index as guidance for propofol-remifentanil anaesthesia in combination with regional anaesthesia compared with a standard clinical practice group. *Anaesth Intensive Care* 2010;**38**:159–66.

63. Kreuer S, Biedler A, Larsen R, Altmann S, Wilhelm W. Narcotrend monitoring allows faster emergence and a reduction of drug consumption in propofol-remifentanil anesthesia. *Anesthesiology* 2003;**99**:34–41. http://dx.doi.org/10.1097/00000542-200307000-00009

64. Kreuer S, Bruhn J, Stracke C, Aniset L, Silomon M, Larsen R, *et al.* Narcotrend or bispectral index monitoring during desflurane-remifentanil anesthesia: a comparison with a standard practice protocol. *Anesth Analg* 2005;**101**:427–44. http://dx.doi.org/10.1213/01.ANE.0000157565.00359.E2

65. Ahmad S, Yilmaz M, Marcus RJ, Glisson S, Kinsella A. Impact of bispectral index monitoring on fast tracking of gynecologic patients undergoing laparoscopic surgery. *Anesthesiology* 2003;**98**:849–52. http://dx.doi.org/10.1097/00000542-200304000-00010

66. Anez C, Papaceit J, Sala JM, Fuentes A, Rull M. [The effect of encephalogram bispectral index monitoring during total intravenous anesthesia with propofol in outpatient surgery.] *Rev Esp Anestesiol Reanim* 2001;**48**:264–9.

67. Basar H, Ozcan S, Buyukkocak U, Akpinar S, Apan A. Effect of bispectral index monitoring on sevoflurane consumption. *Eur J Anaesthesiol* 2003;**20**:396–400. http://dx.doi.org/10.1097/00003643-200305000-00008

68. Boztug N, Bigat Z, Akyuz M, Demir S, Ertok E. Does using the bispectral index (BIS) during craniotomy affect the quality of recovery? *J Neurosurg Anesthesiol* 2006;**18**:1–4. http://dx.doi.org/10.1097/01.ana.0000188028.80960.dd

69. Bruhn J, Kreuer S, Bischoff P, Kessler P, Schmidt GN, Grzesiak A, *et al.* Bispectral index and A-line AAI index as guidance for desflurane-remifentanil anaesthesia compared with a standard practice group: a multicentre study. *Br J Anaesth* 2005;**94**:63–9. http://dx.doi.org/10.1093/bja/aei013

70. Chiu CL, Ong G, Majid AA. Impact of bispectral index monitoring on propofol administration in patients undergoing cardiopulmonary bypass. *Anaesth Intensive Care* 2007;**35**:342–7.

71. Gan TJ, Glass PS, Windsor A, Payne F, Rosow C, Sebel P, *et al.* Bispectral index monitoring allows faster emergence and improved recovery from propofol, alfentanil, and nitrous oxide anesthesia. BIS Utility Study Group. *Anesthesiology* 1997;**87**:808–15. http://dx.doi.org/10.1097/00000542-199710000-00014

72. Hachero A, Alamo F, Caba F, Echevarria M, Merino S, Gomez P, *et al.* [Influence of bispectral index monitoring on fentanyl requirements during total intravenous anesthesia for major gynecological surgery.] *Rev Esp Anestesiol Reanim* 2001;**48**:364–9.

73. Ibraheim O, Alshaer A, Mazen K, El-Dawlaty A, Turkistani A, Alkathery K, *et al.* Effect of bispectral index (BIS) monitoring on postoperative recovery and sevoflurane consumption among morbidly obese patients undergoing laparoscopic gastric banding. *Middle East J Anesthesiol* 2008;**19**:819–30.

74. Leslie K, Myles PS, Forbes A, Chan MT, Short TG, Swallow SK. Recovery from bispectral index-guided anaesthesia in a large randomized controlled trial of patients at high risk of awareness. *Anaesth Intensive Care* 2005;**33**:443–51.

75. Masuda T, Yamada H, Takada K, Sagata Y, Yamaguchi M, Tomiyama Y, *et al.* [Bispectral index monitoring is useful to reduce total amount of propofol and to obtain immediate recovery after propofol anesthesia.] *Masui* 2002;**51**:394–9.

76. Mayer J, Boldt J, Schellhaass A, Hiller B, Suttner SW. Bispectral index-guided general anesthesia in combination with thoracic epidural analgesia reduces recovery time in fast-track colon surgery. *Technol Comput Simulation* 2007;**104**:1145–9.

77. Morimoto Y, Oka S, Mii M, Shinjo Y, Yamashita A, Gohara T, *et al.* [Efficacy of bispectral index monitoring in improving anesthetic management, economics, and use of the operating theater.] *Masui* 2002;**51**:862–68.

78. Muralidhar K, Banakal S, Murthy K, Garg R, Rani GR, Dinesh R. Bispectral index-guided anaesthesia for off-pump coronary artery bypass grafting. *Ann Card Anaesth* 2008;**11**:105–10. http://dx.doi.org/10.4103/0971-9784.41578

79. Myles PS, Leslie K, McNeil J, Forbes A, Chan MT. Bispectral index monitoring to prevent awareness during anaesthesia: the B-Aware randomised controlled trial. *Lancet* 2004;**363**:1757–63. http://dx.doi.org/10.1016/S0140-6736(04)16300-9

80. Nelskyla KA, Yli-Hankala AM, Puro PH, Korttila KT. Sevoflurane titration using bispectral index decreases postoperative vomiting in phase II recovery after ambulatory surgery. *Anesth Analg* 2001;**93**:1165–9. http://dx.doi.org/10.1097/00000539-200111000-00021

81. Paventi S, Santevecchi A, Metta E, Annetta MG, Perilli V, Sollazzi L, *et al.* Bispectral index monitoring in sevoflurane and remifentanil anesthesia. Analysis of drugs management and immediate recovery. *Minerva Anestesiol* 2001;**67**:435–9.

82. Puri GD, Murthy SS. Bispectral index monitoring in patients undergoing cardiac surgery under cardiopulmonary bypass. *Eur J Anaesthesiol* 2003;**20**:451–6. http://dx.doi.org/10.1097/00003643-200306000-00004

83. Recart A, Gasanova I, White PF, Thomas T, Ogunnaike B, Hamza M, *et al.* The effect of cerebral monitoring on recovery after general anesthesia: a comparison of the auditory evoked potential and bispectral index devices with standard clinical practice. *Anesth Analg* 2003;**97**:1667–74. http://dx.doi.org/10.1213/01.ANE.0000087041.63034.8C

84. Song D, Joshi GP, White PF. Titration of volatile anesthetics using bispectral index facilitates recovery after ambulatory anesthesia. *Anesthesiology* 1997;**87**:842–8. http://dx.doi.org/10.1097/00000542-199710000-00018

85. Struys MM, De ST, Versichelen LF, Van D, V, Van den Broecke R, Mortier EP. Comparison of closed-loop controlled administration of propofol using Bispectral Index as the controlled variable versus 'standard practice' controlled administration. *Anesthesiology* 2001;**95**:6–17. http://dx.doi.org/10.1097/00000542-200107000-00007

86. Tufano R, Palomba R, Lambiase G, Giurleo LG. [The utility of bispectral index monitoring in general anesthesia.] *Minerva Anestesiol* 2000;**66**:389–93.

87. White PF, Ma H, Tang J, Wender RH, Sloninsky A, Kariger R. Does the use of electroencephalographic bispectral index or auditory evoked potential index monitoring facilitate recovery after desflurane anesthesia in the ambulatory setting? *Anesthesiology* 2004;**100**:811–17. http://dx.doi.org/10.1097/00000542-200404000-00010

88. Wong J, Song D, Blanshard H, Grady D, Chung F. Titration of isoflurane using BIS index improves early recovery of elderly patients undergoing orthopedic surgeries. *Can J Anaesth* 2002;**49**:13–18. http://dx.doi.org/10.1007/BF03020413

89. Zohar E, Luban I, White PF, Ramati E, Shabat S, Fredman B. Bispectral index monitoring does not improve early recovery of geriatric outpatients undergoing brief surgical procedures. *Can J Anaesth* 2006;**53**:20–5. http://dx.doi.org/10.1007/BF03021523

90. Assare H, Anderson RE, Jakobsson J. Sevoflurane requirements during ambulatory surgery: a clinical study of bispectral index and auditory evoked potential guided anaesthesia. *Ambul Surg* 2002;**9**:207–11. http://dx.doi.org/10.1016/S0966-6532(02)00004-5

91. Luginbuhl M, Wuthrich S, Petersen-Felix S, Zbinden AM, Schnider TW. Different benefit of bispectral index (BIS) in desflurane and propofol anesthesia. *Acta Anaesthesiol Scand* 2003;**47**:165–73. http://dx.doi.org/10.1034/j.1399-6576.2003.00041.x

92. McCormack J, Mehta D, Peiris K, Dumont G, Fung P, Lim J, *et al.* The effect of a target controlled infusion of propofol on predictability of recovery from anesthesia in children. *Paediatr Anaesth* 2010;**20**:56–62. http://dx.doi.org/10.1111/j.1460-9592.2009.03196.x

93. Wennervirta J, Hynynen M, Koivusalo AM, Uutela K, Huiku M, Vakkuri A. Surgical stress index as a measure of nociception/antinociception balance during general anesthesia. *Acta Anaesthesiol Scand* 2008;**52**:1038–45. http://dx.doi.org/10.1111/j.1399-6576.2008.01687.x

94. Ministry of Health and Long-Term Care. *Medical Advisory Secretariat. Bispectral index monitor*. 2004.

95. Choi BM, Shin DH, Noh MH, Kim YH, Jeong YB, Lee SH, *et al.* Temporal linear mode complexity as a surrogate measure of the effect of remifentanil on the central nervous system in healthy volunteers. *Br J Clin Pharmacol* 2011;**71**:871–85. http://dx.doi.org/10.1111/j.1365-2125.2011.03904.x

96. Nordström O, Engström AM, Persson S, Sandin R. Incidence of awareness in total i.v. anaesthesia based on propofol, alfentanil and neuromuscular blockade. *Acta Anaesthesiol Scand* 1997;**41**:978–84. http://dx.doi.org/10.1111/j.1399-6576.1997.tb04823.x

97. Abenstein JP. Is BIS monitoring cost-effective? *Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine & Biology Society 2009*; 7041–4.

98. Ekman A, Lindholm ML, Lennmarken C, Sandin R. Reduction in the incidence of awareness using BIS monitoring. *Acta Anaesthesiol Scand* 2004;**48**:20–6. http://dx.doi.org/10.1111/j.1399-6576.2004.00260.x

99. Elliott RA, Payne K, Moore JK, Davies LM, Harper NJN, St Leger AS, *et al.* Which anaesthetic agents are cost-effective in day surgery? Literature review, national survey of practice and randomised controlled trial. *Health Technol Assess* 2002;**6**(30).

100. Chernin EL. Pharmacoeconomics of Inhaled Anesthetic Agents: Considerations for the Pharmacist. *Am J Health Syst Pharm* 2004;**61**(Suppl. 4):18–22.

101. Dion P. The cost of anaesthetic vapours. *Can J Anaesth* 1992;**39**:633. http://dx.doi.org/10.1007/BF03008331

102. Chatterjee S, Rudra A, Sengupta S. Current concepts in the management of postoperative nausea and vomiting. *Anesthesiol Res Pract* 2011;**2011**:748031.

103. Ku CM, Ong BC. Postoperative nausea and vomiting: a review of current literature. *Singapore Med J* 2003;**44**:366–74.

104. Gan TJ. Postoperative nausea and vomiting – Can it be eliminated? *J Am Med Assoc* 2002;**287**:1233–6. http://dx.doi.org/10.1001/jama.287.10.1233

105. Liu SS. Effects of Bispectral Index Monitoring on Ambulatory Anesthesia. *Anesthesiology* 2004;**101**:311–15. http://dx.doi.org/10.1097/00000542-200408000-00010

106. Deiiner S, Silverstein JH. Postoperative delirium and cognitive function. *Br J Anaesth* 2009;**103**(Suppl. 1):i416.

107. Moller JT, Cluitmans P, Rasmussen LS, Canet J, Rabbitt P, Jolles J, *et al.* Long term postoperative cognitive dysfunction in the elderly: ISPOCD1 study. *Lancet* 2012;**351**:857–61. http://dx.doi.org/10.1016/S0140-6736(97)07382-0

108. Abildstrom H, Rasmussen LS, Rentowl P, Rasmussen H, Kristensen PA, Moller JT. Cognitive dysfunction 1–2 years afternon-cardiac surgery in the elderly. *Acta Anaesthesiol Scand* 2000;**44**:1246–51. http://dx.doi.org/10.1034/j.1399-6576.2000.441010.x

109. The Information Centre for Health and Social Care. *HES online – main procedures and intervention tables*. 2012. URL: www.hesonline.nhs.uk.

110. Jonsson L, Andreasen N, Kilander L, Soininen H, Waldemar G, Nygaard H, *et al.* Patient and proxy-reported utility in Alzheimer Disease using the EuroQol. *Alzheimer Dis Assoc Disord* 2006;**20**:49–55. http://dx.doi.org/10.1097/01.wad.0000201851.52707.c9

111. Lyons G, Macdonald R. Awareness during Caesarean section. *Anaesthesia* 1991;**46**:62–4. http://dx.doi.org/10.1111/j.1365-2044.1991.tb09321.x

112. Ranta S, Jussila J, Hynynen M. Recall of awareness during cardiac anaesthesia: influence of feedback information to the anaesthesiologist. *Acta Anaesthesiol Scand* 1996;**40**:554–60. http://dx.doi.org/10.1111/j.1399-6576.1996.tb04487.x

113. Liu WHD, Thorp TAS, Graham SG, Aitkenhead AR. Incidence of awareness with recall during general anaesthesia. *Anaesthesia* 1991;**46**:435–7. http://dx.doi.org/10.1111/j.1365-2044.1991.tb11677.x

114. Schwender D, Kunze-Kronawitter H, Dietrich P, Klasing S, Forst H, Madler C. Conscious awareness during general anaesthesia: patients' perceptions, emotions, cognition and reactions. *Br J Anaesth* 1998;**80**:133–9. http://dx.doi.org/10.1093/bja/80.2.133

115. Samuelsson P, Brudin L, Sandin RH. Late psychological symptoms after awareness among consecutively included surgical patients. *Anesthesiology* 2007;**106**:26–32. http://dx.doi.org/10.1097/00000542-200701000-00009

116. Evans JM. Patients' experience of awareness during general anaesthesia. In Rosen M, Lunn JN, editors. *Consciousness, awareness and pain in general anesthesia*. London: Butterworth; 1987. pp. 184–92.

117. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*-Fourth Edition. Washington, DC: American Psychiatric Association; 1994.

118. Kessler RC, Sonnega A, Bromet E, Hughes M, Neslon CB. Posttraumatic stress disorder in the national comorbidity survey. *Arch Gen Psychiatry* 1995;**52**:1048–60. http://dx.doi.org/10.1001/archpsyc.1995.03950240066012

119. Doctor JN, Zoellner LA, Feeny NC. Predictors of health-related quality-of-life utilities among persons with posttraumatic stress disorder. *Psychiatr Serv* 2011;**62**:272–7. http://dx.doi.org/10.1176/appi.ps.62.3.272

120. Freed MC, Yeager DE, Liu X, Gore KL, Engel CC, Magruder KM. Preference-weighted health status of PTSD among veterans: an outcome for cost-effectiveness analysis using clinical data. *Psychiatr Serv* 2009;**60**:1230–8. http://dx.doi.org/10.1176/appi.ps.60.9.1230

121. Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002;**21**:271–92. http://dx.doi.org/10.1016/S0167-6296(01)00130-8

122. Laffaye C, Kennedy C, Stein MB. Post-traumatic stress disorder and health-related quality of life in female victims of intimate partner violence. *Violence Vict* 2003;**18**:227–38. http://dx.doi.org/10.1891/vivi.2003.18.2.227

123. Tsai KY, Chou P, Chou FH, Su TT, Lin SC, Lu MK, *et al.* Three-year follow-up study of the relationship between posttraumatic stress symptoms and quality of life among earthquake survivors in Yu-Chi, Taiwan. *J Psychiatr Res* 2007;**41**:90–9. http://dx.doi.org/10.1016/j.jpsychires.2005.10.004

124. Meeske KA, Ruccione K, Globe DR, Stuber ML. Posttraumatic stress, quality of life, and psychological distress in young adult survivors of childhood cancer. *Oncol Nurs Forum* 2001;**28**:481–9.

125. Hernandez-Gancedo C, Pestana D, Perez-Chrzanowska H, Martinez-Casanova E, Criado A. Comparing Entropy and the Bispectral index with the Ramsay score in sedated ICU patients. *J Clin Monit Comput* 2007;**21**:295–302. http://dx.doi.org/10.1007/s10877-007-9087-7

126. Weil G, Passot S, Servin F, Billard V. Does spectral entropy reflect the response to intubation or incision during propofol-remifentanil anesthesia? *Anesth Analg* 2008;**106**:152–9. http://dx.doi.org/10.1213/01.ane.0000296454.00236.fc

127. Magruder KM, Frueh B, Knapp R, *et al.* Prevalence of posttraumatic stress disorder in Veteran's Affairs primary care clinics. *Gen Hosp Psychiatry* 2005;**27**:169–79. http://dx.doi.org/10.1016/j.genhosppsych.2004.11.001

128. Ara R, Brazier J. Predicting the short form-6D preference-based index using the eight mean short form-36 health dimension scores:estimating preference-based health-related utilities when patient level data are not available. *Value Health* 2009;**12**:346–53. http://dx.doi.org/10.1111/j.1524-4733.2008.00428.x

129. Berger W, Figueira I, Maurat AM, Bucassio EP, Vieira I, Jardim SR, *et al.* Partial and full PTSD in Brazilian ambulance workers: prevalence and impact on health and on quality of life. *J Trauma Stress* 2007;**20**:637–42. http://dx.doi.org/10.1002/jts.20242

130. Shiner B, Watts BV, Pomerantz A, Young-Xu Y, Schnurr PP. Sensitivity of the SF-36 to PTSD symptom change in veterans. *J Trauma Stress* 2011;**24**:111–15. http://dx.doi.org/10.1002/jts.20613

131. Evren C, Sar V, Dalbudak E, Cetin R, Durkaya M, Evren B, *et al.* Lifetime PTSD and quality of life among alcohol-dependent men: impact of childhood emotional abuse and dissociation. *Psychiatry Res* 2011;**186**:85–90. http://dx.doi.org/10.1016/j.psychres.2010.07.004

132. SCHARR. *SF-6D*. 2012. URL: www.shef.ac.uk/scharr/sections/heds/mvh/sf-6d.

133. National Institute for Health and Care Excellence (NICE). *Post-traumatic stress disorder (PTSD). The management of PTSD in adults and children in primary and secondary care. National costing report*. 2005. URL: http://guidance.nice.org.uk/CG26/CostingReport/pdf/English (accessed 1 February 2012).

134. National Institute for Health and Care Excellence (NICE). *Post-traumatic stress disorder (PTSD). The management of PTSD in adults and children in primary and secondary care. Clinical Guideline 26*. 2005. URL: www.nice.org.uk/CG26 (accessed 1 February 2012).

135. Singleton N, Bumpstead R, O'Brien M, Lee A, Meltzer H. *Psychiatric morbidity among adults living in private households, 2000*. London: HMSO; 2001. URL: www.ons.gov.uk/ons/index.html (accessed 1 February 2012).

136. Kartha A, Brower V, Saitz R, Samet JH, Keane TM, Liebschutz J. The impact of trauma exposure and post-traumatic stress disorder on health care utilization among primary care patients. *Med Care* 2008;**46**:388–93. http://dx.doi.org/10.1097/MLR.0b013e31815dc5d2

137. Department of Health. *NHS Reference Costs 2010–2011*. 2012. URL: www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_131140 (accessed 17 May 2013).

138. Phillips AA, McLean RF, Devitt JH, Harrington EM. Recall of intraoperative events after general anaesthesia and cardiopulmonary bypass. *Can J Anaesth* 1993;**40**:922–6. http://dx.doi.org/10.1007/BF03010093

139. Russell IF. The Narcotrend 'depth of anaesthesia' monitor cannot reliably detect consciousness during general anaesthesia: an investigation using the isolated forearm technique. *Br J Anaesth* 2006;**96**:346–52. http://dx.doi.org/10.1093/bja/ael017

140. Mashour GA, Tremper KK, Avidan MS. Protocol for the 'Michigan Awareness Control Study': A prospective, randomized, controlled trial comparing electronic alerts based on bispectral index monitoring or minimum alveolar concentration for the prevention of intraoperative awareness. *BMC Anesthesiol* 2009;**9**:7. http://dx.doi.org/10.1186/1471-2253-9-7

141. Mashour GA, Shanks A, Tremper K, Kheterpal S, Turner C, Ramachandran S, *et al.* Prevention of intraoperative awareness with explicit recall in an unselected surgical population: a randomized comparative effectiveness trial. *Anesthesiology* 2012;**117**:717–25.

142. Medical Advisory Secretariat Bispectral index monitor: an evidence-based analysis. *ONT Health Technol Assess Ser* 2004;**4**:1–72.

143. El Menesy TM, Hamden GA, Amin SM, Zakaria D. Continuous intraoperative high state and response entropies reduced desflurane emergence agitation in pediatric laparoscopic surgery. *Egypt J Anaesth* 2008;**24**:85–93.

144. Pelletier N, Passot S, Zufferey P, Molliex S. Benefit of EEG guided anaesthesia: a meta-analysis. *Anesthesiology* 2011;**115**.

145. Smajic J, Praso M, Hodzic M, Hodzic S, Srabovic-Okanovic A, Smajic N, *et al.* Assessment of depth of anesthesia: PRST score versus bispectral index. *Medicinski Arhiv* 2011;**65**:216–20. http://dx.doi.org/10.5455/medarh.2011.65.216-220

146. Bauer M, Wilhelm W, Kraemer T, Kreuer S, Brandt A, Adams HA, *et al.* Impact of bispectral index monitoring on stress response and propofol consumption in patients undergoing coronary artery bypass surgery. *Anesthesiology* 2004;**101**:1096–104. http://dx.doi.org/10.1097/00000542-200411000-00008

147. Riad W, Schreiber M, Saeed AB. Monitoring with EEG entropy decreases propofol requirement and maintains cardiovascular stability during induction of anaesthesia in elderly patients. *Eur J Anaesthesiol* 2007;**24**:684–8. http://dx.doi.org/10.1017/S026502150700018X

148. Singh M, Kulkarni D, Gopinath R. Bispectral index monitoring may not enhance performance of the 'wake-up' test during scoliosis surgery. *J Anaesthesiol Clin Pharmacol* 2009;**25**:179–82.

149. Weber F, Pohl F, Hollnberger H, Taeger K. Impact of the Narcotrend Index on propofol consumption and emergence times during total intravenous anaesthesia with propofol and remifentanil in children: a clinical utility study. *Eur J Anaesthesiol* 2005;**22**:741–7. http://dx.doi.org/10.1017/S0265021505001237

150. Tappenden P, Jones R, Paisley S, Carroll C. Systematic review and economic evaluation of bevacizumab and cetuximab for the treatment of metastatic colorectal cancer. *Health Technol Assess* 2007;**11**(12).

Appendix 1 Report methods for synthesis of evidence of clinical effectiveness and cost-effectiveness as described in the research protocol

Report methods for assessing the outcomes arising from the use of the interventions

The systematic review of clinical effectiveness will adhere to standard methodology as outlined in the CRD guidance for undertaking reviews in health care.

Population

The relevant study population for this assessment is patients receiving GA for surgery, including adults and children in whom the technology is licensed. Elderly and obese patients undergoing GA will be included as sub-groups for this evaluation where data allow.

Studies of patients receiving sedation in settings such as intensive care or high-dependency units are not relevant to this assessment. Studies of anaesthesia monitoring in healthy volunteers, or in non-surgical anaesthesia will not be included. Studies in which only regional or local anaesthesia are given will not be included.

Interventions

- E-Entropy.
- BIS.
- Narcotrend.

Comparators

The comparator in this assessment is standard clinical observation, including one or more of the following clinical markers: end-tidal anaesthetic gas concentrations (for inhaled anaesthesia); pulse measurement; heart rhythm; blood pressure; lacrimation; and sweating.

Outcomes

Studies will be included if they report one or more of the following outcomes:

- probability of intraoperative awareness
- patient distress and other sequelae resulting from intraoperative awareness
- recovery status (e.g. Aldrete scoring system)
- time to emergence from anaesthesia
- time to extubation (if appropriate)
- time to discharge from the recovery room
- consumption of anaesthetic agents
- morbidity and mortality including postoperative cognitive dysfunction from anaesthetic agents, pain-relieving drugs, antibiotics, antisickness drugs and muscle relaxants
- HRQoL.

Data on these indirect outcomes are likely to be used to estimate QALYs as final health outcomes.

Study design

We will prioritise RCTs for inclusion in the systematic review of clinical effectiveness. Where RCTs of technologies are not identified we will consider non-RCTs and controlled observational studies for inclusion, providing they include relevant outcomes.

Systematic reviews will be retrieved only to check their reference lists for potentially relevant studies. However, to ensure the workload is manageable within available time and resources we may include the aforementioned Cochrane systematic review of BIS which included 31 RCTs (Punjasawadwong and colleagues[34]). The Cochrane review had similar inclusion criteria to the current review and was last updated in May 2009. Rather than search for and review all studies of BIS, it is proposed that we summarise the findings of the Cochrane review and supplement it by reviewing any relevant studies published since May 2009.

Search strategy

A comprehensive search strategy will be devised, tested and applied to a number of electronic databases by an experienced Information Scientist (see *Appendix 1* for the MEDLINE strategy). Electronic databases to be searched include: MEDLINE (Ovid); MEDLINE In-Process & Other Non-Indexed Citations (Ovid); EMBASE (Ovid); The Cochrane Library (CDSR; CENTRAL); DARE; HTA; NHS Economic Evaluation Database (NHS EED); and EconLit.

Databases will be searched from 1995 to the present day (for BIS the search will be from May 2009 to the present day, supplementing the Cochrane systematic review). In addition, contact will be made with experts in the field to identify any relevant studies. Reference lists of included studies will be checked for any potentially relevant studies. Research in progress will be identified from the following databases: Current Controlled Trials; ClinicalTrials.gov; NIHR-Clinical Research Network Portfolio; WHO ICTRP (International Clinical Trials Registry Platform).

Studies published in the last two years as abstracts or conference proceedings will be included only if sufficient details are presented to allow appraisal of the methodology and the assessment of results to be undertaken.

Only articles published in the English language will be included.

For the cost-effectiveness assessment, searches for other evidence to inform cost-effectiveness modelling will be conducted as required and may include a wider range of study types.

The titles and abstracts of studies identified by the search strategy will be assessed for potential eligibility using the inclusion/exclusion criteria detailed above. Full papers of studies that appear potentially relevant will be requested for further assessment. These will be screened by one reviewer and checked by a second, and a final decision regarding inclusion will be agreed. Any disagreements will be resolved by discussion, with involvement of a third reviewer where necessary.

Data extraction strategy

All included studies will undergo data extraction using a structured piloted template. Each study will be extracted by one reviewer and checked by a second for accuracy. Any disagreements between reviewers will be resolved by consensus or if necessary by arbitration by a third reviewer.

Quality assessment strategy

The methodological quality of all included studies will be appraised by one reviewer, and checked by a second. Any disagreements between reviewers will be resolved by consensus or if necessary by arbitration by a third reviewer.

RCTs will be appraised using the Cochrane Collaboration Risk of Bias criteria. Any non-randomised and observational studies included will be appraised using criteria developed by Spitzer and colleagues (1990).

Methods of analysis/synthesis

Studies will be synthesised through a narrative review with tabulation of results of included studies. Quantitative synthesis of results will be contingent on the data available. Meta-analysis using Cochrane Review Manager (REVMAN) software will be considered where appropriate (e.g. if there are several high quality studies of the same design) and sources of heterogeneity will be investigated.

Report methods for synthesising evidence of cost-effectiveness

Review of published cost-effectiveness studies

The methods detailed above will be used to systematically review the cost-effectiveness literature. The inclusion and exclusion criteria are similar to that of the systematic review of clinical effectiveness, with the exception of study design and outcomes. Studies will be included if they are full economic evaluations, assessing both costs and consequences, of the specified technologies (e.g. reporting cost per patient, cost per episode of intraoperative awareness or cost per QALY). The quality of the included economic evaluations will be assessed using a critical appraisal checklist based upon that proposed by Drummond and colleagues (2005) and Philips and colleagues (2006). The data from these studies will be tabulated and discussed in a narrative review.

Where presented, HRQoL data will be extracted from studies included in both the systematic review of clinical effectiveness and the systematic review of cost-effectiveness. In addition, a targeted literature search will be conducted specifically for publications reporting HRQoL or health state utility for adults with episodes of intraoperative awareness. Where available, QoL data will be used in our economic model.

Evaluation of costs and cost-effectiveness

A comparison of the costs and consequences of depth of anaesthesia monitoring will be made using decision-analytic models. The structure of the models will be informed by the systematic review of cost-effectiveness and other systematic searches of the literature and, where necessary, using guidelines and expert opinion. The model will be constructed according to standard modelling guidelines (Phillips and colleagues (2006) and a full explanation of our methods for formulating model structure and deriving parameter values will be given in the assessment report. The perspective will be that of the NHS and Personal Social Services (PSS). The outcome will be reported as cost per patient, cost per intraoperative awareness avoided and cost per quality-adjusted life-year (QALY) gained, where possible.

The decision tree model will include the costs of the anaesthesia-monitoring device (including the module, the sensors, and, if applicable, the monitors), and any savings associated with reduced use of anaesthesia, fewer side effects and improved recovery time from the anaesthesia. We will aim to assess the HRQoL impact of episodes of intraoperative awareness. If good HRQoL data are available the model will include health benefits in terms of QALYs. In the case where insufficient published HRQoL data are available it will be necessary to elicit HRQoL values from clinical experts or to conduct threshold analyses using a range of estimates. The time horizon will be a patient's lifetime (or shorter if appropriate) in order to reflect long-term health gains. Both costs and benefits will be discounted at 3.5%.

Parameter values will be obtained from the relevant research literature, including our own systematic review of clinical and cost-effectiveness. Sources for parameters will be stated clearly. Resource use will be specified and valued from the perspective of the NHS and PSS. Costs will be derived from primary data from previous studies, and national and local NHS unit costs. If insufficient data are retrieved from published sources, costs may be obtained from individual NHS Trusts or groups of Trusts.

Uncertainty will be explored through both one-way sensitivity analyses and scenario analyses. A probabilistic sensitivity analysis probabilistic sensitivity analysis will be undertaken if both the data and modelling approach permit this. The outputs of any probabilistic sensitivity analysis will be presented using plots of the cost-effectiveness plane and cost-effectiveness acceptability curves.

The model will be validated by checking the model structure, calculations and data inputs for technical correctness. The structure will be reviewed by clinical experts for appropriateness for the clinical and diagnostic pathways. The robustness of the model to changes in input values will be tested using sensitivity analyses.

References

Drummond M, Sculpher M, Torrance G, O'Brien B, Stoddart G. *Methods for the economic evaluation of health care programmes*. Oxford: Oxford University Press; 2005.

Philips Z, Bojke L, Sculpher M, *et al*. Good practice guidelines for decision-analytic modelling in health technology assessment: A review and consolidation of quality assessment. *Pharmacoeconomics* 2006; **244**, 355–71.

Punjasawadwong Y, Phongchiewboon A, Bunchungmongkol N. Bispectral index for improving anaesthetic delivery and postoperative recovery. *Cochrane Database Syst Rev* 2007; **4**: CD003843.

Spitzer W, Lawrence V, Dales R, Hill G, Archer M, Clarck P, *et al*. Links between passive smoking and disease: a best evidence synthesis. *Clin Invest Med* 1990; **13**, 17–42.

Appendix 2 Literature search strategies

MEDLINE search strategy for Bispectral Index, Narcotrend and E-Entropy used in systematic review of patient outcomes

1. ("E-Entropy" or "M-Entropy" or Narcotrend).mp.
2. (entropy adj5 (module* or technolog* or system* or monitor* or machine*)).tw.
3. (entropy adj2 (state or response or spectral)).tw.
4. 2 or 3
5. 1 or 4
6. monitoring intraoperative/
7. consciousness monitors/
8. ("automated responsiveness" and (monitor* or measur* or machine*)).tw.
9. sedation monitor*.tw.
10. sedation measurement*.tw.
11. exp Anesthesia, General/
12. exp Anesthetics, General/
13. (an?esthetic* or an?esthesia or an?esthetist*).tw.
14. Intraoperative Period/
15. Anesthesia, Intravenous/
16. Anesthetics, Inhalation/
17. Anesthesiology/
18. exp Infusions, Intravenous/
19. Surgical Procedures, Operative/
20. General Surgery/
21. (surgery or surgical).tw.
22. Perioperative Period/
23. Signal Processing, Computer-Assisted/
24. Intraoperative Complications/
25. Perioperative Care/
26. Monitoring, Physiologic/
27. Adjuvants, Anesthesia/
28. Electromyography/
29. exp Electroencephalography/
30. Mental Recall/
31. Wakefulness/
32. Consciousness/
33. Perception/
34. Intraoperative Awareness/ or Awareness/
35. Arousal/
36. Deep Sedation/
37. Conscious Sedation/
38. Drug Therapy, Computer-Assisted/
39. Pain Measurement/
40. cerebral cortex/de
41. Evoked Potentials/ or Evoked Potentials Auditory/
42. Signal Processing, Computer-Assisted/
43. (surgery or surgical or operating or operation*1).tw.
44. (intraoperative* or "intra-operative*" or "intra operative*").tw.

45. (perioperative* or "peri-operative*" or "peri operative*").tw.
46. "depth of anaesthesia monitor*".tw.
47. "depth of anesthesia monitor*".tw.
48. "Anesthesia and Analgesia"/
49. Postoperative Period/
50. (postoperative or post?operative).tw.
51. (recall* or aware* or memory or memories or wake* or awake* or arous* or cry* or sweat* or tear*1 or dream* or remember* or movement* or grimac*).tw.
52. EEG or EMG or FEMG or encephalogra* or electroencephalogra* or electromyogra*).tw.
53. Brice.tw.
54. or/6-53
55. 5 and 54
56. limit 55 to (english language and yr="1995 -Current")
57. animals/
58. 56 not 57
59. (letter or comment or editorial).pt.
60. 58 not 59
61. crystal*.tw.
62. 60 not 61
63. coma/ or coma.tw.
64. 62 not 63
65. (("bispectral Index" or "bi-spectral index" or "bi spectral index") adj5 (module* or technolog* or system* or monitor* or machine*)).mp.
66. ((BIS or BISx) adj5 (module* or technolog* or system* or monitor* or machine*)).mp.
67. (anesth* adj20 (BIS or BISx)).tw.
68. (anaesth* adj20 (BIS or BISx)).tw.
69. or/65-68
70. "behavio?ral inhibition system".tw.
71. 69 not 70
72. ((surg* adj20 "BIS") or "BISx").tw.
73. 71 or 72
74. 54 and 73
75. limit 74 to (English language and humans and yr="2009 - 2011")
76. 75 not 59
77. 76 not 64
78. Anesthesia, Local/
79. (local adj1 an?esth*).tw.
80. 78 or 79
81. 77 not 80

NB. Search for BIS studies was performed separately from Narcotrend and E-Entropy, hence the inclusion of BIS terms at the end of the strategy (from line 65 onwards).

MEDLINE search strategy for Bispectral Index, Narcotrend and E-Entropy used in systematic review of cost-effectiveness

1. ("E-Entropy" or "M-Entropy" or Narcotrend).mp. (73)
2. (entropy adj5 (module* or technolog* or system* or monitor* or machine*)).tw. (380)
3. (entropy adj2 (state or response or spectral)).tw. (300)
4. 2 or 3 (604)
5. 1 or 4 (662)

6. monitoring intraoperative/ (13,101)
7. consciousness monitors/ (119)
8. ("automated responsiveness" and (monitor* or measur* or machine*)).tw. (4)
9. sedation monitor*.tw. (46)
10. sedation measurement*.tw. (6)
11. exp Anesthesia, General/ (46,626)
12. exp Anesthetics, General/ (98,559)
13. (an?esthetic* or an?esthesia or an?esthetist*).tw. (184,909)
14. Intraoperative Period/ (11,282)
15. Anesthesia, Intravenous/ (9798)
16. Anesthetics, Inhalation/ (9572)
17. Anesthesiology/ (15,249)
18. exp Infusions, Intravenous/ (44,602)
19. Surgical Procedures, Operative/ (48,143)
20. General Surgery/ (31,636)
21. (surgery or surgical).tw. (1,018,003)
22. Perioperative Period/ (254)
23. Signal Processing, Computer-Assisted/ (30,306)
24. Intraoperative Complications/ (23,721)
25. Perioperative Care/ (6700)
26. Monitoring, Physiologic/ (41,597)
27. Adjuvants, Anesthesia/ (2653)
28. Electromyography/ (62,736)
29. exp Electroencephalography/ (113,311)
30. Mental Recall/ (25,043)
31. Wakefulness/ (13,087)
32. Consciousness/ (8829)
33. Perception/ (17,362)
34. Intraoperative Awareness/ or Awareness/ (12,290)
35. Arousal/ (26,845)
36. Deep Sedation/ (309)
37. Conscious Sedation/ (5918)
38. Drug Therapy, Computer-Assisted/ (1263)
39. Pain Measurement/ (50,337)
40. cerebral cortex/de (15,104)
41. Evoked Potentials/ or Evoked Potentials Auditory/ (57,136)
42. Signal Processing, Computer-Assisted/ (30,306)
43. (surgery or surgical or operating or operation*1).tw. (1,240,833)
44. (intraoperative* or "intra-operative*" or "intra operative*").tw. (73,745)
45. (perioperative* or "peri-operative*" or "peri operative*").tw. (45,446)
46. "depth of anaesthesia monitor*".tw. (39)
47. "depth of anesthesia monitor*".tw. (31)
48. "Anesthesia and Analgesia"/ (3320)
49. Postoperative Period/ (30,192)
50. (postoperative or post?operative).tw. (257,047)
51. (recall* or aware* or memory or memories or wake* or awake* or arous* or cry* or sweat* or tear*1 or dream* or remember* or movement* or grimac*).tw. (767,912)
52. (EEG or EMG or FEMG or encephalogra* or electroencephalogra* or electromyogra*).tw. (103,627)
53. Brice.tw. (18)
54. or/6-53 (2,633,781)
55. 5 and 54 (326)
56. limit 55 to (English language and yr="1995 -Current") (277)

57. animals/ (4,924,118)
58. 56 not 57 (259)
59. (letter or comment or editorial).pt. (1,097,745)
60. 58 not 59 (240)
61. crystal*.tw. (146,923)
62. 60 not 61 (229)
63. coma/ or coma.tw. (25,605)
64. 62 not 63 (228)
65. exp economics/ (449,064)
66. exp economics hospital/ (17,691)
67. exp economics pharmaceutical/ (2299)
68. exp economics nursing/ (3854)
69. exp economics medical/ (13,581)
70. exp "Costs and Cost Analysis"/ (161,041)
71. Cost Benefit Analysis/ (52,655)
72. exp models economic/ (8329)
73. exp fees/ and charges/ (7794)
74. exp budgets/ (11,145)
75. (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic*).tw. (350,335)
76. (value adj1 money).tw. (20)
77. budget$.tw. (14,911)
78. or/65-77 (681,466)
79. ((energy or oxygen) adj cost).tw. (2386)
80. (metabolic adj cost).tw. (626)
81. ((energy or oxygen) adj expenditure).tw. (13,708)
82. or/79-81 (16,090)
83. 78 not 82 (677,823)
84. (letter or editorial or comment or historical article).pt. (1,367,063)
85. 83 not 84 (624,009)
86. 64 and 85 (2)
87. 1 and 85 (0)
88. 5 and 11 and 85 (1)
89. 86 or 88 (3)
90. (entropy and device*).tw. (80)
91. 85 and 90 (7)
92. 89 or 91 (9)
93. (entropy and surg*).tw. (167)
94. 85 and 93 (6)
95. 92 or 94 (11)
96. from 95 keep 3,5,8,10 (4)
97. ("depth of an?esth*" and cost).tw. (23)
98. 97 not 96 (22)
99. ((BIS or BISx) adj5 (module* or technolog* or system* or monitor* or machine*)).mp. (951)
100. (("bispectral Index" or "bi-spectral index" or "bi spectral index") adj5 (module* or technolog* or system* or monitor* or machine*)).mp. (533)
101. (anesth* adj20 (BIS or BISx)).tw. (584)
102. (anaesth* adj20 (BIS or BISx)).tw. (278)
103. (surg* adj20 (BIS or BISx)).tw. (424)
104. or/99-103 (1768)
105. 85 and 104 (68)
106. limit 105 to yr="2009 -Current" (9)

107. 96 or 98 or 106 (35)
108. 105 NOT 107 (51)
109. limit 108 to yr="1995 -Current" (50)

NB. Search for BIS studies was performed separately from Narcotrend and E-Entropy; hence, the inclusion of BIS terms at the end of the strategy (from line 99 onwards).

Appendix 3 Inclusion/exclusion worksheet used in systematic review of patient outcomes

Study name or number:			
Population:	Yes	Unclear	No
Adults and children aged > 2 years receiving general anaesthesia for surgery.	↓	↓	→
Not included:	next question	next question	EXCLUDE1
Patients receiving sedation in settings such as intensive care or high-dependency units;			
Healthy volunteers, or non-surgical anaesthesia (e.g. diagnostic investigations);[a]			
Patients receiving only regional or local anaesthesia.			
Technology:	Yes	Unclear	No
Any of the following:	↓	↓	→
E-Entropy[b]	next question	next question	EXCLUDE2
BIS			
Narcotrend			
Comparators:			
Standard clinical observation,[c] including one or more of the following markers:			
end-tidal anaesthetic gas concentrations/MAC (for inhaled anaesthesia)			
heart rhythm			
blood pressure			
oxygen levels (pulse oximeter)			
lacrimation			
sweating			
Outcomes:	Yes	Unclear	No
One or more of the following:	↓	↓	→
Probability of intraoperative awareness	next question	next question	EXCLUDE3
Patient distress and sequelae resulting from intraoperative awareness			
Recovery status (e.g. Aldrete scoring system)			
Time to emergence from anaesthesia			
Time to extubation			
Time to discharge from the recovery room			
Consumption of anaesthetic agents			
Morbidity and mortality including postoperative cognitive dysfunction from anaesthetic agents, use of pain-relieving drugs, use of antibiotics, use of antisickness drugs and muscle relaxants.			
HRQoL			
Study design:	Yes	Unclear	No
RCT; quasi-randomised or non-RCT; controlled before and after study[d]	↓	↓	→
Systematic reviews to be retrieved for reference checking only	next question	next question	EXCLUDE4
Conference abstracts prior to 2010 not for inclusion			
English language only			
Final decision	**INCLUDE**	**UNCLEAR (Discuss)**	**EXCLUDE**

a In some cases diagnostic instruments can also be used surgically to treat a condition (e.g. endoscopy). If it is unclear whether or not such an instrument has been used for treatment retrieve the paper for further inspection.

b Also includes M-Entropy.

c Studies may use a variety of terms to describe this including 'conventional clinical variables', 'standard practice', 'clinical assessment', 'and haemodynamic parameters'. They may not always define which markers they assessed in which case retrieve the paper for further inspection.

d Once screening on title/abstract is complete, only include non-RCT for a technology if no RCT have already been identified.

Appendix 4 Reasons for the exclusion of full-text publications from systematic review of patient outcomes

Of the 31 full-text publications that were screened against the systematic review eligibility criteria, 10 were excluded for the following reasons.

Exclusion criterion = study design (five publications)

Not primary research (two studies):

- Punjasawadwong *et al.*[34] – a Cochrane review comparing BIS against standard practice.
- Anon[142] – a systematic review comparing BIS against standard practice, but pre-dating the Cochrane review by Punjasawadwong *et al.*[34]

Primary research other than RCTs (three studies):

- El Menesy *et al.*[143]
- Pelletier *et al.*[144]
- Smajic *et al.*[145]

Exclusion criterion = comparator (standard practice unclear or not defined) (four publications)

- Bauer *et al.*[146]
- Riad *et al.*[147]
- Singh *et al.*[148]
- Weber *et al.*[149]

Publication retracted by journal (one publication)

- Mayer *et al.*[76]

Appendix 5 Data extraction and critical appraisal forms used in the systematic review of patient outcomes

Aime *et al.*

Reviewer 1: JS	Reviewer 2: GF		
Reference and design	***Technology***	***Participants***	***Outcome measures***
Author: Aime *et al.*[61] **Year:** 2006 **Study design:** RCT **Number of centres:** one **Country:** France **Sponsor:** GE Healthcare Monitoring Solutions loaned the authors a S5 monitor and provided the probes. No other funding source reported **Trial name:** NR	**Group 1:** BIS (Version 4.0 XP, Aspect Medical Systems), using Datex-Ohmeda S/5™ monitor Target device/ index value: 40–60 Commencement of monitoring: started in the operating room. Not stated when monitoring ceased **Group 2:** Entropy module (GE Healthcare) using Datex-Ohmeda S/5™ monitor Target device/ index value: response entropy and state entropy 40–60. Intermittent bolus doses of sufentanil given if response entropy–state entropy difference >10 for >2 minutes Commencement of monitoring: started in the operating room. Not stated when monitoring ceased **Group 3:** Standard practice (routine clinical signs) Hypertension/ hypotension, tachycardia Length of experience/ training of anaesthetist: described as 'more than 3 months of routine use'	**Total numbers involved:** $n = 140$; group 1, $n = 40$; group 2, $n = 40$; group 3, $n = 60$ Premedication used: 100 mg hydroxyzine orally 1 hour before surgery General anaesthetic used: i.v. propofol 2–3 mg/kg (induction). Sevoflurane in 60% nitrous oxide with oxygen Regional anaesthesia used: none Analgesia used: i.v. sufentanil 0.2–0.3 µg/kg injected over 15–30 seconds (induction), 0.15–0.20 µg/kg/hour with 5 µg bolus given 5 minutes before surgical incision. Intravenous morphine for postoperative analgesia started approximately 20 minutes prior to scheduled end of surgery (0.1–0.15 mg/kg), plus paracetamol, nefopam, non-steroidal anti-inflammatory drugs Muscle relaxants used: i.v. atracurium 0.5 mg/kg Antinausea drugs used: not stated Other drugs used: esmolol (for tachycardia), nicardipine 1–2 mg (hypertension), ephedrine 3–6 mg i.v./ phenylephrine 20–100 µg i.v. (for hypotension), atropine 0.5 mg i.v. (bradycardia) Type of surgery: abdominal; gynaecological, urological, orthopaedic Duration of surgery: precise duration not stated. Minimum 1 hour Duration of GA: ranged from 170.8 (± 90.6) minutes (standard practice group) to 190.8 (± 84.9) minutes (spectral entropy-guided group) **Inclusion criteria:** aged 18–80 years, ASA physical status I, II, III, scheduled for elective abdominal, gynaecological, urological or orthopaedic surgery expected to last at least 1 hour **Exclusion criteria**: history of any disabling central nervous or cerebrovascular disease, hypersensitivity to opioids or substance abuse, treatment with opioids or any psychoactive medication, or a body weight <70% or more than 130% of ideal body weight **Baseline measurements:** Sex (male), *n* (%): group 1 = 14 (41); group 2 = 23 (62%); group 3 = 23 (43%) Age years, mean (SD): group 1 = 57;(± 19); group 2 = 58 (± 18); group 3 = 54 (± 15) Ethnic groups, *n* (%): NR Weight kg: group 1 = 73 (± 18.2); group 2 = 77.6 (± 17.3); group 3 = 68.8 (± 13.4) ASA grade, *n* (I/II/III): group 1 = 13/16/5; group 2 = 14/19/4; group 3 = 26/24/4 Risk factors for awareness: none reported **Comorbidities:** none reported **Losses to follow-up:** none reported **Place of anaesthetic administration:** operating room	**Primary outcome:** • Reduction in sevoflurane consumption **Secondary outcomes:** • Sufentanil consumption • BIS and E-Entropy device values • Haemodynamic profiles (bradycardia, tachycardia, normal range of arterial blood pressure) • Treatment of adverse events (hypotension/ hypertension/ tachycardia/ bradycardia) • % of time passed with hypotension/ hypertension/ tachycardia/ bradycardia • Time to spontaneous eye opening • Time to extubation • Intraoperative recall **Length of follow-up**: intraoperative recall assessed on first and third postoperative days **Methods of assessing outcomes:** sevoflurane consumption measured by sevoflurane vaporiser weight: mean for one patient; mean for one patient normalised to the duration of anaesthetic; mean for one patient normalised to the duration of anaesthetic and also to the weight of the patient Intraoperative recall measured by standardised interview (Brice *et al.*[24])

NR, not reported; SD, standard deviation.

Outcome	Group 1	Group 2	Group 3	*p*-value
Intraoperative awareness/recall	0	0	0	NR
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR	NR
Time to spontaneous eye opening (minutes)	7.6 (± 4.1)	7.2 (± 4.7)	8.0 (± 3.9)	NR
Time to extubation (minutes)	11.1 (± 5.1)	11.5 (± 5.8)	14.2 (± 9.0)	NR
Time to discharge to/from the recovery room	NR	NR	NR	NR
Anaesthetic consumption (for one patient) mean (SD)				
Sevoflurane consumption (g)	21.3 (± 11.1)	22.8 (± 14.4)	25.6 (± 17.2)	0.49
Sevoflurane consumption normalised (g/hour)	7.2 (± 3.0)	7.8 (± 3.4)	9.4 (± 5.6)	0.07
Sevoflurane consumption normalised (g/kg/hour)	0.10 (± 0.04)	0.10 (± 0.05)	0.14 (± 0.09)	0.003
HRQoL	NR	NR	NR	NR
Nausea/vomiting/antisickness drugs	NR	NR	NR	NR
Pain/pain relieving drugs (for one patient)				
Sufentanil induction dose				
Sufentanil induction dose (μg/kg)	0.22 (± 0.05)	0.21 (± 0.05)	0.23 (± 0.06)	0.18
Sufentanil induction dose (μg/hour)	14.0 (± 6.7)	13.6 (± 6.1)	14.9 (± 8.3)	0.66
Sufentanil maintenance consumption (μg/kg/hour)	0.20 (± 0.09)	0.18 (± 0.09)	0.22 (± 12)	0.26
Other morbidity				
Ephedrine use (*n*)	3	2	4	NR
Nicardipine use (*n*)	1	2	2	NR
Esmolol	0	0	1	NR
Atropine (*n*)	1	0	0	NR
Mortality	NR	NR	NR	NR

NR, not reported; SD, standard deviation.

Additional results/comments (e.g. early response factors, QoL)

Percentage of time passed (induction, maintenance, recovery and total) with bradycardia (< 75% of baseline values), normal range of heart rate, tachycardia (> 125% of baseline values), hypotension (< 75% of baseline values), normal range of mean arterial blood pressure, and hypertension (> 125% of baseline values) were similar among groups (data not extracted)

Results demonstrate that BIS and spectral entropy guidance for the titration of sevoflurane results in a reduction of 29% in sevoflurane consumption

Sevoflurane consumption was statistically significantly different between study arms only when normalised for patient weight and duration of anaesthesia

Methodological comments

Allocation to treatment groups: random using a randomisation list performed with computer-generated random numbers

Allocation concealment: NR

Blinding: NR

Analysis by ITT: analysis excluded those who became ineligible post randomisation

Comparability of treatment groups at baseline: reported to be similar in demographics except that patients in the E-Entropy-guided group (group 2) were statistically significantly heavier ($p = 0.04$). More males were included in the E-Entropy-guided group

Method of data analysis: chi-squared test for nominal data. One-way analysis of variance with Bonferroni's test for multiple comparisons used for numerical data

Sample size/power analysis: previous open study from the authors' institution in the same surgical population showed that sevoflurane consumption was 0.16 ± 0.10 g/kg/hour. Applying an a priori power analysis, at least 34 patients had to be enrolled in each treatment group to detect a reduction of 50% in the sevoflurane consumption with a risk α of 0.05 and a statistical power of 0.9. The authors included 60 patients in the standard practice group and 40 in the BIS and spectral E-Entropy-guided groups

Attrition/dropout: six patients excluded from group 3 (one not extubated at the end of surgery due to hypothermia, three required intraoperative propofol administration, and missing data in two cases), six patients excluded from group 1 (three not extubated at the end of surgery because of hypothermia, two required intraoperative propofol administration, and monitor data were lost in one case), and three from group 2 (all were not extubated at the end of surgery because of hypothermia, two required intraoperative propofol administration)

General comments

Generalisability: general surgical population receiving an inhaled maintenance anaesthetic, not specifically identified as at increased risk for intraoperative awareness

Intercentre variability: NA

Conflict of interests: none declared. Some of the monitoring equipment used was provided by GE Healthcare

NA, not applicable; NR, not reported.

Domain	Author's judgement (state: low/high/unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Low	Computer-generated randomisation
Allocation concealment	Unclear	No information given
Performance bias		
Blinding of participants and personnel	Unclear	No information given
Detection bias		
Blinding of outcome assessment	Unclear	No information given
Attrition bias		
Incomplete outcome data	Low	Exclusions generally balanced between groups, and generally similar reasons given
Reporting bias		
Selective reporting	Low	No evidence to suggest selective reporting

Avidan *et al.*

Reviewer 1: JS	Reviewer 2: GF		
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Avidan[44] Year: 2011 **Study design:** RCT **Number of centres:** three **Countries:** USA/Canada **Sponsors:** Foundation for Anaesthesia Education & Research; American Society of Anaesthetists Winnipeg Regional Health Authority & University of Manitoba Department of Anaesthesia; Department of Anaesthesiology at Washington in St. Louis; University; Department of Anaesthesiology at University of Chicago **Trial name:** BIS or Anaesthetic Gas to Reduce Explicit Recall trial (BAG-RECALL)	**Group 1:** BIS (Covidien) Target device/index value: 40–60 (audible alarms used outside of this range) **Group 2:** ETAC (audible alarms used outside of 0.7 to 1.3 age-adjusted MAC range in group 2 only) Patients in group 2 had monitors configured to conceal the BIS value and did not receive a BIS audible alarm Commencement of monitoring: not stated Length of experience/training of anaesthetist: summaries of BIS and ETAC protocols were given to the practitioners to provide education and to increase adherence. Signs were affixed to anaesthesia machines to remind practitioners to check BIS/ETAC and consider patient awareness	**Total numbers involved:** 6041 randomised; 3021 (group 1); 3020 (group 2) Premedication used: midazolam used in 80.8% patients (group 1); 79.7% of patients (group 2) General anaesthetic used: isoflurane, sevoflurane or desflurane (further information not reported) Regional anaesthesia used: none (except for 13 patients who were excluded from the study) Analgesia used: not stated Muscle relaxants used: not stated Antinausea drugs used: not stated Other drugs used: not stated Type of surgery: not explicitly reported, but inclusion criteria refer to open heart surgery (see below) Duration of surgery: not stated Duration of GA: not stated **Inclusion criteria:** 18 years or older, undergoing GA with isoflurane, sevoflurane or desflurane. At high risk for intraoperative awareness for one or more of the following risk factors: planned open heart surgery; aortic stenosis; pulmonary hypertension; use of opiates; use of benzodiazepines; use of anticonvulsant drugs; daily alcohol consumption; ASA status 4; end-stage lung disease; history of intraoperative awareness; history of or anticipated difficult intubation; cardiac ejection fraction <40%; marginal exercise tolerance **Exclusion criteria:** patients with dementia, unable to provide written informed consent, or had a history of stroke with residual neurological deficits. 'Minor risk factors' for awareness as used in the B-Aware study were not used as enrolment criteria **Baseline measurements:** Sex (male), *n* (%): group 1 = 1621 (56.7); group 2 = 1679 (58.9) Age years, mean (SD): group 1 = 60 (± 14.2); group 2 = 61 (± 14.4)	**Primary outcome:** • Incidence of definite intraoperative awareness **Secondary outcomes:** • Definite or possible awareness (pre-specified secondary outcome) • Distressing experience of awareness (post hoc secondary outcome) **Length of follow-up:** up to 30 days post extubation **Methods of assessing outcomes:** awareness assessed by modified Brice questionnaire (references cited). Assessments made 72 hours after surgery, and 30 days after extubation. Patients who reported memories of the period between 'going to sleep' and 'waking up' were contacted by a different evaluator, who asked additional structured questions. Three experts independently reviewed responses to the questionnaire from patients who had reported memories and determined whether the reported event involved definite awareness, possible awareness or no awareness. Experts assigned each event of definite or possible awareness to one of the categories of the Michigan Awareness Classification Instrument. In the event of divergence of opinion a fourth expert reviewer who reviews cases for the Anaesthesia Awareness Registry of the ASA, made the final determination

Reviewer 1: JS Reviewer 2: GF			
Reference and design	**Technology**	**Participants**	**Outcome measures**
		Ethnic groups, *n* (%): White: group 1 = 2405 (84.1); group 2 = 2388 (83.7) Black: group 1 = 357 (12.5); group 2 = 369 (12.9) Other: group 1 = 99 (3.5); group 2 = 95 (3.3) Weight BMI (SD): group 1 = 30 (± 8.4); group 2 = 30 (± 8.3) ASA grade, *n* (%): 1: group 1 = 23 (0.8); group 2 = 19 (0.7) 2: group 1 = 468 (16.4); group 2 = 407 (14.3) 3: group 1 = 1416 (49.5); group 2 = 1407 (49.3) 4: group 1 = 954 (33.3); group 2 = 1019 (35.7) Composite number of inclusion criteria met (risk factors as defined above under 'inclusion criteria') • Median: 2 (group 1); 2 (group 2) • Interquartile range: 1–3 (group 1); 1–3 (group 2) Comorbidities: Composite number of pre-existing medical conditions (as above) • Median: 2 (group 1); 2 (group 2) • Interquartile range: 1–3 (group 1); 1–3 (group 2) **Losses to follow up:** 46 (group 1); 50 (group 2) **Place of anaesthetic administration:** NR	

NR, not reported; SD, standard deviation.

Outcome	Group 1	Group 2	Difference, BIS-ETAC percentage points (95% CI)	*p*-value
Intraoperative awareness, *n/N* (%)				
Definite	7/2861 (0.24)	2/2852 (0.07)	0.17 (–0.03 to 0.38)	0.98
Definite or possible	19/2861 (0.66)	8/2852 (0.28)	0.38 (0.03 to 0.74)	0.99
Patient distress and sequelae resulting from perioperative awareness, *n* (%)	8/2861 (0.28)	1/2852 (0.04)	0.24 (0.04 to 0.45)	0.99
Time to emergence from anaesthesia	NR	NR	NR	NR
Time to extubation	NR	NR	NR	NR
Time to discharge to/from the recovery room	NR	NR	NR	NR
Anaesthetic consumption	NR	NR	NR	NR
HRQoL	NR	NR	NR	NR
Nausea/vomiting/antisickness drugs	NR	NR	NR	NR
Pain/pain-relieving drugs	NR	NR	NR	NR
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR	NR
Mortality				
Died before first interview	33/2907 (1.14%)	38/2902 (1.31%)	NR	NR
30-day mortality	57/2907 (1.96%)	64/2902 (2.21%)	0.24 (–0.50 to 0.99)	NR

NR, not reported.

Additional results/comments

In total, 49 patients, including patients from all three enrolment sites, reported having memories of the period between 'going to sleep' and 'waking up' at the end of surgery

Experts determined that nine patients had definite intraoperative awareness (incidence 0.16%, 95% CI 0.08 to 0.30), and 27 patients had definite or possible awareness (incidence 0.47%, 95% CI 0.32 to 0.68)

A classification of awareness events is given, according to the Michigan Awareness Classification (data not extracted)

Patients who experienced awareness compared with patients who did not, met a median of one additional inclusion criterion and had a median of one additional pre-existing medical condition

A total of five of the nine patients who experienced possible awareness did not have either BIS values of >60 or ETAC values of <0.7 age-adjusted MAC

Overall, during the maintenance of anaesthesia the BIS was <60, a median of 94.0% of the time (interquartile range, 93.6–100), and the ETAC was >0.7 age-adjusted MAC, a median of 84.8% of the time (interquartile range, 67.2–95.3)

In both groups the median length of stay in the hospital was 7.0 days, and the median length of stay in the ICU was 2.1 days

There were no important differences between the groups in the doses of sedative, hypnotic, opioid analgesic or neuromuscular-blocking drugs administered

Methodological comments

Allocation to treatment groups: 6100 pre-randomisation designations were generated electronically *n* blocks of 100, divided equally between the groups

Allocation concealment: labels indicating BIS group or ETAC group were sealed in opaque, numbered envelopes

Blinding: the anaesthesia practitioners were aware of the patients' group assignments, but the patients, the postoperative interviewers, the expert reviewers and the statisticians were not

Analysis by ITT: a modified ITT analysis was performed, which included all patients who underwent randomisation and who were assessed for intraoperative awareness. All the patients were treated with the protocol to which they had been randomly assigned

Comparability of treatment groups at baseline: Statistically significant differences were found for two variables: use of anticonvulsant drugs (slightly higher in group 1); cardiac ejection fraction <40% (slightly higher in group 2)

Method of data analysis: Fisher's exact test for primary and secondary analysis. Chi-squared test, Fisher's exact test, unpaired Mann–Whitney *U*-test or unpaired Student's *t*-test used for other comparisons

Sample size/power analysis: it is estimated that with 6000 patients the study would have 87% power to detect a clinically significant reduction of 0.4 percentage points in the incidence of definite awareness with the BIS protocol, compared with the ETAC protocol (from 0.5% in the ETAC group to 0.1% in the BIS group), at a one-tailed alpha level of 0.05 with the use of Fisher's exact test

Attrition/dropout: of the 3021 patients randomised to group 1, 114 (3.8%) were excluded post randomisation. Of the remaining 2907 patients, 46 (1.6%) were lost to follow-up and 2861 were assessed for intraoperative awareness. Of the 3020 patients randomised to group 2, 118 (3.9%) were excluded. Of the remaining 2902, 50 (1.7%) were lost to follow-up and 2852 were assessed for intraoperative awareness. Reasons given for exclusions and loss to follow-up in both groups were similar (primarily death before awakening). 5713 (98.3%) completed at least one postoperative interview and were included in the primary outcome analysis. 5413 (93.2%) completed the postoperative interviews at both times (within 72 hours after surgery and at 30 days after extubation)

General comments

Generalisability: surgical population classified at high risk of intraoperative awareness receiving inhaled anaesthesia. Not applicable to the general surgical population, and those receiving i.v. anaesthesia. BIS and ETAC were used as part of structured protocols. It was not the intention of the protocols to prescribe or restrict the use of anaesthetic agents. Practitioners could decrease anaesthetic administration at their discretion if a patient's condition was haemodynamically unstable. The protocols were designed to increase vigilance and to provide warnings that patients might be aware

Intercentre variability: median BIS and ETAC values were similar between the three study sites

Conflict of interests: states that no potential conflict of interest was reported

Domain	Author's judgement (state: low/high/unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Low	Electronic randomisation
Allocation concealment	Low	Sealed opaque envelopes
Performance bias		
Blinding of participants and personnel	Unclear	
Detection bias		
Blinding of outcome assessment	Low	Postoperative interviewers, the expert reviewers and the statistician were not aware of group assignment
Attrition bias		
Incomplete outcome data	Low	Level of missing data from postrandomisation exclusions and loss to follow-up and reasons were similar between study arms
Reporting bias		
Selective reporting	Low	No evidence to suggest selective reporting

Bannister *et al.*

Reviewer 1: GF Reviewer 2: JS			
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Bannister *et al.*[45] **Year:** 2001 **Study design:** RCT **Number of centres:** not reported; appears to be one **Country:** USA **Sponsor:** supported in part by a grant from Aspect medical systems (device manufacturer)	**Group 1:** BIS (version 3.3, Aspect Medical Systems) using an A-1050 EEG monitor Target device/ index value: 40–60 during maintenance and 60–70 during last 15 minutes of surgery Commencement of monitoring: prior to anaesthesia; location not reported **Group 2:** standard practice (at anaesthesiologist's discretion using unspecified clinical signs and haemodynamic changes). BIS was recorded but the anaesthesiologist was blinded to BIS data Length of experience/training of anaesthetist: NR	**Total numbers involved:** $n = 75$; group 1, $n = 40$, group 2, $n = 35$ NB part of a wider study (total $n = 202$) that included patients aged 0–3 years and 3–18 years, with patients randomised within age groups. Only the 3- to 18-years age group meets the systematic review age inclusion criterion and is reported here (mean age in the younger group ≤ 2.2 years) Premedication used: midazolam 0.3–0.75 mg/kg (group 1, 77.5%, group 2, 88.6%) General anaesthesia (induction and maintenance): sevoflurane in 60% N_2O in oxygen (8% sevoflurane in induction; not stated for maintenance) Regional anaesthesia: none Analgesia: fentanyl 1–2 μg/kg or morphine 0.05–0.1 mg/kg Muscle relaxants: non-polarising i.v. neuromuscular block (no other details) Antinausea drugs: none reported Other drugs: opioids (dose not specified) Type of surgery: tonsillectomy and/or adenoidectomy Duration of surgery, mean ± SD: group 1, 27.7 ± 17.1 minutes; group 2, 33.2 ± 20.3 minutes Duration of GA: not reported **Inclusion criteria:** not reported other than age 6–18 years and undergoing tonsillectomy and/or adenoidectomy **Exclusion criteria:** NR Baseline measurements: Sex (male), *n* (%): group 1, 26 (65.0); group 2, 23 (65.7) Age (years), mean ± SD: group 1, 6.7 ± 2.5; group 2, 6.1 ± 2.6 Ethnic groups, *n* (%): NR Weight (kg), mean ± SD: group 1, 26.9 ± 10.6; group 2: 27.7 ± 14.7 ASA grade: NR Risk factors for awareness: none reported Comorbidities: none reported **Losses to follow-up:** none reported **Place of anaesthetic administration:** NR	**Outcomes (not reported whether primary or secondary):** • Sevoflurane consumption • BIS device values • Time to first movement response • Time to extubation • Time to PACU discharge • Haemodynamic parameters (mean arterial pressure and heart rate) **Length of follow-up:** limited to period up to discharge from PACU **Methods of assessing outcomes:** sevoflurane concentration was measured with a Capnomac Ultima gas analyser (Datex Medical Instrumentation Inc., Helsinki, Finland) and end-tidal concentration was continuously recorded by a computer PACU discharge readiness was defined as a score of ≥ 12, with no zeros, on a modified Aldrete scale and in a room air O_2 saturation ≥ 94%

NR, not reported.

Outcome	Group 1: BIS ($n = 40$)	Group 2: Standard clinical practice ($n = 35$)	p-value
Intraoperative awareness/recall	NR	NR	NR
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time to emergence from anaesthesia: mean ± SD time to first movement response, minutes	4.2 ± 3.7	7.0 ± 3.9	<0.05
Mean ± SD time to extubation, minutes	7.1 ± 3.7	11.3 ± 5.9	<0.05
Mean ± SD time to discharge from the PACU	20.0 ± 7.9	26.7 ± 11.2	<0.05
Anaesthetic consumption: mean ± SD end-tidal sevoflurane concentration (%)			
Maintenance of GA	1.8 ± 0.4	2.4 ± 0.	<0.05
Last 15 minutes of GA	1.6 ± 0.6	2.1 ± 0.7	<0.05
End of procedure	1.1 ± 0.6	1.5 ± 0.7	NS
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs	NR	NR	NR
Pain/pain-relieving drugs			
Opioid use, *n* (%)	37 (92.5)	35 (100)	NR
Other morbidity	NR	NR	NR
Mortality	NR	NR	NR

NR, not reported; NS, not statistically significant ($p \geq 0.05$).

Additional results/comments (e.g. early response factors, QoL)

Primary outcome not specified but the main focus appears to be on anaesthetic consumption and recovery times

Stated there were no statistically significant differences among groups for mean arterial pressure or heart rate recorded during surgery (no quantitative data or *p*-values provided)

Stated there were no intergroup differences in any measured variables between group 2 and a historical control group – showing no change in clinical practice during the trial

Methodological comments

Allocation to treatment groups: stated random allocation but sequence generation method not reported

Allocation concealment: NR

Blinding: single observer blinded to the patient groups was responsible for all PACU discharge assessments

Analysis by ITT: unclear: ITT not mentioned and sample sizes not reported for outcomes

Comparability of treatment groups at baseline: stated no statistically significant differences in demographic data between the groups (no *p*-values reported), but data were only provided for age, weight and sex, which were similar in the two study groups. No information was provided on ethnicity or health status

Method of data analysis: non-normally distributed variables (not specified) were identified by Kolmogorov–Smirnov statistic then log-transformed. Parametric data (not specified) were compared between group 1 and group 2 using Bonferroni-corrected *t*-tests. Chi-squared test was used to compare sex distribution

Sample size/power analysis: NR

Attrition/dropout: none reported

General comments

Generalisability: North American paediatric population aged 6–18 years undergoing tonsillectomy and/or adenoidectomy under sevoflurane for GA; socioeconomic details not reported. Not specifically identified as at risk for intraoperative awareness

Intercentre variability: NA (appears to be a single-centre study)

Conflict of interests: funded in part by Aspect Medical Systems (AMS) who supplied the BIS monitor. One author was employed by AMS; another author was a paid consultant to AMS

NA, not applicable; NR, not reported.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Unclear	No information given
Allocation concealment	Unclear	No information given
Performance bias		
Blinding of participants and personnel	Unclear	No information given
Detection bias		
Blinding of outcome assessment	Unclear	Single observer blinded to the patient groups was responsible for all PACU discharge assessments. Not reported whether or not observers were blinded for other outcomes
Attrition bias		
Incomplete outcome data	Unclear	Attrition and sample sizes for outcomes not reported
Reporting bias		
Selective reporting	Low	No evidence to suggest selective reporting
Other bias		
Other sources of bias	High	Notable conflict of interest declared likely to favour results supporting the utility of BIS-guided anaesthesia

Bhardwaj and Yaddanapudi

Reviewer 1: JS Reviewer 2: JB			
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Bhardwaj and Yaddanapudi[46] **Year:** 2010 **Study design:** RCT **Number of centres:** one **Country:** India **Sponsor:** not stated	**Group 1:** BIS Monitor Model A-2000 IP X 2 (Aspect Medical Systems Inc., Newton, MA, USA) (propofol infusion rate manually altered by 20 µg/kg/minute to achieve a BIS value between 45 and 60) **Group 2:** Standard clinical practice (propofol infusion rate manually altered by 20 µg/kg/minute if systolic blood pressure changed by >20% of baseline) Commencement of monitoring: following transition to the operating theatre and just before start of induction of anaesthesia. Monitoring continued in recovery room and monitored until patients achieved discharge criteria (Steward score of 6) BIS monitoring took place in both groups, but monitor was kept covered in group 2 Length of experience/training of anaesthetist: not stated	**Total numbers involved:** 50; group 1 = 25; group 2 = 25 Premedication used: midazolam 0.5 mg/kg General anaesthetic used: propofol 3 mg/kg (induction). Propofol 150 µg/kg/minute with nitrous oxide in oxygen (FiO_2 0.33) (maintenance) Regional anaesthesia used: none Analgesia used: morphine 0.1 mg/kg (induction). Additional dose of opioid (fentanyl or morphine) was administered if signs of inadequate anaesthesia detected Muscle relaxants used: atracurium (0.5 mg/kg) used to facilitate tracheal intubation Antinausea drugs used: NR Other drugs used: atropine used to treat bradycardia (heart rate <80 of baseline). Neostigmine (0.05 mg/kg and atropine (0.025 mg/kg) used for reversal of neuromuscular blockade Type of surgery: elective urogenital surgery Duration of surgery (minutes), mean (SD): group 1 = 65.6 (29.2); group 2 = 71.8 (27.3) Duration of GA (minutes), mean (SD): group 1 = 88.6 (31.8); group 2 = 95.1 (28.3) **Inclusion criteria:** ASA 1 children aged 2–12 years undergoing elective urogenital surgery of about 1 hour in duration under GA **Exclusion criteria:** Patients with epilepsy and those taking drug known to affect EEG **Baseline measurements:** Sex (male), *n* (%): group 1 = 21/25 (84%); group 2 = 24/25 (96%) Age (years), mean (SD): group 1 = 6.3 (3.2); group 2 = 6 (3) Ethnic groups, *n* (%): NR Weight (kg), mean (SD): group 1 = 18.7 (8.1); group 2 = 18.5 (5.9) ASA grade: all grade 1 Risk factors for awareness: NR Comorbidities: NR **Losses to follow-up:** NA **Place of anaesthetic administration:** premedication took place prior to transfer to the operation theatre. GA was initiated in the operation theatre	**Primary outcome:** • Reduction in consumption of propofol **Secondary outcome:** • Recovery from anaesthesia **Length of follow-up:** NA (all outcomes measured at the end of surgery) **Methods of assessing outcomes:** Steward recovery scoring system used to assess eligibility for discharge from the recovery room (eligibility = score of 6) Duration of anaesthesia was defined as the time from the start of propofol bolus for induction to extubation of trachea. Duration of surgery was defined as the time from surgical incision to the application of last suture

NA, not applicable; NR, not reported; SD, standard deviation.

Outcome	Group 1	Group 2	*p*-value
Intraoperative awareness/recall	NR	NR	NR
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time to emergence from anaesthesia	Time to eye-opening and time to response to commands reported to be comparable in the two groups No difference in the time interval between end of anaesthesia and return of consciousness between the groups on basis of log-rank test; ($p = 0.86$)		
Time to extubation	Time to extubation reported to be comparable in the two groups		
Time to discharge to/from the recovery room	Time to achieve a Steward recovery score of 6 (for discharge from the recovery room) reported to be comparable in the two groups		
Anaesthetic consumption			
Propofol consumption during maintenance of anaesthesia, mean (SD)	108.6 µg/kg/minute (37.8)	106.6 µg/kg/minute (38.9)	NR Mean difference 1.9 (95% CI –19.9 to 23.7)
Total propofol consumption, mean (SD)	232.6 mg (136.7)	250.8 mg (118.2)	NR Mean difference –18.1 (95% CI –68.2 to 76)
Duration of propofol infusion, mean (SD)	82 minutes (29.2)	86 minutes (28.5)	NR Mean difference –4 (95% CI –20 to 13.5)
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs	NR	NR	NR
Pain/pain-relieving drugs			
Morphine consumption, Mean (SD)	1.9 (08)	1.9 (0.6)	NR Mean difference –0.01 (95% CI –0.4 to 0.4)
Other morbidity, *n*/*N* (%)			
Hypertension	5/25 (20%)	5/24 (21%)	NR
Hypotension	6/25 (24%)	7/24 (29%)	NR
Bradycardia	8/25 (32%)	6/24 (25%)	NR
Mortality	NR	NR	NR

NR, not reported.

Additional results/comments (e.g. early response factors, QoL)

Mean propofol infusion rates at various time intervals during the course of surgery were similar in the two groups

The number of patients requiring additional opioids was similar in both groups (two patients in group 1 compared with three patients in group 2)

Mean heart rate and systolic blood pressure were not statistically different between the groups during the duration of surgery

Methodological comments

Allocation to treatment groups: computer-generated randomisation table

Allocation concealment: randomisation to the two groups was performed by opening a sealed envelope

Blinding: NR

Analysis by ITT: all patients received their allocated intervention. Only one patient was excluded from the analysis (group 2) because the child received lower propofol infusion rate owing to wrong dose calculation. Note that table 1 which provides demographic data and study outcomes lists there being 25 patients in each group

Comparability of treatment groups at baseline: authors state that the two study groups were comparable in terms of demographic variables (age, weight, sex)

Method of data analysis: age, weight, heart rate, systolic blood pressure, and duration of anaesthesia, surgery and propofol infusion were compared between groups using Student's *t*-test, whereas the BIS values were compared between groups using Mann–Whitney *U*-test

Sample size/power analysis: calculated that 22 patients required in each study group to detect a 20% difference in propofol consumption [average requirement of propofol 150 μg/kg/minute (SD 30) with an alpha error of 0.05 and power of 90%]. To compensate for any exclusion 25 patients were studied in each group

Attrition/dropout: as above, one patient was excluded from the analysis from group 2

General comments

Generalisability: authors state that they used the three-sensor device for BIS monitoring and that it does not use the new XP technology. The newer version became available later in the study but was not used as the algorithm in the newer device may be different and may affect results. Results of this study may therefore not be applicable to newer versions of BIS monitors

Intercentre variability: NA

Conflict of interests: reported as 'Nil'

Other: the authors note that the Steward score for anaesthetic recovery has never been formally validated for the paediatric patient population, although is widely accepted as a tool in paediatric anaesthesia research

NA, not applicable; NR, not reported.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Low	Computer-generated randomisation table
Allocation concealment	Unclear	Sealed envelopes were used although it does not say whether or not they were opaque
Performance bias		
Blinding of participants and personnel	Unclear	NR
Detection bias		
Blinding of outcome assessment	Unclear	NR
Attrition bias		
Incomplete outcome data	Low	Only one exclusion from the study
Reporting bias		
Selective reporting	Low	No evidence to suggest selective reporting

NR, not reported.

Chan *et al.*

Reviewer 1: GF Reviewer 2: JS			
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Chan *et al.*[47] **Year:** 2010 **Study design:** RCT **Number of centres:** two **Country:** China **Sponsor:** none reported Note: abstract only	**Group 1:** BIS (no further details) Target device/index value: 40–60 during maintenance of GA Commencement of monitoring: NR **Group 2:** routine practice Anaesthesia adjusted according to traditional clinical signs and haemodynamic parameters (no further details). BIS was measured but values were not revealed to the anaesthesiologist Length of experience/ training of anaesthetist: NR	**Total numbers involved:** Starting number: 921; group 1, 449; group 2, 452 Number randomised per group not stated. Difference (20 patients) between starting number and sample size reported for outcomes but unclear whether this reflects attrition before or after randomisation NB. There was also a matched control group of 211 non-surgery patients which were outside of the randomised cohort – unclear in the presentation of one outcome whether 'control' refers to this group or to the routine practice group Premedication used: NR General anaesthetic used: not explicitly reported but implied that both an inhalational agent and i.v. propofol were involved Regional anaesthesia used: not reported Analgesia used: NR Muscle relaxants used: NR Antinausea drugs used: NR Other drugs used: NR Type of surgery: stated as major non-cardiac surgery (no other details) Duration of surgery: NR Duration of GA: NR **Inclusion criteria:** elderly patients (>60 years) undergoing major non-cardiac surgery. No other details reported **Exclusion criteria:** none reported **Baseline measurements:** stated that patient characteristics and surgical details were similar between groups. No baseline data reported **Losses to follow-up:** NR **Place of anaesthetic administration:** NR	**Outcomes (not stated whether primary or secondary):** • POCD • BIS device values • Anaesthetic consumption **Length of follow-up:**1 week and 3 months after surgery **Methods of assessing outcomes:** POCD assessed by a battery of eight neuropsychology tests before and at 1 and 3 weeks after surgery (no information on the tests reported). POCD was confirmed when two or more test parameters or the combined *z*-score >1.96 (no further information given)

NR, not reported.

Outcome	Group 1 (BIS) (*n* = 449)	Group 2 (routine care) (*n*=452)	*p*-value
Intraoperative awareness/recall	NR	NR	NR
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time to emergence from anaesthesia	NR	NR	NR
Time to extubation	NR	NR	NR
Time to discharge to/from the recovery room	NR	NR	NR
Anaesthetic consumption			
ETAC	25.3% reduction vs group 2[a]	NR	NR
Target plasma propofol concentration	20.7% reduction vs group 2[a]	NR	NR
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs	NR	NR	NR
Pain/pain-relieving drugs	NR	NR	NR
Other morbidity (e.g. cognitive dysfunction), *n* (%)[b]			
POCD, 1 week post surgery	146 (32.5)	177 (39.1)	0.07
POCD, 3 months post surgery	36 (8.1)	54 (12.0)	0.03 [OR (95% CI) 1.6 (1.0 to 2.4)]
Mortality	NR	NR	NR

NR, not reported.

a Assumed by reviewer that this comparison was between groups 1 and 2; however, the wording of the results does not rule out that the comparison may instead have been between group 1 and the matched 'control' group.

b Percentages only were provided in the abstract; numbers of patients estimated by reviewer.

Additional results/comments (e.g. early response factors, QoL)

Only an abstract is available, hence, the information reported is limited

Reported ETAC and target plasma propofol concentration outcomes which would correspond, respectively, to inhaled and i.v. anaesthesia; unclear how the patients received these different types of anaesthesia, as no subgroups were specified

Methodological comments:

Allocation to treatment groups: random assignment. No further details given

Allocation concealment: NR

Blinding: NR

Analysis by ITT: not discernible as the number randomised and the analysis methods were not reported

Comparability of treatment groups at baseline: stated patient characteristics and surgical details similar between groups, but no data provided for any variables

Method of data analysis: NR

Sample size/power analysis: NR

Attrition/dropout: NR. The starting number of patients (921) is 20 more than the total sample size indicated for outcomes data (449 + 452 = 901); unclear whether or not this difference reflects attrition pre or post randomisation

General comments

Generalisability: elderly Chinese patients (> 60 years) undergoing major non-cardiac surgery under GA, but limited information on the types of anaesthesia (appears to include both inhaled and i.v.); unclear population characteristics (sex, weight, comorbidities not reported); unclear surgical procedures (no information reported); and unclear which groups some outcomes were reported for. Not reported whether or not population was at high risk of intraoperative awareness

Intercentre variability: NR

Conflict of interests: none reported

NR, not reported.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Unclear	No information given
Allocation concealment	Unclear	No information given
Performance bias		
Blinding of participants and personnel	Unclear	No information given
Detection bias		
Blinding of outcome assessment	Unclear	No information given
Attrition bias		
Incomplete outcome data	Unclear	No information given – number randomised not discernible
Reporting bias		
Selective reporting	Unclear	Stated that postoperative complications were recorded, but these were not reported

Choi *et al.*

Reviewer 1: JS	Reviewer 2: GF		
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Choi *et al.*[54] **Year:** 2010 **Study design:** RCT **Number of centres:** not stated (presume single-centre) **Country:** South Korea **Sponsor:** Dong-A University	**Group 1:** E-Entropy (GE Datex-Ohmeda S/5 Anaesthesia monitor, Helsinki, Finland) Target device/index value: state entropy 40–50 Entropy sensor stripes were applied upon arrival in the operating room **Group 2:** standard practice Sevoflurane adjusted to maintain heart rates and systolic blood pressures within 20% of the baseline values Entropy indices were recorded with the anaesthesiologist blinded to them Length of experience/ training of anaesthetist: not stated	**Total numbers involved:** 80 patients enrolled. 39 were included in each group Premedication used: i.v. midazolam (0.15 mg/kg) General anaesthetic used: 5% vol% sevoflurane in oxygen at fresh gas flow of 5 l/minute (induction). Sevoflurane administration was started at 2.5 vol% in air and oxygen 1.5 l/minute Regional anaesthesia used: not stated Analgesia used: intraoperative analgesics were not used as their sedative effect may not be detected by entropy monitoring. ketorolac (non-steroidal anti-inflammatory) 0.5 mg/kg i.v. administered following sevoflurane cessation Muscle relaxants used: rocuronium 0.6 mg/kg i.v. used for endotracheal intubation Antinausea drugs used: NR Other drugs used: NR Type of surgery: tonsillectomy/adenoidectomy Duration of surgery (minutes), mean (SD): group 1 = 41.4 (± 14.8); group 2 = 48.1 (± 17.8) Duration of GA (minutes), mean (SD): group 1 = 64.3 (± 16.4); group 2 = 67.9 (± 19.7) **Inclusion criteria:** ASA physical status I-II, aged 3–12 years, scheduled for tonsillectomy/ adenoidectomy **Exclusion criteria:** children with any neurological disease or on any antiseizure medication **Baseline measurements:** Sex (male), *n* (%): group 1 = 25/39 (64); group 2 = 27/39 (69) Age (years), median (range): group 1 = 4.0 (3.0–12.0); group 2 = 6.0 (3.0–11.0) Ethnic groups, *n* (%): NR Weight (kg), median (range): group 1 = 24.0 (13.0–35.0); group 2 = 22.0 (14.0–52.0) ASA grade: physical status I–II Risk factors for awareness: none reported Comorbidities: none reported **Losses to follow-up:** NR **Place of anaesthetic administration:** not stated	**Primary outcome:** • Reduction in sevoflurane use, as expressed by end-tidal sevoflurane concentration (described as the 'final end-point') **Secondary outcomes:** • Time to extubation • Time to eye opening • Time to orientation • Time to complete recovery • Intraoperative recall • Haemodynamic parameters (heart rate; systolic and diastolic blood pressure) • Entropy values (state and response entropy) **Length of follow-up:** longest follow-up appears to be the first postoperative day (for intraoperative recall) **Methods of assessing outcomes:** end-tidal sevoflurane concentration, entropy values and heart rate were continuously recorded using the S/5 Collect software program (GE Healthcare) on a computer hard drive for off-line analysis. The average end-tidal sevoflurane concentration, entropy values and haemodynamic parameters during anaesthetic maintenance were calculated using data collected from the application of the gag retractor to the end of surgery Patients were interviewed about intraoperative recall in the PACU and on the first postoperative day by an independent nurse Time to the various recovery parameters was measured following discontinuation of sevoflurane. Complete recovery was defined as a score of 9 or more on a modified Aldrete score

NR, not reported.

Outcome	Group 1	Group 2	*p*-value
Intraoperative awareness/recall	Anaesthesia and surgery-related memories were not reported by any patients in the postoperative interview		
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time (minutes) to emergence from anaesthesia, mean (SD)			
Eye-opening	14.3 (3.6)	18.0 (3.3)	NS
Orientation	18.2 (4.0)	23.3 (5.0)	<0.05
Complete recovery	24.3 (7.3)	28.8 (5.7)	<0.05
Time (minutes) to extubation, mean (SD)	8.3 (1.4)	11.9 (2.5)	<0.05
Time (minutes) to discharge to/from the recovery room	NR	NR	NR
Anaesthetic consumption, end-tidal sevoflurane%, mean (SD)	2.2 (0.3)	2.6 (0.4)	<0.05
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs	NR	NR	NR
Pain/pain-relieving drugs	NR	NR	NR
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

NR, not reported; NS, not statistically significant.

Additional results/comments (e.g. early response factors, QoL)

Systolic and diastolic blood pressure were significantly higher in group 1 compared with group 2 during anaesthesia maintenance ($p<0.05$)

Methodological comments

Allocation to treatment groups: random, no further information given

Allocation concealment: parents opened a sealed envelope

Blinding: not stated

Analysis by ITT: NR. Analysis excludes two patients out of the 80 enrolled because of 'technical problems'. It is not clear whether this was pre or post randomisation

Comparability of treatment groups at baseline: authors state that there were no statistically significant demographic differences between the groups or in the anaesthetic times or duration of surgery

Method of data analysis: nominal data were compared using the chi-squared test and parametric data were compared using the two-sided *t*-test

Sample size/power analysis: applying a priori analysis, at least 33 patients had to be enrolled in each group to detect a reduction of 20% in end-tidal sevoflurane concentration with an alpha of 0.05 and a statistical power of 0.9. Forty patients were enrolled in each group for redundancy

Attrition/dropout: two patients out of the 80 enrolled were excluded from the analysis because of 'technical problems'

General comments

Generalisability: results applicable to Korean children without any apparent comorbidities undergoing tonsillectomy/adenoidectomy. Not stated to be at increased risk for intraoperative awareness

Intercentre variability: NA (presumed single centre)

Conflict of interests: none reported

NA, not applicable; NR, not reported.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Unclear	No information given on randomisation method
Allocation concealment	Unclear	States that parents opened a sealed envelope, although it is not reported whether or not the envelope was opaque
Performance bias		
Blinding of participants and personnel	Unclear	No information given
Detection bias		
Blinding of outcome assessment	Unclear	No information given
Attrition bias		
Incomplete outcome data	Low	Two patients were excluded from the analysis, although it is not clear at when or why these exclusions happened (other than for 'technical problems'). As this is a relatively low number, and given that the study recruited a greater number of participants than were needed (as estimated from the power calculation), attrition bias may be low
Reporting bias		
Selective reporting	Low	No evidence to suggest selective reporting

Ellerkmann *et al.*

Reviewer 1: JB Reviewer 2: JS

Reference and design	Technology	Participants	Outcome measures
Author: Ellerkmann *et al.*[62] **Year:** 2010 **Study design:** RCT **Number of centres:** 1 **Country:** Germany **Sponsor:** not stated	**Group 1:** Entropy module (GE Healthcare, version not stated) with BIS monitor A-2000 Propofol adjusted State entropy to target value of 50 during maintenance Target state entropy value of 60 to facilitate rapid emergence from anaesthesia (15 minutes before expected end of surgery) **Group 2:** BIS Monitor A-2000 (version XP, software version 4.0) Propofol adjusted to target value of 50 during maintenance Target value of 60 to facilitate rapid emergence from anaesthesia (15 minutes before expected end of surgery) In the E-Entropy and BIS group, a propofol bolus of 0.25 mg/kg could be given in the presence of a sudden increase in state entropy or BIS above the index value of 65 Group 3: standard practice (blood pressure, heart rate, sweating, tear production, movement) Propofol increased in steps of 1 mg/kg/hour as necessary for clinical parameters During maintenance of anaesthesia, all patients assessed for signs of inadequate anaesthesia, hypotension or bradycardia Commencement of monitoring: in operating room Further details unclear In group 3 both BIS and E-Entropy monitors were covered behind a curtain; in the BIS and E-Entropy group, either only the BIS monitor or only the E-Entropy module was uncovered Length of experience/training of anaesthetist: 'experienced anaesthesiologist'	**Total numbers involved:** 90; group 1, 30; group 2, 30; group 3, 30 Premedication used: midazolam 7.5 mg orally on morning of surgery General anaesthetic used: bolus of 2 mg/kg propofol and a continuous propofol infusion of 6 mg/kg/hour. A propofol bolus of 0.5 mg/kg given in the presence of unexpected somatic intraoperative response Regional anaesthesia used: mentioned in abstract but no further details given Analgesia used: remifentanil infusion at 0.4 µg/kg/minute to induce anaesthesia followed 5 minutes later by propofol Muscle relaxants used: 0.1 mg/kg cis-atracurium to allow tracheal intubation after which remifentanil reduced to 0.08 µg/kg/minute in order to tolerate tube Antinausea drugs used: NR Other drugs used: 0.3 ml of i.v. vasopressor (Akrinor, 1 ml contains 100 mg cafedrine and 5 mg theodrenaline to treat hypotension). 0.5 mg atropine (to treat brachycardia) Type of surgery: orthopaedic of upper or lower extremity Duration of surgery: NR Duration of GA (minutes), mean (SD): group 1 = 123.7 (44.6); group 2 = 100.0 (30.7); group 3 = 119.5 (50.6) **Inclusion criteria:** ASA I, II or III adults 18–80 years undergoing minor surgery expected to last at least 1 hour **Exclusion criteria:** history of disabling central nervous or cerebrovascular disease, hypersensitivity to opioids or substance abuse, or treatment with opioids or any psychoactive medication **Baseline measurements:** Sex (male) *n* (%): group 1 = 15/25 (60%); group 2 = 18/27 (67%); group 3 = 15/27 (56%) Age (years), mean (SD): group 1 = 58.1 (14.2); group 2 = 50.6 (15.7); group 3 = 53.6 (18.4)	**Primary outcome:** • Reduction in propofol consumption **Secondary outcome:** • Remifentanil consumption, recovery time, duration of anaesthesia, intraoperative awareness, BIS and E-Entropy values **Length of follow-up:** third postoperative day for awareness **Methods of assessing outcomes:** Method of assessing reduction in propofol consumption not reported End of surgery defined as the final surgical suture Recovery from anaesthesia assessed by measuring time between last suture and spontaneous opening of eyes allowing extubation Aldrete score evaluated at extubation Modified Aldrete score for assessing discharge from PACU Intraoperative awareness by 'standardised interview' (first and third day postoperative days) (Nordström *et al.*[96])

NR, not reported.

Reviewer 1: JB Reviewer 2: JS			
Reference and design	**Technology**	**Participants**	**Outcome measures**
		Ethnic groups, *n* (%): NR Weight (kg), mean (SD): group 1 = 76.4 (16.4); group 2 = 82.4 (15.7); group 3 = 76.7 (14.1) ASA grade, I/II/III: group 1 = 4/15/6; group 2 = 10/16/1; group 3 = 10/10/7 Risk factors for awareness: NR Comorbidities: NR **Losses to follow-up:** none **Place of anaesthetic administration:** premedication prior to operating theatre; GA initiated in operating theatre	

NR, not reported.

Outcome	Group 1: E-Entropy (*n* = 25)	Group 2: BIS (*n* = 27)	Group 3: SP (*n* = 27)	*p*-value
Intraoperative awareness/recall	0	0	0	
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR	
Time (minutes) to emergence from anaesthesia, mean (SD) NB. Abstract states this is time to extubation	9.2 (3.9)	6.8 (2.9)	7.3 (2.9)	$p = 0.023$ Group 1 vs group 2 NS (no *p-value* given) for group1/2 vs group 3
Time (minutes) to extubation	NR	NR	NR	
Time (minutes) to discharge to/from the recovery room	NR	NR	NR	
Anaesthetic consumption				
Propofol (μg/kg/minute), mean (SD)	106 (24)	104 (20)	101 (22)	$p = 0.27$ Group 1/2 vs group 3
Remifentanil (μg/kg/minute), mean (SD)	0.08 (0.02)	0.08 (0.02)	0.09 (0.02)	$p = 0.56$
Bolus of propofol following rise in BIS or Entropy (state entropy) above 65 or sudden unexpected somatic response, *n*	12	8	10	
HRQoL	NR	NR	NR	
Nausea/vomiting/antisickness drugs	NR	NR	NR	
Pain/pain-relieving drugs	NR	NR	NR	
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR	
Mortality	NR	NR	NR	

NR, not reported, SP, standard practice.

Additional results/comments

Aldrete scores (10/10) at extubation were group 1 = 8.4 (SD 0.6), group 2 = 8.6 (SD 0.5), group 3 = 8.8 (SD 0.4); group 1 vs group 3 $p = 0.045$

Aldrete scores similar 1 minute after extubation

Various E-Entropy and BIS values reported for all three groups; differences between groups not significant

Methodological comments

Allocation to treatment groups: randomised by drawing lots from a closed box

Allocation concealment: NR

Blinding: NR

Analysis by ITT: no

Comparability of treatment groups at baseline: no differences between groups in age, weight and height by analysis of variance; not reported for sex and ASA status

Method of data analysis: normally distributed data compared with between-group analysis of variance and Tukey's HSD (honestly significant difference) post hoc test if global analysis of variance result was significant; a covariance analysis of variance was performed for 'recovery time' and the covariate 'duration of anaesthesia'. Data not normally distributed compared using Kruskal-Wallis analysis

Sample size/power analysis: calculated that at least 25 patients had to be investigated in each group to detect a reduction of 20% in propofol consumption with a standard deviation of 20% in propofol consumption in each group with a type I error of 0.05 and a statistical power of 0.86

Attrition/dropout: patients excluded from analysis because of insufficient regional anaesthesia or EEG data loss were group 1 = 5, group 2 = 3, group 3 = 3

General comments

Generalisability: to separate hypnotic and analgesic components of anaesthesia, all patients received regional anaesthesia catheters for intra- and postoperative pain control prior to investigation (i.e. pain perception completely blocked), which could limit generalisability. Also more than one type of surgery was included and more than one regional anaesthesia technique that might contribute to different levels of analgesia. Authors state that similar results may not have been obtained with less experienced anaesthetists. Results applicable to adult patients receiving i.v. GA (and regional anaesthesia) assumed not to have significant morbidities

Intercentre variability: NA

Conflict of interests: NR

NA, not applicable; NR, not reported.

Domain	Reviewer's judgement (state: low/high/unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Low	Drawing lots
Allocation concealment	Unclear	No details reported
Performance bias		
Blinding of participants and personnel	Unclear	Monitors covered as appropriate
Detection bias		
Blinding of outcome assessment	Unclear	No details
Attrition bias		
Incomplete outcome data	High	Group 1, 17% patients excluded from analysis; group 2 and group 3, 10%. Not balanced between groups, although reasons similar across groups
Reporting bias		
Selective reporting	Low	No evidence of selective reporting
Other bias		
Other sources of bias		

Gruenewald *et al.*

Reviewer 1: GF Reviewer 2: JS

Reference and design	Technology	Participants	Outcome measures
Author: Gruenewald *et al.*[55] **Year:** 2007 **Study design:** RCT **Number of centres:** 1 (not explicitly stated) **Country:** Germany **Sponsor:** GE Healthcare supplied the M-Entropy module and electrodes	**Group 1:** E-Entropy + standard practice S/5™ M-Entropy module (GE Healthcare); BIS XP monitor (Aspect Medical Systems Inc.); anaesthetist viewed only the entropy monitor Target device/index value: 40–60 for state entropy (>60 acceptable in final 15 minutes of surgery); <10 for response-state entropy difference Commencement of monitoring: prior to induction of anaesthesia, after arrival in the operating theatre **Group 2:** standard practice only Dosage adjustments of anaesthesia at the discretion of the anaesthetist based on standard clinical signs (hypertension (blood pressure >120% of baseline), hypotension (blood pressure<80% of baseline), tachycardia (>90 beats/minute), bradycardia (heart rate <80% of baseline), somatic arousal (coughing, chewing, grimacing), somatic response (purposeful movement) Also monitored by same entropy and BIS devices as group 1, but the monitor screen was covered to obscure the processed EEG parameters Both groups: anaesthesia was guided to achieve rapid recovery Length of experience/training of anaesthetist: stated only that anaesthesia was supervised by an experienced staff anaesthetist	**Total numbers involved:** 72; group 1, 37; group 2, 35 Premedication used: oral benzodiazepine (dipotassium chlorazepate) 20 mg; midazolam 7.5 mg General anaesthetic used: Induction: Propofol 2 mg/kg; remifentanil 0.3–0.5 µg/kg/minute Maintenance: propofol and remifentanil (dose adjusted according to entropy or clinical signs) Regional anaesthesia used: none reported Analgesia used: piritramide 0.1 mg/kg 15 minutes before end of surgery Muscle relaxants used: rocuronium 0.6 mg/kg Antinausea drugs used: none reported Other drugs used: hypotension and bradycardia were managed where appropriate with unspecified pharmacologic agents (dose not reported) Type of surgery: routine elective gynaecological laparoscopy Duration of surgery: ≥1 hour Duration of GA, minute, mean ± SD: group 1, 110 ± 39; group 2, 111 ± 46 **Inclusion criteria:** NR (implied adult female population) **Exclusion criteria:** pregnancy, neurological or neuromuscular disease, use of CNS-active medication, abuse of alcohol or illicit drugs **Baseline measurements:** Sex (male) *n* (%): 0 (0) Age (years) mean ± SD: group 1, 38 ± 9; group 2, 33 ± 9 Ethnic groups, *n* (%): NR Weight (kg) mean ± SD: group 1, 68 ± 15; group 2, 68 ± 13 ASA grade 1/2, *n*: group 1, 14/23; group 2, 11/24 Risk factors for awareness: NR Comorbidities: NR **Losses to follow-up:** NR **Place of anaesthetic administration:** NR	**Primary outcomes:** • Recovery time (from discontinuation of propofol and remifentanil to eye-opening) **Secondary outcomes:** • Intraoperative awareness • Pain, nausea, vomiting • Anaesthetic consumption • Device values (BIS, state entropy, response entropy, state-response entropy difference); • Haemodynamic variables • Somatic responses (purposeful movement) • Cumulative probability of emergence • Patient satisfaction **Length of follow-up:** on arrival in the recovery room (Observer Assessment of Alertness and Sedation scale, nausea and vomiting, and pain questionnaires), and 24 hours post surgery (memory or awareness and satisfaction) **Methods of assessing outcomes:** Intraoperative awareness: Questions about memory or awareness during the ward, induction room, surgery, extubation or recovery room stages Postoperative pain rating: 0–10 scale PONV: assessed by unspecified questions Patient satisfaction: 0–100 scale (100 = totally satisfied) Awareness and satisfaction outcomes assessed by patient interview by an anaesthesiologist blinded to the treatment groups Method of assessing anaesthetic consumption not reported

NR, not reported.

Outcome	Group 1 (entropy + standard practice)	Group 2 (standard practice only)	*p*-value
Intraoperative awareness/recall			
Patients reporting awareness during the procedure when assessed at 24 hours post surgery, *n* (%)	0 (0)	1 (2.8)[a]	NR
Stated no difference between groups in awareness or explicit memory assessed 24 hours post surgery (no further quantitative data provided)			
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time (minutes) to emergence from anaesthesia			
Median [interquartile] (range) time to eye-opening	3 [1–5] (0–9)	4 [3–6] (0–14)	NS
Time (minutes) to extubation	NR	NR	NR
Time (minutes) to discharge to/from the recovery room	NR	NR	NR
Anaesthetic consumption (induction + maintenance; μg/kg/minute), mean (SD)			
Propofol	81 ± 22	95 ± 14	<0.01
Remifentanil	0.46 ± 0.08	0.39 ± 0.08	<0.001
HRQoL	NR	NR	NR
Nausea/vomiting			
Nausea and vomiting, *n* (%) (on arrival in recovery room)	15 (41)	13 (37)	NS
Antisickness drugs: none reported			
Pain			
Median [interquartile] (range) pain intensity score (on arrival in recovery room)	6 [4–7] (2–10)	4 [3–5] (1–10)[b]	0.03
Pain-relieving drugs			
Stated analgesia (piritramide) did not differ between groups (no quantitative data reported)			
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

NR, not reported; NS, not statistically significant ($p \geq 0.05$).

a Implied this was a female patient who did not report feeling any pain.

b As reported with the original data: meaning not stated.

Additional results/comments (e.g. early response factors, QoL)

Patients in group 2 had significantly more hypertension, hypotension, tachycardia, bradycardia and somatic responses (purposeful movements) compared with those in group 1 (47 vs 27 total events, respectively; $p<0.01$). However, the incidence of purposeful movement alone (15 vs 18 total events respectively) did not differ significantly ($p\geq0.05$) between group 2 and group 1

In addition to the emergence data above, cumulative probability of non-emergence was reported in a Kaplan–Meier survival analysis graph (data not extracted)

Median [interquartile] (range) patient satisfaction score 24 hours post surgery: group 1: 93 [80–100] (50–100); group 2: 90 [80–100] (50–100); difference not statistically significant ($p\geq0.05$)

Three patients in group 2 and one patient in group 1 had EEG-derived variables that were considered out of range after skin incision (no further explanation provided)

Methodological comments

Allocation to treatment groups: randomisation to group 1 or group 2 was done by opening a sealed envelope. Sequence generation method and nature of the envelope contents not reported

Allocation concealment: sealed envelope used, not stated whether or not opaque

Blinding: Observer Assessment of Alertness and Sedation Scale, PONV, pain, and recall questions were completed by patient interview by an anaesthesiologist who was blinded to the treatment groups. Postoperative care was supervised by a recovery room nurse blinded to treatment groups. However, stated that entropy and standard practice guidance could not be performed in a blinded fashion

Analysis by ITT: stated that all patients were included into the final analysis

Comparability of treatment groups at baseline: patients in group 1 had mean age 5 years older than group 2; group 1 had a slightly higher ratio of ASA class 1 to class 2 (i.e. slightly less severe illness rating) than group 2. Height (not extracted) and weight were similar in the two groups. Ethnicity not reported. Stated that there were no significant differences in patients' characteristics (*p*-values not reported)

Method of data analysis: t-tests for normally distributed data; Mann–Whitney *U*-tests for non-normally distributed data; repeated measures analysis of variance 'as appropriate' (no further details given). Distribution of emergence times by study group compared using Kaplan–Meier log-rank survival analysis (calculating the cumulative probability of patients remaining unconscious after discontinuation of the anaesthetic drugs)

Sample size/power analysis: sample size of 34 based on a previous study by Kreuer *et al.*,[63] assuming a difference in emergence (eye-opening) of 3 minutes, an error of 0.05 and 90% power. Study was powered for time to eye-opening; stated that there were too few subjects to show a significant effect on intraoperative awareness, given the low incidence rate

Attrition/dropout: NR

General comments

Generalisability: women-only study, mid-30s age group, with ASA score <3. Population does not appear to be at high risk of intraoperative awareness

Intercentre variability: NA; appears to be a single centre

Conflict of interests: none explicitly reported, but the M-Entropy module and electrodes were provided by the module manufacturer

NA, not applicable; NR, not reported.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Unclear	No information given
Allocation concealment	Unclear	Sealed envelopes, not stated whether or not opaque and sequentially numbered
Performance bias		
Blinding of participants and personnel	Unclear	Group 2 anaesthesiologists were blinded to entropy values but group 1 anaesthesiologists were not blinded to clinical practice guidelines; authors stated that entropy and standard practice guidance could not be performed in a blinded fashion, so bias cannot be totally excluded (relevant to performance bias as unclear how much of group 2 intervention was also received by group 1 patients)
Detection bias		
Blinding of outcome assessment	Unclear	Anaesthesiologist who interviewed patients for awareness and satisfaction was blinded to the treatment groups; not reported whether or not assessors of recovery time and anaesthesia consumption were blinded
Attrition bias		
Incomplete outcome data	Unclear	Attrition not reported
Reporting bias		
Selective reporting	Low	All outcomes mentioned in the methods section were reported in the results

Kamal

Reviewer 1: JB Reviewer 2: JS

Reference and design	Technology	Participants	Outcome measures
Author: Kamal *et al.*[48] **Year:** 2009 **Study design:** RCT **Number of centres:** 1 **Country:** Egypt **Sponsor:** not stated	**Group 1:** BIS plug-in modules connected to monitor model A-2000 (Aspect medical Systems, Newton, MA, USA). Software program Datex-Ohmeda S/5 Collect (v4.0) Target BIS index: 50–60. If patient exhibited hypertension or tachycardia treatment depended on BIS value – if BIS >60 then sevoflurane was increased If BIS in target range fentanyl 25–50 µg i.v. given; if BIS<50 sevoflurane decreased and patient checked for lack of analgesia If lack of analgesia fentanyl 25–50 µg i.v. given; if no lack of analgesia labetalol 5–10 mg i.v. given at end of surgery BIS 55–70 to facilitate recovery **Group 2:** standard clinical practice and such that provides early recovery If patient showed hypertension (mean arterial blood pressure >25% above baseline) and tachycardia (heart rate >90 beats/minute) anaesthesia was deepened by increasing inspired sevoflurane or adjusting fentanyl 25–50 µg i.v. or labetalol 5–10 mg i.v. according to anaesthesiologist's discretion Commencement of monitoring: all patients monitored; place and time not explicitly stated In group 2 the monitor display was customised to make BIS values invisible to the attending anaesthesiologist Length of experience/ training of anaesthetist: not stated	**Total numbers involved:** 60; group 1 = 30; group 2 = 30 Premedication used: none used General anaesthetic used: Propofol 1–2 mg/kg i.v. and fentanyl 2–3 µg/kg i.v. (induction) Sevoflurane and 50% nitrous oxide with oxygen 2 l/ minute (continued) Nitrous oxide discontinued, sevoflurane adjusted for BIS index in group 1 and as usual practice in group 2 (10 minutes before last stitch) Sevoflurane discontinued (end of skin closure, beginning of recovery period) Regional anaesthesia used: none used Analgesia used: not stated Muscle relaxants used: atracurium 0.5 mg/kg i.v. Intermittent boluses of atracurium 0.2–0.3 mg/kg i.v. Antinausea drugs used: NR Other drugs used: ephedrine 3–6 mg i.v. or phenylephrine 20–100 µg i.v. (for hypotension). Atropine 0.02 mg/kg i.v. (for bradycardia). Glycopyrate 0.01 mg/kg and neostigmine 0.05 mg/kg i.v. 5 minutes before discontinuation of anaesthesia (to reverse residual neuromuscular blockade) Type of surgery: elective moderate abdominal surgery Duration of surgery (minutes), mean (SD): group 1 = 91.7 (11.3); group 2 = 85.8 (17.4) Duration of GA (minutes), mean (SD): group 1 = 111.7 (14.6); group 2 = 108.7 (10.5) **Inclusion criteria:** ASA I, II, III adults 45-60 years undergoing surgery with expected durations of at least 2 hours **Exclusion criteria:** history of any disabling central nervous or cerebrovascular disease, hypersensitivity to opioids, substance abuse, treatment with opioids or any psychoactive medication and a BMI >40 kg/m^2 **Baseline measurements:** Sex (male), *n* (%): group 1 = 18 (62%); group 2 = 20 (71%) Age (years), mean (SD): group 1 = 51.6 (7.4); group 2 = 52.1 (5.2) Ethnic groups, *n* (%): NR Weight (kg), mean (SD): group 1 = 87.6 (8.2); group 2 = 91.4 (6.5) ASA grade: not reported by group Risk factors for awareness: NR Comorbidities: NR **Losses to follow-up:** none **Place of anaesthetic administration:** NR	**Primary outcomes:** • Not specified **Secondary outcomes:** • Not specified **Outcomes:** • Recovery times (awakening, tracheal extubation, orientation, arrival at PACU, discharge from PACU) • BIS index values • Anaesthetic drug consumption **Length of follow-up:** third postoperative day for awareness **Methods of assessing outcomes:** Sevoflurane used calculated using Dion's formula Recovery starting point was immediately after last surgical stitch Aldrete score for assessment of discharge from PACU (>9), at 15-minute intervals by research assistant blinded to group assignment Awakening defined as eye-opening Orientation to place, person and time For intraoperative awareness patients visited on first, second and third day postoperatively and questioned for recall of events, hearing vague sounds, feeling surgical instruments or dressing application, or dreaming

NR, not reported.

Outcome	Group 1 ($n = 29$)	Group 2 ($n = 28$)	p-value
Intraoperative awareness/recall	0	0	
Patient distress and sequelae resulting from perioperative awareness	NR	NR	
Time (minutes) to emergence from anaesthesia after termination of anaesthesia (awakening eye-opening)	4.1 (1.6)	4.4 (1.9)	NS
Time (minutes) to extubation	4.3 (2.1)	4.8 (2.3)	NS
Time (minutes) to discharge to/from the recovery room			
Arrival at PACU	9.4 (1.9)	14.1 (2.8)	$p<0.01$
PACU discharge (minutes)	53.9 (14.7)	78.6 (21.5)	$p<0.01$
Anaesthetic consumption			
Sevoflurane (ml), mean (SD)	5.7 (1.9)	8.4 (2.3)	$p<0.01$
End-tidal sevoflurane (vol%), mean (SD)	0.43 (0.3)	0.59 (0.1)	$p\leq0.01$
Propofol (mg), mean (SD)	161.7 (27.5)	157.9 (35.8)	NS
Fentanyl (μg), mean (SD)	383.7 (62.6)	389.4 (41.5)	NS
HRQoL	NR	NR	
Nausea/vomiting/antisickness drugs	NR	NR	
Pain/pain-relieving drugs	NR	NR	
Other morbidity (e.g. cognitive dysfunction)	NR	NR	
Mortality	NR	NR	

NR, not reported; NS, not statistically significant..

Additional results/comments (e.g. early response factors, QoL)

Orientation (minutes) group 1 = 7.4 (1.5), group 2 = 11.2 (1.9), $p<0.01$

Average BIS index values were statistically significantly lower in group 2 than group 1 during surgery and during anaesthesia (both $p<0.01$)

Patient disorientation (%) after discontinuation of inhalational anaesthetic agents was statistically significantly higher at 15 and 20 minutes postoperatively in group 2 than group 1 ($p<0.01$)

Methodological comments

Allocation to treatment groups: randomised (no details reported)

Allocation concealment: no details reported

Blinding: anaesthetists in the control group (group 2) were blinded to the BIS values. No other blinding reported

Analysis by ITT: no, as three patients not included in analysis

Comparability of treatment groups at baseline: authors state groups comparable but no *p-value*s reported (although results suggest groups are comparable)

Method of data analysis: comparison between groups performed using Mann–Whitney *U*-test. Categorical data were compared using chi-squared test

Sample size/power analysis: NR

Attrition/dropout: as above. One patient in group 1 was desaturated intraoperatively necessitating discontinuation of nitrous oxide, and two in group 2 received excessive fentanyl near the end of surgery

General comments

Generalisability: authors state that anaesthetists vary in the way and timing of reducing anaesthetic drug administration towards the end of surgery and this could have an effect on results (i.e. starting point of recovery process variable). Results applicable to adults receiving inhaled anaesthesia for moderate abdominal surgery

Intercentre variability: NA

Conflict of interests: NR

NA, not applicable; NR, not reported.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Unclear	No method reported
Allocation concealment	Unclear	No method reported
Performance bias		
Blinding of participants and personnel	Unclear	Reported that anaesthetists for control group were blinded to BIS values
Detection bias		
Blinding of outcome assessment	Unclear	Only reported that research assistant collecting Aldrete score was blinded
Attrition bias		
Incomplete outcome data	Low	Only three patients not included in analysis (see above)
Reporting bias		
Selective reporting	Low	No evidence of selective reporting
Other bias		
Other sources of bias		

Kerssens *et al.*

Reviewer 1: GF	Reviewer 2: JS		
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Kerssens *et al.*[49] **Year:** 2009 **Study design:** RCT **Number of centres:** NR **Country:** USA **Sponsor:** lead author received an educational grant in support of her salary from Aspect Medical Systems Inc.; one co-author was a paid consultant to Aspect Medical Systems Inc.; stated that Aspect Medical Systems did not financially support the study	**Group 1:** BIS, BIS monitor (XP, algorithm 3.4; Aspect Medical Systems Inc.) Target device/ index value: 50–60 Commencement of monitoring: NR **Group 2:** standard practice Standard clinical signs such as heart rate and blood pressure-guided anaesthesia BIS was recorded but not available to the attending clinician for drug dosing Length of experience/ training of anaesthetist: NR	**Total numbers involved:** 128 Number randomised: group 1, 67; group 2, 61 Premedication used: stated benzodiazepines were not given to any patients pre- or intraoperatively General anaesthetic used: Induction: propofol 2 mg/kg Maintenance: sevoflurane in oxygen using standard ventilation parameters (not specified) Regional anaesthesia used: used only for postoperative pain management Analgesia used: fentanyl 3 µg/kg (induction); 50–100 µg (maintenance) Muscle relaxants used: vecuronium bromide 0.1 mg/kg with additional doses as necessary (tracheal intubation) Antinausea drugs used: none reported Other drugs used: esmolol 0.5 mg/kg for hypertension and phenylephrine 100 µg for hypotension as needed Type of surgery: major orthopaedic surgery (hip or knee replacement) Duration of surgery: NR Duration of GA, minutes, mean ± SD: group 1, 126 ± 51; group 2, 112 ± 48 **Inclusion criteria:** patients aged ≥ 18 years scheduled for hip or knee replacement surgery, primary or revision, under GA **Exclusion criteria:** medical history or status that could compromise or skew EEG recordings; history of illicit drug use; antipsychotic medication treatment; head trauma resulting in the loss of consciousness; CNS disorders (e.g. epilepsy); persons scoring < 24 on the preoperatively administered MMSE (reference cited); severe visual or auditory handicaps; non-fluent-English speakers Baseline measurements (only reported for subset of patients assessed after attrition: group 1, $n = 62$; group 2, $n = 47$, but stated that characteristics of the full sample were similar)	**Main outcomes:** • Word recognition memory (implicit recall) • Recall assessment (explicit recall) **Secondary outcomes:** • Anaesthetic consumption • BIS device values **Length of follow-up:** 6 hours post surgery **Methods of assessing outcomes:** physiological parameters, BIS, end-tidal gas concentrations (every 5 seconds) and vital signs (every 3 seconds) were automatically recorded to a computer using Rugloop (Demed, Belgium) Recall assessment: 6 hours after surgery, consisting of five questions (listed in the paper, similar to Brice interview questions), with additional questions asked as necessary Recognition memory test: conducted after recall assessment. An auditory test in which sequences of predetermined neutral words was played to patients through headphones (rationale of the word selection and language characteristics reported). Word presentation typically started 15 minutes after induction and lasted approximately 42 minutes. The memory test involved playing predetermined combinations of words that had been used during anaesthesia, and distractor words, to patients though headphones. Patients were instructed to listen to each test sequence and select the word played during surgery, or to guess if necessary (three-alternative forced choice)

NR, not reported.

Reviewer 1: GF	Reviewer 2: JS		
Reference and design	**Technology**	**Participants**	**Outcome measures**
		Sex (male), *n* (%): group 1, 28 (45); group 2, 16 (34) Age (years) mean ± SD: group 1, 61.2 ± 11.4; group 2, 63.9 ± 11.8 Ethnic groups, *n* (%): NR Weight (kg) mean ± SD: group 1, 87.9 ± 18.9; group 2: 84.4 ± 14.8 BMI (kg/m^2), mean ± SD: group 1, 30.2 ± 5.6; group 2, 28.9 ± 3.7 ASA grade: ASA I-II: about 50%; ASA III: 50%; stated no differences between groups Baseline data were also reported for MMSE and STAI scores (values were similar in both study groups) Risk factors for awareness: not explicitly reported but population undergoing major orthopaedic surgery and appears to have BMI around 30 kg/m^2 Comorbidities: none reported (patients with comorbidities were excluded) **Losses to follow-up:** attrition reported, with reasons, both pre and post randomisation **Place of anaesthetic administration:** NR	

NR, not reported. STAI, State–Trait Anxiety Inventory.

Outcome	Group 1: BIS ($n = 67$)	Group 2: Standard practice ($n = 61$)	p-value
Intraoperative awareness/recall			
Recall of time period between falling asleep and waking up from anaesthesia, *n* (%)	2 (3.0)	1 (1.6)	Not tested (outcome not powered)
Memory recall: probability of postoperatively selecting a word presented during anaesthesia (target) or not presented during anaesthesia (distractor), mean ± SD			
Target	0.371 ± 0.132	0.323 ± 0.132	NR[a]
Distractor	0.315 ± 0.117	0.338 ± 0.119	NR[a]
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time to emergence from anaesthesia	NR	NR	NR
Time to extubation	NR	NR	NR
Time to discharge to/from the recovery room	NR	NR	NR
Anaesthetic consumption, end-tidal gas concentration (%) mean ± SD			
Maintenance phase	1.31 ± 0.29[b]	1.56 ± 0.29[c]	< 0.001
During word presentation	1.30 ± 0.31[b]	1.60 ± 0.37[c]	NS[d]
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs	NR	NR	NR
Pain/pain-relieving drugs – fentanyl analgesia			
Preoperative (μg/kg), mean ± SD	0.27 ± 0.43[b]	0.40 ± 0.47[c]	NS[d]
Intraoperative (μg/kg/hour), including induction dose	2.83 ± 1.04[b]	2.70 ± 1.18[c]	NS[d]
Postoperative (μg/kg)	0.47 ± 0.66[b]	0.55 ± 1.10[c]	NS[d]
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

NR, not reported; NS, not statistically significant.

a See additional comments for interpretation of within-group differences.

b Reported for post-attrition subgroup ($n = 62$).

c Reported for post-attrition subgroup ($n = 47$).

d Authors only reported p-values that were considered significant ($p < 0.05$); reviewers have assumed that comparisons reported without p-values were not significant (i.e. $p \geq 0.05$).

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Low	Random assignment using a computer-generated list
Allocation concealment	Unclear	No information provided
Performance bias		
Blinding of participants and personnel	Unclear	BIS was recorded in group 2 but not available to the attending clinician for drug dosing, but unclear whether or not anaesthetist was still aware of group assignment
Detection bias		
Blinding of outcome assessment	Low	Outcome assessors (two of the study authors) were blinded to study group allocation. Note that the method of blinding was not stated; hence, the likelihood of blinding being broken cannot be assessed
Attrition bias		
Incomplete outcome data	Unclear	Attrition with reasons was reported, but not separately by study group
Reporting bias		
Selective reporting	Unclear	STAI scores were reported only for baseline; stated that postoperative STAI score results can be found elsewhere, together with results of a depression questionnaire, but no references were provided

STAI, State–Trait Anxiety Inventory.

Kreuer *et al.*

Reviewer 1: JB Reviewer 2: JS			
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Kreuer *et al.*[64] **Year:** 2005 **Study design:** RCT **Number of centres:** one **Country:** Germany **Sponsor:** solely supported by departmental funding	**Group 1:** BIS A-2000 monitor (version XP) Desflurane maintenance anaesthesia adjusted to target value of 50 BIS 15 minutes before expected end of surgery desflurane adjusted to target value of BIS 60 **Group 2:** Narcotrend monitor (software version 2.0 AF). Desflurane maintenance anaesthesia adjusted to target value of D_0 15 minutes before expected end of surgery desflurane adjusted to target value of C_1 In groups 1 and 2: if anaesthesia judged inadequate, although target value achieved, infusion rate of remifentanil increased by 0.05 µg/kg/minute **Group 3:** standard anaesthetic practice protocol If anaesthesia inadequate desflurane concentration increased in steps of 0.5% volume as necessary. If insufficient remifentanil increased by 0.05 µg/kg/minute Hypotension treated with desflurane concentration reduced in steps of 0.5 vol%. Desflurane reduced 15 minutes before end of surgery as much as judged clinically possible without intraoperative awakening Inadequate anaesthesia in all patients defined as hypertension, tachycardia or patient movement, eye-opening, swallowing, grimacing, lacrimation or sweating Commencement of monitoring: in operating theatre Both monitors covered behind curtain for group 3 and invisible to anaesthesiologist; in groups 1 and 2 either only the Narcotrend or only the BIS monitor was uncovered Length of experience/training of anaesthetist: one experienced anaesthesiologist	**Total numbers involved:** 120; group 1 = 40; group 2 = 40; group 3 = 40 Premedication used: midazolam 7.5 mg orally in the evening and on the morning before surgery General anaesthetic used: Induction: remifentanil infusion 0.4 µg/kg/minute, 5 minutes later 2 mg/kg propofol for hypnosis After intubation remifentanil reduced to constant rate of 0.2 µg/kg/minute Desflurane adjusted according to EEG target values or clinical variable 15 minutes before expected end of surgery desflurane reduced in all groups to facilitate rapid emergence from anaesthesia; remifentanil infusion rate remained unchanged throughout end of surgery Regional anaesthesia used: NR Analgesia used: 100 ml infusion of 0.9% NaCl + metamizol 25 mg/kg for postoperative pain relief Muscle relaxants used: 0.5 mg/kg atracurium Antinausea drugs used: NR Other drugs used: hypotension treated with an i.v. vasopressor (Akrinor, 1 ml contains 100 mg of cafedrine and 5 mg of theodrenaline) given at dose chosen by investigator. Atropine 0.5 mg for bradycardia Type of surgery: minor orthopaedic surgery Duration of surgery: NR Duration of GA (minutes), mean (SD): group 1 = 113 (57); group 2 = 122 (50); group 3 = 125 (51) (reported in table 1, although text states this is duration of surgery) **Inclusion criteria:** ASA I, II, III adults 18–80 years scheduled for minor orthopaedic surgery expected to last at least 1 hour **Exclusion criteria:** history of disabling central nervous or cerebrovascular disease, hypersensitivity to opioids or substance abuse, or a treatment with opioids or any psychoactive medication **Baseline measurements:** Sex (male), *n* (%): group 1 = 20/40 (50); group 2 = 20/40 (50); group 3 = 20/40 (50) Age (years), mean (range): group 1 = 46.5 (14.1); group 2 = 44.7 (15.6); group 3 = 43.6 (16.0)	**Primary outcome:** • Time taken to spontaneous opening of eyes **Secondary outcome:** • Not explicitly stated (times to tracheal extubation and arrival at PACU, consumption of desflurane) **Length of follow-up:** third day postoperative for recall **Methods of assessing outcomes:** end of surgery defined as final surgical suture when anaesthesia was stopped Emergence from anaesthesia assessed by measuring times to spontaneous opening of eyes, tracheal extubation and arrival at PACU Desflurane vaporiser weighed before and after anaesthesia to calculate consumption Intraoperative recall assessed by interview in PACU and on first and third postoperative days

Reviewer 1: JB Reviewer 2: JS			
Reference and design	**Technology**	**Participants**	**Outcome measures**
		Ethnic groups, *n* (%): NR Weight (kg), mean (SD): group 1 = 79.3 (16.2); group 2 = 83.6 (18.3); group 3 = 79.0 (17.4) ASA grade, *n*, I/II/III: group 1 = 7/30/3; group 2 = 13/23/4 group 3 = 11/27/2 Risk factors for awareness: NR Comorbidities: NR **Losses to follow-up:** NR **Place of anaesthetic administration:** in the operating room	

NR, not reported.

Outcome	Group 1 BIS	Group 2 Narcotrend	Group 3 Standard care	*p*-value
Intraoperative awareness/recall	0	0	0	
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR	
Time (minutes) to eye opening, mean (SD)	4.2 (2.1)	3.7 (2.0)	4.7 (2.2)	NS
Reduction compared with standard practice (%)	–10.6	–21.3	NA	
Time (minutes) to extubation, mean (SD)	4.4 (2.2)	3.6 (2.0)*	5.0 (2.4)	*$p<0.05$ Group 2 vs group 3
Reduction compared with standard practice (%)	–12.0	–28.0	NA	
Time (minutes) to discharge to PACU (minutes), mean (SD)	8.4 (2.4)*	8.0 (1.9)*	9.4 (2.4)	*$p<0.05$ Group 1 and 2 vs group 3
Reduction compared with standard practice (%)	–10.6	–15.0	NA	
Anaesthetic consumption per patient				
Desflurane mg, mean (SD)	4861.7 (2948.3)	4655.9 (2891.7)	5547.3 (2396.4)	NS
Reduction compared with standard practice (%)	–12.4	–16.1	NA	
Desflurane mg/minute, mean (SD)	416.2 (99.1)*	374.6 (124.2)*	443.6 (71.2)	*$p<0.05$
Reduction compared with standard practice (%)	–6.2	–15.7	NA	
Normalised remifentanil infusion rates (µg/kg/minute), mean (SD)	0.22 (0.05)	0.22 (0.06)	0.23 (0.07)	NS
HRQoL	NR	NR	NR	
Nausea/vomiting/antisickness drugs	NR	NR	NR	
Pain/pain-relieving drugs	NR	NR	NR	
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR	
Mortality	NR	NR	NR	

NA, not applicable; NR, not reported; NS, not statistically significant.
The asterisks refer to a statistical significance of 0.05.

Additional results/comments (e.g. early response factors, QoL)

End-tidal desflurane concentration reported to be significantly smaller with BIS and Narcotrend compared with standard practice (graph only)

Mean arterial blood pressure at various times points during anaesthesia similar between groups

Vasopressor was necessary in 19 BIS patients, in 19 Narcotrend patients and in 17 standard practice patients

Five patients in each group needed 0.5 mg atropine for treatment of bradycardia

Mean BIS values in the Narcotrend group were higher than those in the BIS group and standard care group (but not statistically significantly so at all time points)

Methodological comments

Allocation to treatment groups: randomised by drawing lots from a closed box

Allocation concealment: no details reported

Blinding: for standard practice group attending anaesthesiologist blinded to EEG readings; in EEG groups either only BIS or only Narcotrend monitor uncovered. Recovery times recorded by blinded investigator. No details reported for desflurane consumption or interview for intraoperative recall

Analysis by ITT: yes

Comparability of treatment groups at baseline: groups reported to be similar at baseline (no statistically significant differences reported)

Method of data analysis: chi-squared test or one-way analysis of variance with Student-Newman-Keuls test for multiple comparisons as appropriate; all tests two-tailed with statistical significance defined as $p < 0.05$. Recovery time to opening of eyes also compared using Kaplan–Meier survival analysis

Sample size/power analysis: 35 patients had to be enrolled in each treatment group to provide 80% power to detect a difference of 1.5 minutes at an $\alpha = 0.05$

Attrition/dropout: none

General comments

Generalisability: observed differences were minimal and not clinically significant. Results applicable to patients receiving GA with desflurane-remifentanil for minor orthopaedic surgery

Intercentre variability: NA

Conflict of interests: funding source stated but no other details reported

NA, not applicable.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Low	Drawing lots
Allocation concealment	Unclear	Method not reported
Performance bias		
Blinding of participants and personnel	Unclear	Not all details reported
Detection bias		
Blinding of outcome assessment	Low	Recovery times recorded by blinded investigator. No details reported for other outcomes
Attrition bias		
Incomplete outcome data	Low	ITT analysis
Reporting bias		
Selective reporting	Low	No evidence of selective reporting
Other bias		
Other sources of bias		

Kreuer *et al.*

Reviewer 1: JB Reviewer 2: JS

Reference and design	Technology	Participants	Outcome measures
Author: Kreuer *et al.*[63] **Year:** 2003 **Study design:** RCT **Number of centres:** one **Country:** Germany **Sponsor:** support solely from departmental sources	**Group 1:** BIS A-2000 monitor (software version 3.0) Propofol TCI continuously adjusted to target value of 50 BIS 15 minutes before end of surgery propofol TCI adjusted to target value of BIS 60 **Group 2:** Narcotrend monitor (software version 2.0 AF) Propofol TCI continuously adjusted to target value of D_0 15 minutes before end of surgery propofol TCI adjusted to target value of C_1 Group 3: standard anaesthetic practice protocol During maintenance all patients were assessed for signs of inadequate anaesthesia (hypertension, tachycardia, movement, eye opening, swallowing, grimacing, lacrimation or sweating), hypotension or bradycardia If anaesthesia inadequate, propofol concentration increased in steps of 0.5 µg/ml as necessary. If insufficient remifentanil increased by 0.05 µg/kg/minute Hypotension treated with propofol concentration reduced in steps of 0.5 µg/ml Propofol reduced 15 minutes before end of surgery as much as judged clinically possible without intraoperative awakening Commencement of monitoring: in operating theatre Both monitors covered behind curtain for group 3 and invisible to anaesthesiologist; in groups 1 and 2 either only the Narcotrend or only the BIS monitor was uncovered Length of experience/ training of anaesthetist: one anaesthesiologist experienced in BIS and Narcotrend monitoring	**Total numbers involved:** 120; group 1 = 40; group 2 = 40; group 3 = 40 Premedication used: 0.15 mg/kg diazepam orally in the evening and on the morning before surgery General anaesthetic used: Induction: remifentanil infusion 0.4 µg/kg/minute, 5 minutes later propofol TCI, initially started at 3.5 µg/ml After intubation remifentanil reduced to constant rate of 0.2 µg/kg/minute Propofol TCI adjusted according to EEG target values or clinical variables 15 minutes before expected end of surgery propofol reduced in all groups to facilitate rapid emergence from anaesthesia; remifentanil infusion rate remained unchanged throughout end of surgery Regional anaesthesia used: NR Analgesia used: 100 ml infusion of 0.9% NaCl + metamizol 25 mg/kg for postoperative pain relief Muscle relaxants used: 0.1 mg/kg cisatracurium Antinausea drugs used: NR Other drugs used: hypotension treated with an i.v. vasopressor (Akrinor, 1 ml contains 100 mg of cafedrine and 5 mg of theodrenaline) given at dose chosen by investigator. Atropine 0.5 mg for bradycardia Type of surgery: minor orthopaedic surgery Duration of surgery: NR Duration of GA (minutes), mean (SD): group 1 = 121.2 (40.9); group 2 = 126.9 (67.7); group 3 = 108.2 (44.2) (reported in *table 1*, although text states this is duration of surgery) **Inclusion criteria:** ASA I, II, III adults 18–80 years scheduled to undergo minor orthopaedic surgery expected to last at least 1 hour **Exclusion criteria:** history of disabling central nervous or cerebrovascular disease, hypersensitivity to opioids or substance abuse, or a treatment with opioids or any psychoactive medication **Baseline measurements:** Sex (male), *n* (%): group 1 = 20/40 (50); group 2 = 20/40 (50); group 3 = 20/40 (50) Age (years), mean (SD): group 1 = 43.8 (4.2); group 2 = 44.8 (15.9); group 3 = 46.1 (14.5) Ethnic groups, *n* (%): NR Weight (kg), mean (SD): group 1 = 78.3 (13.8); group 2 = 76.6 (11.7); group 3 = 82.7 (17.8)	**Primary outcomes:** • Time taken to spontaneous opening of eyes **Secondary outcomes:** • Other outcomes reported – recovery times and consumption of remifentanil and propofol **Length of follow-up:** third day postoperative for recall **Methods of assessing outcomes:** End of surgery defined as final surgical suture when anaesthesia was stopped Emergence from anaesthesia defined as spontaneous opening of eyes, tracheal extubation and arrival at PACU Mean propofol infusion rate normalised to weight was calculated from induction and maintenance doses Intraoperative recall assessed by interview in PACU and on first and third postoperative day

Reviewer 1: JB Reviewer 2: JS			
Reference and design	**Technology**	**Participants**	**Outcome measures**
		ASA grade, *n*, I/II/III: group 1 = 12/25/3; group 2 = 13/24/3; group 3 = 12/24/4 Risk factors for awareness: NR Comorbidities: NR **Losses to follow-up:** NR **Place of anaesthetic administration:** in the operating room	

NA, not reported; TCI, target-controlled infusion.

Outcome	Group 1 BIS	Group 2 Narcotrend	Group 3 Standard care	*p*-value
Intraoperative awareness/recall	0	0	0	
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR	
Time (minutes) to emergence from anaesthesia, mean (SD)	3.5 (2.9)*	3.4 (2.2)*	9.3 (5.2)	*$p<0.001$ Group1/2 vs group 3
Reduction compared with standard practice (%)	–63.4	–62.4	NA	
Time (minutes) to extubation, mean (SD)	4.1 (2.9)*	3.7 (2.2)*	9.7 (5.3)	*$p<0.001$ Group 1/2 vs group 3
Reduction compared with standard practice (%)	–57.7	–61.9	NA	
Time (minutes) to discharge to PACU, mean (SD)	7.0 (3.2)*	6.6 (2.8)*	12.4 (5.7)	*$p<0.001$ Group 1/2 vs group 3
Reduction compared with standard practice (%)	–43.5	–46.7	NA	
Anaesthetic consumption per patient				
Propofol (mg), mean (SD)	720.6 (245.3)*	721.3 (401.2)**	970.5 (384.4)	*$p<0.001$ **$p<0.05$
Reduction compared with standard practice (%)	–25.7	–25.7	NA	
Propofol (mg/kg/hour), mean (SD)	4.8 (1.0)*	4.5 (1.1)*	6.8 (1.2)	*$p<0.001$
Reduction compared with standard practice (%)	–29.4	–33.8	NA	
Normalised remifentanil infusion rates (µg/kg/minute), mean (SD)	0.22 (0.07)	0.21 (0.07)	0.20 (0.07)	ns
HRQoL	NR	NR	NR	
Nausea/vomiting/antisickness drugs	NR	NR	NR	
Pain/pain-relieving drugs	NR	NR	NR	
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR	
Mortality	NR	NR	NR	

NA, not applicable; NR, not reported.
The asterisk(s) refer to a statistical significance of 0.001 (*) or 0.05 (**).

Additional results/comments

Mean arterial blood pressure at various times points during anaesthesia similar between groups

Vasopressor was necessary in significantly more patients ($n = 27$) with standard practice than in Narcotrend ($n = 14$) or in the BIS group ($n = 17$) ($p < 0.05$). The mean drug amount was also significantly higher in the standard practice group

Five patients in each group needed 0.5 mg atropine for treatment of bradycardia

Recovery times were significantly shorter in women than men in the standard practice group with comparable amounts of propofol

Propofol consumption was significantly lower for men than women in the BIS group

BIS values comparable for patients in Narcotrend and BIS groups; significantly lower BIS values were observed in standard practice group vs BIS or Narcotrend group at various time points of anaesthesia

Methodological comments

Allocation to treatment groups: randomised by drawing lots from closed box

Allocation concealment: no details reported

Blinding: for standard practice group attending anaesthesiologist blinded to EEG readings; in EEG groups either only BIS or only Narcotrend monitor uncovered. Recovery times and propofol consumption recorded by a blinded investigator

Analysis by ITT: yes

Comparability of treatment groups at baseline: groups reported to be similar at baseline (no statistically significant differences reported)

Method of data analysis: for nominal data chi-squared test; for numerical data statistical analysis by *t*-test, Mann–Whitney *U*-test, or one-way analysis of variance with Student–Newman–Keuls test for multiple comparisons as appropriate; all tests two tailed with statistical significance defined as $p < 0.05$. Recovery time to opening of eyes also compared using Kaplan–Meier survival analysis

Sample size/power analysis: at least 26 patients had to be enrolled in each treatment group to provide 90% power to detect a difference of 3 minutes at $\alpha = 0.05$

Attrition/dropout: none reported

General comments

Generalisability: Sex differences observed within groups (see above). Results applicable to patients receiving i.v. GA with propofol–remifentanil for minor orthopaedic surgery

Intercentre variability: NA

Conflict of interests: NR

NA, not applicable; NR, not reported.

Domain	Author's judgement (state: low/high/unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Low	Drawing lots
Allocation concealment	Unclear	Method not reported
Performance bias		
Blinding of participants and personnel	Unclear	Not all details reported; anaesthesiologist blinded
Detection bias		
Blinding of outcome assessment	Low	Blinded investigator for recovery times and propofol consumption
Attrition bias		
Incomplete outcome data	Low	ITT analysis
Reporting bias		
Selective reporting	Low	No evidence of selective reporting
Other bias		
Other sources of bias		

Lai *et al.*

Reviewer 1: GF Reviewer 2: JS

Reference and design	Technology	Participants	Outcome measures
Author: Lai *et al.*[59] **Year:** 2010 **Study design:** RCT **Number of centres:** one **Country:** China **Sponsor:** NR	**Group 1:** Narcotrend Narcotrend monitor (MonitorTechnik, Germany), with three-pole Blue sensor (Medicotest, Olstykke, Denmark) (skin impedance reported) Stated that vasoactive agents were used to target the appropriate NT range Target device/index value: Narcotrend (NT) index maintained between D_2 and E_0, then the fentanyl infusion rate was adjusted 10 minutes before end of surgery to target NT values between D_0 and D_1 Commencement of monitoring: not explicitly stated but appears to be the CT room (venue of the surgery) **Group 2:** standard clinical monitoring Monitoring of heart rate (normal = 50–100 b.p.m.), mean arterial pressure (normal = baseline value ± 20%) and body movement Length of experience/ training of anaesthetist: NR	**Total numbers involved:** 40; group 1, 20; group 2, 20 Premedication used: none reported General anaesthetic used (TIVA): Induction: Propofol 3 mg/kg/hour Maintenance: Propofol 4–8 mg/kg/hour Stated anaesthesia was lightened 10 minutes before the end of surgery (group 2; no further details provided) Regional anaesthesia used: none reported (local anaesthetic (lidocaine) used at the puncture site) Analgesia used: Induction: fentanyl 2 μg/kg Maintenance: fentanyl 1 μg/kg as necessary (see below); 10 minutes before end of surgery fentanyl was titrated to NT values between D0 and D1 (group 1) Muscle relaxants used: none (patients maintained spontaneous breathing) Antinausea drugs used: none reported Other drugs used: Tachycardia (> 100 b.p.m.): fentanyl 1 μg/kg, with metoprolol 1 mg added as necessary Hypertension (> 20% above baseline value): urapidil 10–15 mg Body movement: fentanyl 1 μg/kg Bradycardia (< 50 b.p.m.): atropine 0.2–0.5 mg Hypotension (> 20% below baseline value): ephedrine 5–10 mg Note: mentioned for group 1 only that if tachycardia, hypertension or body movement occurred, propofol infusion rate was increased as necessary Type of surgery: microwave coagulation for liver cancer Duration of surgery: NR Duration of GA (minutes) mean ± SD:[a] group 1, 91 ± 30; group 2, 88 ± 31; difference NS **Inclusion criteria:** patients with liver cancer scheduled to undergo microwave coagulation under the guidance of computed tomography (CT) **Exclusion criteria:** neurological or psychiatric problems; hearing defects; alcohol or drug dependence **Baseline measurements:** Sex (male), *n* (%): NR Age, years, mean (range): group 1, 44 (25–69); group 2, 41 (20–70); difference NS Ethnic groups, *n* (%): probably Chinese (NR)	**Outcomes (not stated whether primary or secondary):** • Changes in haemodynamic parameters • Arousal time • Recovery of orientation • Anaesthetic consumption • Postoperative nausea and vomiting • Intraoperative awareness • Postoperative (VASs) **Length of follow-up:** outcomes were assessed within 24 hours after surgery **Methods of assessing outcomes:** intraoperative awareness: stated that this was inquired within 24 hours after the operation, but no details of the method were provided Arousal time: defined as the time between cessation of drugs and the patient being able to open their eyes on command Time for recovery of orientation: defined as the time between a patient opening their eyes on command and the restoration of orientation Restoration of orientation: not defined VAS scores: no explanation of scale provided

Reviewer 1: GF Reviewer 2: JS			
Reference and design	**Technology**	**Participants**	**Outcome measures**
		Weight (kg) mean ± SD:[a] group 1, 60 ± 8; group 2, 60 ± 7; difference NS ASA grade: all patients were grade II to III Risk factors for awareness: none reported Comorbidities: Hypertension, *n* (%): group 1, 3 (15); group 2, 4 (20); difference NS **Losses to follow-up:** none reported; outcome data reported for all randomised patients (*n* = 20 per group) **Place of anaesthetic administration:** not explicitly stated but appears to be the CT room (venue of the surgery)	

b.p.m., beats per minute; CT, computed tomography; NR, not reported; NS, not statistically significant ($p > 0.05$); NT, Narcotrend index; SD, standard deviation.

Outcome[b]	Group 1 (*n* = 20)	Group 2 (*n* = 20)	*p*-value
Intraoperative awareness/recall			
Intraoperative awareness followed up 24 hours post surgery (no methodological details provided), *n* (%)	0 (0)	0 (0)	NA
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time (minutes) to emergence from anaesthesia, mean ± SD			
Arousal time	4.9 ± 2.2	9.5 ± 2.9	<0.01
Duration of orientation recovery	6.6 ± 3.2	12.2 ± 3.5	<0.01
Time to extubation	NA	NA	NA
Time to discharge to/from the recovery room	NR	NR	NR
Anaesthetic consumption			
Propofol dose (mg), mean ± SD[c]	380 ± 35	460 ± 30	<0.01
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs			
Nausea or vomiting reported after surgery, *n* (%)	0 (0)	0 (0)	NA
Pain/pain relieving drugs			
Fentanyl dose, mg, mean ± SD[c]	0.15 ± 0.03	0.13 ± 0.03	0.68
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

NA, not applicable; NR, not reported.

Additional results/comments (e.g. early response factors, QoL)

Stated there were no differences in heart rate or blood pressure between the two groups preoperation, at anaesthesia induction, at the beginning of surgery, at the end of surgery, or at anaesthesia emergence ($p > 0.05$) (data reported in charts, not extracted by reviewer)

Stated that the uses of vasoactive agents (ephedrine, atropine, metoprolol and urapidil) were not statistically different ($p > 0.05$) (no quantitative data reported)

Methodological comments

Allocation to treatment groups: stated random allocation but no details of sequence generation provided

Allocation concealment: NR

Blinding: NR

Analysis by ITT: not explicitly stated, but it appears that there were no withdrawals and that the outcomes data were reported for all randomised patients

Comparability of treatment groups at baseline: sex was not reported. Stated there was no significant difference between the two groups in terms of age, body weight, hypertension ($p > 0.05$)

Method of data analysis: stated that quantitative data were analysed with a chi-squared test and categorical data were analysed with independent *t*-tests or an analysis of variance. No other details of the analysis were reported

Sample size/power analysis: NR

Attrition/dropout: not explicitly reported but there do not appear to have been any dropouts

General comments

Generalisability: liver cancer patients eligible for microwave coagulation. Sex and ethnicity not reported, but appears to be a Chinese population. Early 40s in age, with ASA grade <III, most without concurrent hypertension, receiving TIVA with propofol and fentanyl. No specific risk factors for intraoperative awareness identified

Intercentre variability: NA (one centre)

Conflict of interests: NR

NA, not applicable; NR, not reported.

a Variance parameter not specified; assumed by reviewer to be SD.

b Postoperative VASs reported as an outcome: data not extracted by reviewer as no explanation or interpretation of the scores was provided.

c Not stated whether or not this was the total dose for all phases of anaesthesia.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Unclear	No information provided
Allocation concealment	Unclear	No information provided
Performance bias		
Blinding of participants and personnel	Unclear	No information provided
Detection bias		
Blinding of outcome assessment	Unclear	No information provided
Attrition bias		
Incomplete outcome data	Low	Attrition not explicitly reported, but outcome data appear to have been reported for all randomised patients
Reporting bias		
Selective reporting	Low	No evidence to suggest selective reporting
Other bias		
Other sources of bias	Unclear	The paper was translated from Chinese to English prior to publication. It is unclear whether or not any checks were made to ensure fidelity of the published version to the original work

Liao *et al.*

Reviewer 1: GF Reviewer 2: JS

Reference and design	Technology	Participants	Outcome measures
Author: Liao *et al.*[51] **Year:** 2011 **Study design:** RCT **Number of centres:** not reported but appears to be single centre **Country:** China **Sponsor:** supported in part by grants from Shin Kong Wu Ho-Su Memorial Hospital and Taipei Veterans General Hospital	**Group 1:** BIS, Philips BIS module (Aspect Medical Systems' XP platform technology) with Paediatric BIS Sensor Target device/index value: BIS 40–60 Commencement of monitoring: operating room Involved two anaesthesiologists, one of whom ensured proper functioning of the monitors during surgery **Group 2:** standard clinical practice Involved a single anaesthesiologist Goal: to maintain haemodynamic stability while avoiding patient movement and achieving a rapid recovery Group 3: auto-regressive index (AAI)-guided anaesthesia (data not extracted) Patients in all groups received both BIS and AAI sensors, and headphones, placed before induction in the operating room. In group 1, the AAI monitor was positioned out of the anaesthesiologist's line of sight. In group 2 the AAI and BIS monitors were positioned out of the anaesthesiologist's line of sight Length of experience/ training of anaesthetist: NR; all patients were induced by the same staff anaesthesiologist; patient behaviour during induction was assessed by a trained observer using the Induction Compliance Checklist (reference cited)	**Total numbers involved:** 160; group 1, 52. group 2, 54 (group 3, 54 – data not extracted) Premedication used: stated none GA used: inhaled: Induction: sevoflurane, initially 8 vol% fraction inspired with 50% N_2O in oxygen Maintenance: sevoflurane titrated by BIS values (group 1) or in 0.5% increments according to clinical signs (group 2), or in response to patient movement (either group) Recovery: sevoflurane was stopped at the time of the final surgical suture and fresh gas flow was increased Regional anaesthesia used: none reported Analgesia used: i.v. fentanyl 1 µg/kg 5 minutes before incision Muscle relaxants used: stated none (patients breathed spontaneously) Antinausea drugs used: none reported Other drugs used: in the PACU for patients who cried or suffered pain: meperidine 1.0 mg/kg; if agitation persisted, further meperidine 0.5 mg/kg and then midazolam 0.1 mg/kg (routes of administration not stated) Type of surgery: paediatric outpatient urologic surgery Duration of surgery, minutes, mean ± SD: group 1, 28.4 ± 11.2; group 2, 30.2 ± 14.0 ($p = 0.70$ for 3-group comparison) Duration of GA, minutes, mean ± SD: group 1, 39.5 ± 11.7; group 2, 41.8 ± 14.0 ($p = 0.44$ for three-group comparison) Duration of GA maintenance phase, minutes, mean ± SD: group 1, 36.8 ± 9.7; group 2, 38.7 ± 14.8 ($p = 0.79$ for three-group comparison) **Inclusion criteria:** pre-puberty children, aged 3–12 years, with ASA physical status I or II, scheduled for elective urologic outpatient surgery **Exclusion criteria:** history of premature delivery; reported developmental delay; deafness; significant cardiovascular, respiratory or neurological disease; receiving medication known to affect the central nervous system Baseline measurements (*p*-values refer to three-group comparisons; data for group 3 not extracted):	**Primary outcome:** • Recovery time (time to first spontaneous movement) **Secondary outcomes:** • Emergence delirium • Postoperative nausea and vomiting • Parental satisfaction • Anaesthetic consumption • Anaesthesia duration • Maintenance duration • Intraoperative recall • Device values • Haemodynamic parameters **Length of follow-up:** varied with outcome: up to 30 minutes after awakening for PACU; up to time of discharge for patient satisfaction; unclear for intraoperative recall (nurses appear to have assessed this at a separate follow-up interview, the date of which was not reported) **Methods of assessing outcomes:** Anaesthesia time: defined as the time from induction to discontinuation Sevoflurane maintenance time: defined as the time from insertion of laryngeal mask airway to discontinuation of sevoflurane Surgery time: defined as the time from incision to the final surgical suture End of surgery: defined as the time of the final surgical suture Responses: times of first movement response, phonation or eye-opening were assessed after discontinuation of sevoflurane (i.e. after the final surgical suture)

Reviewer 1: GF Reviewer 2: JS

Reference and design	Technology	Participants	Outcome measures
		Sex (male), *n* (%): group 1, 41 (79); group 2, 45 (83); $p = 0.15$ Age, years, mean ± SD: group 1, 6.0 ± 2.8; group 2, 6.1 ± 2.8; $p = 0.39$ Ethnic groups: probably Chinese (NR) Weight (kg) mean ± SD: group 1, 24.7 ± 11.1; group 2, 23.5 ± 9.3; $p = 0.54$ Height, cm, mean ± SD: group 1, 116.7 ± 17.5; group 2, 115.8 ± 15.4; $p = 0.52$ BMI, kg/m^2, mean ± SD: group 1, 16.4 ± 3.2; group 2, 16.3 ± 2.5; $p = 0.88$ ASA grade I/II, *n*: group 1, 46/6; group 2, 50/4; $p = 0.74$ Risk factors for awareness: none specifically reported Comorbidities: none reported **Losses to follow-up:** none reported **Place of anaesthetic administration:** induction commenced in a pre-anaesthetic clinic; full anaesthetic given in the operating room	PAED score (reference cited): assessed by a trained observer in the PACU every 5 minutes after awakening for 30 minutes. The highest score during this period was used in the final PAED score Readiness for PACU discharge (= full hospital discharge): defined as a score of 9 or more, with no zeros in any domains, on the Aldrete score, and a room air O_2 saturation of ≥96% Intraoperative recall: patients were asked at a follow-up interview (timing not specified) by a nurse of the Anaesthesia Department of the hospital whether they could recall any event or dreaming during the intraoperative period Parent satisfaction with child's treatment: assessed at PACU discharge and rated on a scale from very good, good, acceptable to a bad experience

NR, not reported.

Outcome	Group 1 ($n = 52$)	Group 2 ($n = 54$)	*p*-value (a) for three-group comparison; (b) post hoc comparison group 1 v group 2
Intraoperative awareness with explicit recall, *n* (%)	0 (0)	0 (0)	NA
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time (minutes) to emergence from anaesthesia, mean ± SD:			
Spontaneous movement	3.6 ± 2.7	6.1 ± 5.7	(a) 0.02; (b) < 0.05
Phonation	8.4 ± 5.2	12.9 ± 9.0	(a) 0.11
Eyes opening	15.0 ± 16.4	16.1 ± 11.3	(a) 0.17
Time to extubation: NA			
Time (minutes) to laryngeal mask airway removal, mean ± SD	1.8 ± 1.6	2.1 ± 2.4	(a) 0.93
Time (minutes) to discharge from the recovery room, mean ± SD	64.5 ± 10.1	66.8 ± 9.0	(a) 0.03; (b) < 0.05
Anaesthetic consumption			
Sevoflurane, (g/minute), mean ± SD	0.6 ± 0.2	0.9 ± 0.3	(a) < 0.001; (b) < 0.01
Mean end-tidal sevoflurane concentration,%, during maintenance	2.5 ± 0.4	2.9 ± 0.5	(a) 0.001 (b) < 0.01
(See also additional comments below concerning anaesthetic consumption at different time points)			
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs			
Postoperative nausea, *n* (%)	5 (10)	6 (11)[a]	(a) 0.95
Postoperative vomiting, *n* (%)	2 (4)	3 (6)[a]	(a) 0.88
Pain/pain relieving drugs, *n* (%)			
Did not receive analgesic or sedative agents	4 (8)[a]	5 (9)	(a) 0.83
Rescue requiring more analgesic or sedative agents	9 (17)	6 (11)[a]	(b) 0.6
Fentanyl use (μg) mean ± SD	24.8 ± 11.1	23.4 ± 9.1	(a) 0.54
Other morbidity			
PAED score, median (interquartile range)	18 (14–16)	15 (13–15)	(a) 0.94
Mortality, *n* (%)	0 (0)	0 (0)	NA

NA, not applicable; NR, not reported.

Additional results/comments (e.g. early response factors, QoL)

Baseline data were reported for the number (%) of patients in each group who underwent the following types of surgery: herniorrhaphy; circumcision; herniorrhaphy and circumcision; orchiopexy; hydrocelectomy; varicocele ligation (p-values for three-group comparisons of these variables all >0.7; data not extracted). Baseline data were also reported for the BMI-for-age percentile (three-group comparison, $p = 0.52$) and Induction Compliance Checklist score (three-group comparison, $p = 0.96$) (data not extracted)

Mean arterial pressure did not differ significantly between the groups at baseline ($p \geq 0.05$), but was significantly higher in group 1 than group 2 during and at the end of surgery ($p < 0.01$) (reported in a graph; data not extracted)

Mean heart rate and mean respiratory rate did not differ significantly between the groups at any time point ($p \geq 0.05$) (data not reported)

Mean end-tidal sevoflurane concentration (%) was reported in a graph for six time points from start of induction to end of surgery and was significantly higher ($p < 0.01$) in group 1 than group 2 at four times: at the start of surgery; 5 minutes after incision; 10 minutes after incision; and at the end of surgery (data not extracted)

The number (%) of patients who moved during surgery was 11 (21) in group 1 and 10 (19) in group 2 ($p = 0.94$ for three-group comparison)

The number (%) of patients whose parents gave a satisfaction score of very good, good, acceptable or bad was reported and did not differ significantly between the groups ($p = 1.00$ for each rating class; there were no bad experiences reported) (data not extracted)

Stated there were no adverse respiratory events in any of the groups

Methodological comments

Allocation to treatment groups: patients were allocated randomly to three groups after induction of anaesthesia, using a computer-generated randomisation table

Allocation concealment: NR

Blinding: two anaesthesiologists were involved in the study, a third investigator assessed the patient during the emergence and recovery period, and a nurse of the Anaesthesia Department assessed intraoperative recall at a follow-up interview. Stated that both anaesthesiologists were blinded to the anaesthetic technique and all three investigators were blinded to the grouping of the patient. However, the methods used to achieve blinding were not reported, and it was not stated whether or not the nurse who assessed intraoperative recall was blinded to the patient group

Analysis by ITT: not reported, but there appears to have been no attrition; all randomised patients would appear to have been analysed

Comparability of treatment groups at baseline: groups appear comparable for age, weight, ASA health status, types of surgery being undertaken and haemodynamic parameters; no statistically significant differences were reported at baseline

Method of data analysis: group comparisons of continuous variables were made by one-way analysis of variance for normally distributed variables or by Kruskal–Wallis rank-sum test for non-normally distributed variables. Where differences were significant, post hoc comparisons between groups were by Bonferroni correction (normally distributed variables) or by Mann–Whitney U-test (non-normally distributed variables). Categorical data were analysed by chi-squared or Fisher's exact test as appropriate

Sample size/power analysis: stated that an a priori power analysis was based on a previous study (Bannister *et al.*[45]) which suggested that a sample size of 44 patients for each group should be adequate to achieve a 30% or greater reduction in the time to first movement response with a power of 0.9 ($\alpha = 0.05$)

Attrition/dropout: none reported, but sample sizes for postoperative outcomes suggest there were no dropouts

General comments

Generalisability: pre-pubertal predominantly male, probably Chinese, paediatric outpatient population with ASA health status <3, who received GA with sevoflurane. Not identified as being at high risk of intraoperative awareness

Intercentre variability: NA (appears to be one centre)

Conflict of interests: NR

NA, not applicable; NR, not reported.

a Rounded percentage as calculated by reviewer (difference of 1% from that reported by the authors).

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Low	Randomisation sequence generated by computer
Allocation concealment	Unclear	No information provided
Performance bias		
Blinding of participants and personnel	Unclear	Stated that both anaesthesiologists were blinded to the anaesthetic technique and all three investigators were blinded to the grouping of the patient. However, the methods used to achieve blinding were not reported so it is unclear how easily blinding could be broken
Detection bias		
Blinding of outcome assessment	Low	Not reported whether or not the nurse who assessed intraoperative recall was blinded. The investigator who assessed other outcomes was blinded (method of blinding not reported)
Attrition bias		
Incomplete outcome data	Low	None reported, but sample sizes for postoperative outcomes suggest there were no dropouts
Reporting bias		
Selective reporting	Low	No evidence to suggest reporting bias

Messieha

Reviewer 1: JS Reviewer 2: GF			
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Messieha *et al*.[52] **Year:** 2004 **Study design:** RCT **Number of centres:** one (presumed) **Country:** USA **Sponsor:** not stated	**Group 1:** 'BIS known' – BIS (Aspect Medical Systems), no further detail given Target device/index value: 60-70 Adjustment of inhalation anaesthetic also based on patient vital signs (heart rate, blood pressure, surgical stimulation) **Group 2:** 'BIS unknown' Adjustment of inhalation anaesthetic based on patient vital signs (heart rate, blood pressure, surgical stimulation) BIS was recorded but anaesthesiologist was not aware of the BIS number Commencement of monitoring: not stated when monitoring started, but BIS was continued until PACU discharge Length of experience/ training of anaesthetist: not stated	**Total numbers involved:** 20 children recruited, 10 in each study arm Premedication used: ketamine 3 mg/kg; midazolam 0.05 mg/kg; glycopyrrolate 0.2 mg, intramuscular injection General anaesthetic used: sevoflurane, dose not stated Regional anaesthesia used: none stated Analgesia used: fentanyl, 1 µg/kg (maintenance) Muscle relaxants used: rocuronium bromide 1 mg/kg Antinausea drugs used: ondansetron 0.15 mg/kg, given near the end of the procedure Other drugs used: none stated Type of surgery: complete dental rehabilitation Duration of surgery (minutes), mean (SD): group 1 = 139 (± 43); group 2 = 162 (± 35); $p = 0.2$ Duration of GA: not stated **Inclusion criteria:** scheduled to undergo complete dental rehabilitation under general anaesthetic. Patients with mild cerebral palsy without significant neurological deficit also enrolled **Exclusion criteria:** none stated **Baseline measurements:** Sex (male), *n* (%): group 1 = 4 (40); group 2 = 7 (70) ($p = 0.3$) Age (years), mean (SD): group 1 = 7.4 (± 3), range 3–13 years; group 2 = 5.5 (± 3), range 2-12 years ($p = 0.2$) Ethnic groups, *n* (%): NR Weight (kg), mean (SD): group 1 = 28 (± 15); group 2 = 21 (± 9); $p = 0.2$ ASA physical status grade, mean (range): group 1 = II (I–III); group 2 = II (I–III); $p = 1.0$ Risk factors for awareness: none reported Comorbidities – cerebral palsy, *n* (%): group 1 = 2 (20%); group 2 = 2 (20%); $p = 1.0$ **Losses to follow-up:** NR **Place of anaesthetic administration:** presedation was given prior to transfer to the operating room. Upon transfer GA was started	**Primary outcome:** • Study focused on the reduction in time from end of general anaesthesia to extubation and to PACU discharge **Secondary outcomes:** • Length of PACU stay • Duration of surgery • BIS values **Length of follow-up:** not stated **Methods of assessing outcomes:** not stated other than BIS values were recorded by an independent observer. Not clear whether or not assessment of other outcomes was blinded

NR, not reported.

Outcome	Group 1	Group 2	*p*-value
Intraoperative awareness/recall	NR	NR	NR
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time (minutes) to emergence from anaesthesia	NR	NR	NR
Time (minutes) to extubation, mean (SD)	9 (± 5)	13 (± 5)	0.07
Time (minutes) to PACU discharge, mean (SD)	60 (± 13)	90 (± 11)	<0.001
Duration (minutes) of PACU stay, mean (SD)	45 (± 8)	71 (± 9)	<0.001
Anaesthetic consumption	NR	NR	NR
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs	NR	NR	NR
Pain/pain-relieving drugs	NR	NR	NR
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

NR, not reported; SD, standard deviation.

Additional results/comments

BIS values recorded at key points before, during and after the surgical and anaesthetic procedure showed no statistically significant differences between groups

Duration of surgery did not differ statistically significantly between the two study arms

The level of the surgical care and the procedure were similar in all patients

Methodological comments

Allocation to treatment groups: random, no further information given

Allocation concealment: NR

Blinding: describes the study as observer blind, but no other information provided. Presume that the observer recording BIS values was not aware of allocation to study arm

Analysis by ITT: NR

Comparability of treatment groups at baseline: described as comparable. No statistically significant differences reported between groups at baseline

Method of data analysis: student's *t*-test and Mann–Whitney rank-sum test

Sample size/power analysis: NR

Attrition/dropout: NR

General comments

Generalisability: relevant to US paediatric patients undergoing dental procedures under general anaesthetic with use of premedication and muscle relaxant. Not clear which version of the BIS module was used, so results may not necessarily be comparable to studies using later or earlier versions

Intercentre variability: NA (presumed to be one centre)

Conflict of interests: NR

NA, not applicable; NR, not reported.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Unclear	No information given on the randomisation method used
Allocation concealment	Unclear	NR
Performance bias		
Blinding of participants and personnel	Unclear	NR
Detection bias		
Blinding of outcome assessment	Unclear	BIS values recorded by blinded observer. Not clear whether or not assessment of other outcomes was blinded
Attrition bias		
Incomplete outcome data	Unclear	NR
Reporting bias		
Selective reporting	Low	No evidence to suggest selective reporting

NR, not reported.

Messieha *et al.*

Reviewer 1: JS Reviewer 2: GF

Reference and design	Technology	Participants	Outcome measures
Author: Messieha *et al.*[53] **Year:** 2005 **Study design:** RCT **Number of centres:** one (presumed) **Country:** USA **Sponsor:** not stated	**Group 1:** 'BIS known' – BIS (Aspect Medical Systems), no further detail given Target device/index value: 55-65 Adjustment of inhalation anaesthetic also based on patient vital signs (heart rate, blood pressure, surgical stimulation) **Group 2:** 'BIS unknown' Adjustment of inhalation anaesthetic based on patient vital signs (heart rate, blood pressure, surgical stimulation) BIS was recorded but anaesthesiologist was not aware of the BIS number End-tidal carbon dioxide maintained at the standard operation room level of 30–35 in all patients (both groups) Commencement of monitoring: not stated when monitoring started, but BIS was continued until PACU discharge Length of experience/ training of anaesthetist: not stated	**Total numbers involved:** 29 children recruited; group 1 = 15; group 2 = 14 Premedication used: Versed (midazolam) 0.7 mg/kg orally General anaesthetic used: titrated sevoflurane, dose not stated Regional anaesthesia used: none stated Analgesia used: fentanyl, 1 µg/kg, i.v. administered at the start of the case Muscle relaxants used: rocuronium bromide 1 mg/kg, single dose administered at the beginning of the case. Reversal was administered at the end of the case (drug not stated) Antinausea drugs used: ondansetron 0.15 mg/kg, i.v. Other drugs used: none stated Type of surgery: complete dental rehabilitation Duration of surgery (minutes), mean (SD): group 1 = 133 (± 31); group 2 = 143 (± 33) Duration of GA: not stated **Inclusion criteria:** aged 2–18 years, scheduled to undergo complete dental rehabilitation under general anaesthetic. Patients with mild cerebral palsy without significant neurological deficit also enrolled **Exclusion criteria:** none stated **Baseline measurements:** Sex male–female ratio: group 1 = 4 : 10; group 2 = 2 : 3 (numbers not reported) Age (years), mean (SD): group 1 = 4 (± 2); group 2 = 4 (± 2) Ethnic groups, *n* (%): NR Weight (kg), mean (SD): group 1 = 17 (± 5); group 2 = 18 (± 5) ASA physical status grade: group 1 = I–II; group 2 = I–II Risk factors for awareness: none reported Comorbidities – Children with mild cerebral palsy were eligible, but it is not stated how many were included **Losses to follow-up:** NR **Place of anaesthetic administration:** presedation was given 15–20 minutes prior to transfer to the operating room. Upon transfer GA was started	**Primary outcome:** • Purpose of the study to evaluate time to extubation (from the end of general anaesthetic or turning off the sevoflurane) and time between anaesthesia termination and discharge from PACU **Secondary outcomes:** • Length of PACU stay • Duration of surgery • BIS values **Length of follow-up:** not stated **Methods of assessing outcomes:** criteria for discharge from PACU included consciousness, normal vital signs, no pain, no nausea or vomiting, ability to pass urine BIS values were recorded by an independent observer. Not clear whether or not assessment of other outcomes was blinded

NR, not reported.

Outcome	Group 1	Group 2	*p*-value
Intraoperative awareness/recall	NR	NR	NR
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time to emergence from anaesthesia	NR	NR	NR
Time (minutes) to extubation, mean (SD)	5 (± 2)	10 (± 7)	0.04
Duration (minutes) of PACU stay, mean (SD)	47 (± 17)	63 (± 17)	0.02
Anaesthetic consumption	NR	NR	NR
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs	NR	NR	NR
Pain/pain-relieving drugs	NR	NR	NR
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

NR, not reported; SD, standard deviation.

Additional results/comments

States that none of the patients experienced postoperative pain or postoperative nausea and vomiting

BIS values recorded at key points before, during and after the surgical and anaesthetic procedure in both arms showed no statistical significance

Duration of surgery did not differ statistically significantly between the two study arms

Stated that the level of the surgical care and the procedure were similar in all patients

Methodological comments

Allocation to treatment groups: random, no further information given

Allocation concealment: NR

Blinding: describes the study as observer blind, but no other information provided. The observer recorded BIS values. Unclear whether or not the measurement of other outcomes was blinded

Analysis by ITT: not reported and not discernible (attrition not reported)

Comparability of treatment groups at baseline: described by authors as comparable in terms of ASA physical status, weight and sex

Method of data analysis: t-test and Mann–Whitney rank-sum test

Sample size/power analysis: NR

Attrition/dropout: NR

General comments

Generalisability: relevant to US paediatric patients undergoing dental procedures under general anaesthetic with sevoflurane with use of oral premedication. Ethnicity not stated; no specific risk factors for intraoperative awareness

Intercentre variability: NA (presumed to be one centre)

Conflict of interests: NR

NA, not applicable; NR, not reported.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Unclear	No information given on the randomisation method used
Allocation concealment	Unclear	NR
Performance bias		
Blinding of participants and personnel	Unclear	NR
Detection bias		
Blinding of outcome assessment	Unclear	BIS values recorded by blinded observer. Not clear whether or not assessment of other outcomes was blinded
Attrition bias		
Incomplete outcome data	Unclear	NR
Reporting bias		
Selective reporting	Low	No evidence to suggest selective reporting

NR, not reported.

Rundshagen *et al.*

Reviewer 1: GF	Reviewer 2: JS		
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Rundshagen *et al.*[60] **Year:** 2007 **Study design:** RCT **Number of centres:** not stated (appears to be single) **Country:** not stated, appears to be Germany (multinational authors) **Sponsor:** study supported by Astra Zeneca and a university institutional research grant	**Group 1:** Narcotrend (NCT) (Narcotrend Monitor version 2.0 AF; MonitorTechnik, Bad Bramstedt, Germany; with Blue Sensor; Medicotest S/A, Istykke, Denmark) Target device/index value: NCT D2 – E0 If outside target NCT level, protocol was to first adapt the stepwise target-controlled propofol infusion ± 0.5 µg/kg/minute then the remifentanil infusion ± 0.1 µg/kg/minute Commencement of monitoring: 5–10 minutes before induction of anaesthesia **Group 2:** standard clinical practice (anaesthesia guided by clinical parameters according to the individual decision of the anaesthetist) Both groups: implied (not stated explicitly) that BIS (A-2000TM, version 2.21; Aspect Medical Systems) and NCT were both monitored, with the anaesthesiologist being blinded to BIS values in group 1 and blinded to both BIS and NCT values in group 2 Length of experience/ training of anaesthetist: stated that all patients were treated by one experienced consultant anaesthetist; no details provided	**Total numbers involved:** 48; group 1, 24; group 2, 20 (after attrition) Premedication used: midazolam 0.1 mg/kg orally, 45 minutes pre surgery General anaesthetic used (i.v.): Induction: remifentanil 0.5 µg/kg/minute continuous infusion followed 1 minute later by target-controlled infusion of propofol, with an estimated plasma concentration 3 µg/ml Maintenance: remifentanil and propofol (doses not stated). FIO_2 was kept at 0.3 (except for one-lung ventilation: 1.0 then 0.5 if blood gas analysis acceptable) Regional anaesthesia used: none reported Analgesia used: novaminsulfone 2 g for 20 minutes before and piritramide 7.5 mg for 5 minutes before the suggested end of surgery. Piritramide or morphine (doses not stated) as needed for early postoperative pain in PACU Muscle relaxants used: rocuronium 0.6 mg/kg, before intubation Antinausea drugs used: metoclompramid (dose not stated) used as rescue medication for nausea Other drugs used: see additional comments for full list Type of surgery: stated only that patients were undergoing all kinds of elective surgery, which included surgery for 'malignoma' and peripheral vascular surgery Duration of surgery: NR Duration of GA (minutes) mean ± SD: group 1, 111.1 ± 59.36; group 2, 104.75 ± 54.01; $p = 0.712$ **Inclusion criteria:** none reported **Exclusion criteria:** neurological diseases; consumption of medication affecting the central nervous system; cardiac surgery; neurosurgery; history of drug dependence; alcoholism; pregnancy; or a known intolerance of the used drugs **Baseline measurements:** Sex, male, *n* (%): group 1, 8 (33); group 2, 8 (40); $p = 0.651$ Age, years, mean: group 1, 48.8 (maximum 70); group 2, 58 (maximum 78); $p = 0.041$ Ethnic groups, *n* (%): NR Weight (kg) mean ± SD: group 1, 80.2 ± 17.19; group 2, 77.7 ± 23.03; $p = 0.680$ ASA grade I/II/III (*n*): group 1, 6/12/4; group 2, 4/13/3; $p = 0.836$	**Primary (powered) outcome:** • Time to extubation **Secondary outcomes:** • Postoperative nausea and fatigue • Total anaesthetic doses • Duration of anaesthesia • Memory during anaesthesia • Clinical parameters (heart rate, pulse oximetry, rectal temperature, end-expiratory CO_2, systolic and diastolic arterial pressure) • NCT and BIS values **Length of follow-up:** longest follow-up appears to be on the first postoperative day (for memory questioning) **Methods of assessing outcomes:** plasma propofol concentration was analysed by high-performance liquid chromatography (details of method, calibration and validation reported) Postoperative nausea and fatigue was assessed after 10, 30 and 90 minutes in the PACU using a 100-mm VAS (no details of scaling given) Memory during anaesthesia was assessed by questioning the patient on the first postoperative day (no details of method given) Heart rate, pulse oximetry, rectal temperature, and end-expiratory CO_2 were measured continuously (Ohmeda Modulus CD; Madison, WI, USA) NCT and BIS values were recorded continuously and stored for off-line analyses

Reviewer 1: GF Reviewer 2: JS			
Reference and design	**Technology**	**Participants**	**Outcome measures**
		Risk factors for awareness: none reported Comorbidities: none reported that would be likely to affect EEG (for other comorbidities see additional comments) Losses to follow up: NR. Attrition reported but unclear whether pre or post randomisation **Place of anaesthetic administration:** GA was induced upon arrival in the operating room	

NR, not reported.

APPENDIX 5

Outcome	Group 1	Group 2	*p*-value
Intraoperative awareness/recall			
Explicit memory during anaesthesia, *n* (%)	0 (0)	0 (0)	NR
Recalled dreaming during anaesthesia, *n* (%)	2 (8)	0 (0)	NR
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time to emergence from anaesthesia	NR	NR	NR
Time (minutes) to extubation, mean ± SD	10.6 ± 7.19	9.29 ± 6.23	0.525
Time to discharge to/from the recovery room	NR	NR	NR
Anaesthetic consumption			
Propofol dose (μg/kg/minute), mean ± SD	0.093 ± 0.042	0.114 ± 0.035	0.089
Remifentanil dose (μg/kg/minute), mean ± SD	0.31 ± 0.10	0.34 ± 0.11	0.449
Propofol plasma concentration, μg/ml, mean ± SD[a]			
Intubation	3.7 ± 1.6	2.9 ± 1.4	>0.05
Skin incision	3.4 ± 1.5	3.1 ± 1.2	>0.05
Extubation	1.5 ± 1.3	1.5 ± 1.4	>0.05
10 minutes after extubation	1.5 ± 1.6	1.0 ± 0.9	>0.05
90 minutes after extubation	0.9 ± 1.3	0.7 ± 1.0	>0.05
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs			
Nausea and fatigue VAS scores, mean ± SD[b]			
Nausea, 10 minutes post surgery	6.88 ± 15.2	24.06 ± 34.04	0.005
Nausea, 30 minutes post surgery	15.44 ± 23.8	18.58 ± 24.9	0.146
Nausea, 90 minutes post surgery	9.18 ± 19.0	12.00 ± 27.4	0.095
Fatigue, 10 minutes post surgery	47.74 ± 20.7	45.31 ± 18.9	0.740
Fatigue, 30 minutes post surgery	57.30 ± 22.4	46.32 ± 23.3	0.088
Fatigue, 90 minutes post surgery	74.73 ± 22.5	63.00 ± 30.2	0.164
Metoclopramid for nausea, *n* (%)	1 (4)	3 (15)	NR
Pain/pain-relieving drugs			
Morphine in PACU, *n* (%)	3 (13)	3 (15)	NR
Piritramide in PACU, *n* (%)	10 (42)	8 (40)	NR
Morphine dose in PACU (mg), mean ± SD[a]	5 ± 0	8 ± 3	NR
Piritramide dose in PACU (mg), mean ± SD[a]	6 ± 2	7 ± 3	NR
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

NR, not reported.

Additional results/comments (e.g. early response factors, QoL)

Baseline data for patients' height, type of operation (peripheral/abdominal/thorax), and Apfel score (risk of postoperative nausea and vomiting) were reported; *p*-values for inter-group differences were all >0.05

Four patients in group 1 (17%) and five patients in group 2 (25%) required surgery because of 'malignoma', but none received preoperative radiation or chemotherapy

Changes in the anaesthetic regimen (titration of dose up or down) were reported for propofol and remifentanil (data not extracted); differences between the study groups were not statistically significant ($p>0.05$)

Average temperature during anaesthesia was reported and was identical in both study groups

Stated that all patients except one were extubated earlier in group 1

Other drugs used during anaesthesia:

Theoadrenaline plus cafedrine (Akrinor) (doses reported), *n* (%): group 1, 14 (58%); group 2, 12 (60%)

Atropine 0.5 mg during induction, *n* (%): group 1, 2 (8); group 2, 0 (0)

Dopamine 1–5 mg/kg/minute to maintain mean arterial pressure >80 mmHg (peripheral vascular surgery patients only), *n* (%): group 1, 4 (17); group 2, 2 (10)

Nitroglycerin spray (antihypertensive), *n* (%): group 1, 1 (4); group 2, 0 (0)

Urapidil 20 mg (antihypertensive), *n* (%): group 1, 1 (4); group 2, 0 (0)

Clonidine 75–150 µg during extubation, *n* (%): group 1, 2 (8); group 2, 2 (10)

Variances of diastolic blood pressure and mean arterial pressure were significantly larger in group 2 ($p\leq0.034$ for both parameters combined), but the combined difference was not significant when age-corrected data were analysed

Comorbidities requiring perioperative medication:

Arterial hypertension, *n* (%): group 1, 6 (25); group 2, 4 (20)

Cardiac arrhythmia, *n* (%): group 1, 3 (13); group 2, 2 (10)

Diabetes type II, *n* (%): group 1, 1 (4); group 2, 2 (10)

Asthma, *n* (%): group 1, 3 (13); group 2, 0 (0)

Miscellaneous, *n* (%): group 1, 7 (29); group 2, 3 (15)

None, *n* (%): group 1, 5 (21); group 2, 8 (40)

Methodological comments

Allocation to treatment groups: stated random allocation but no details provided

Allocation concealment: NR

Blinding: NR

Analysis by ITT: unclear. Analysis does not include all the patients who started but it is unclear whether or not attrition happened pre or post randomisation

Comparability of treatment groups at baseline: groups were similar for the reported variables of sex, height, weight, ASA physical status, type of operation and risk of postoperative nausea and vomiting (Apfel score). However, patients were slightly younger in group 1 ($p=0.041$) (data given above) and no information on ethnicity was provided

Method of data analysis: normality of distribution was tested for all variables using a Kolmogorov–Smirnov test. Intergroup comparisons for propofol concentrations and visual analogue scores were tested by repeated-measures analysis of variance or non-parametric statistics. Intergroup comparisons for time of anaesthesia, doses of anaesthetics and times to extubation were tested by Mann–Whitney *U*-test. Effects of patients' characteristics were tested by analysis of variance and a posteriori Scheffé test. EEG parameters were adjusted for patient characteristics

Sample size/power analysis: To achieve a power of at least 80%, standard deviations of the mean difference in time to extubation reported by Kreuer *et al.*[63] were utilised for comparisons between BIS, NCT and standard clinical practice. Given $\alpha=5\%$, and $d=1.0$, the required sample size was estimated using a power table to be 13 subjects per group

Attrition/dropout: stated that out of 48 patients, the data for 44 patients were included in the final analyses. Reasons for four withdrawals were reported, but it was not stated the withdrawals occurred pre or post randomisation nor how they were distributed among the two study groups

Generalisability: appears to be a German adult population, predominantly of ASA grade II, but some grade I and III, with cardiovascular comorbidities, undergoing various elective surgical procedures, and receiving propofol and remifentanil GA. Ethnicity not reported. No explicit risk factors for intraoperative awareness identifiable

Intercentre variability: NA (appears to be a single-centre study)

Conflict of interests: none reported

NA, not applicable; NR, not reported.

a Assumed by reviewers to be mean and SD values (not explicitly stated).

b Direction of scale not reported: assumed higher values indicate worse nausea and fatigue.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Unclear	No information provided
Allocation concealment	Unclear	No information provided
Performance bias		
Blinding of participants and personnel	Unclear	No information provided
Detection bias		
Blinding of outcome assessment	Unclear	No information provided
Attrition bias		
Incomplete outcome data	Unclear	Attrition reasons reported but distribution of attrition across study groups not reported. Unclear whether attrition was pre or post randomisation
Reporting bias		
Selective reporting	Low	No evidence to suggest selective reporting

Talawar *et al.*

Reviewer 1: GF Reviewer 2: JB

Reference and design	Technology	Participants	Outcome measures
Author: Talawar *et al.*[56] **Year:** 2010 **Study design:** RCT **Number of centres:** one **Country:** India **Sponsor:** stated no external funding used	**Group 1:** E-Entropy (S/5 Avance; GE Healthcare, Datex-Ohmeda Division, Helsinki, Finland) Target device/ index value: state entropy between 45 and 65 during the procedure and between 65 and 70 during the last 15 minutes of surgery Commencement of monitoring: In operating room after anaesthesia induction **Group 2:** 'Control' Anaesthesia was titrated to maintain heart rate and mean arterial pressure within 20% of baseline. Simultaneously monitored entropy values were obscured from the anaesthesiologist Length of experience/training of anaesthetist: NR	**Total numbers involved:** 50; group 1, 25; group 2, 25 Premedication used: none reported General anaesthetic used: Induction: i.v. propofol 3–5 mg/kg for patients with an i.v. line in situ; otherwise inhaled sevoflurane in N_2O and O_2 (50 : 50). Patients receiving propofol/ sevoflurane (*n*/*n*) for induction were: group 1, 14/11; group 2, 17/8 (difference: $p = 0.38$) Maintenance: N_2O, O_2 (50 : 50) and isoflurane at inspired concentration 1% (0.8– 0.9 MAC) with 1 l-flow once steady state achieved. Group 2 only: anaesthetic concentration was increased to 1.3 MAC if movement in response to surgical stimulation, lacrimation, or an increase in heart rate or mean arterial pressure by 20% occurred Recovery: inhalational agent was discontinued after skin closure Regional anaesthesia used: caudal block using 0.25% bupivacaine 0.75–1 ml/kg Analgesia used: i.v. fentanyl 1 µg/kg (appears to be after insertion of the laryngeal mask airway) Maintenance: i.v. fentanyl 0.5 µg/kg was administered if the state entropy–response entropy difference increased by more than 10 (group 1), or if signs did not subside or haemodynamic parameters did not settle after increasing the inhaled anaesthesia to 1.3 MAC (group 2) Post surgery: children with a pain score of ≥6 were administered i.v. boluses of fentanyl 0.5 µg/kg every 10 minutes until pain subsided Muscle relaxants used: none used Antinausea drugs used: none reported Other drugs used: none reported Type of surgery: lower abdominal or urological day care surgery Duration of surgery, minutes, median (range): group 1, 29 (16–95); group 2, 30 (15–94); difference $p = 0.47$ Duration of GA, minutes, median (range): group 1, 68 (32–125); group 2, 72 (47–180); difference $p = 0.23$ **Inclusion criteria:** patients undergoing lower abdominal or urological day care surgery between March 2006 and March 2008. No other criteria reported **Exclusion criteria:** parents refused consent; known neurological disorder; history of major head injury; on antiepileptic drugs; any contraindications to laryngeal mask airway insertion	**Primary (powered) outcome:** • Time to awakening **Secondary outcomes:** • Device values • Haemodynamic parameters (ECG, blood pressure, O_2 saturation, end-tidal CO_2 concentration) • End tidal anaesthesia concentration • Recovery score • Time to discharge for PACU • Postoperative pain score **Length of follow-up:** longest duration of follow-up appears to be up to 2 hours in the recovery area for pain assessment **Methods of assessing outcomes:** Blood pressure was assessed non-invasively Time to awakening was the period from discontinuation of anaesthesia Awakening was defined as spontaneous eye-opening, the onset of purposeful limb movements or phonation Recovery was assessed according to modified Steward Recovery score (reference cited); the time to achieve a maximal Steward score was recorded Time to discharge for PACU was the time to transfer from the operating theatre after switching off inhalational anaesthetic agents Pain was assessed in the recovery area by CHEOPS (reference cited) every 30 minutes for the first 2 hours. Note non-independence of postoperative analgesia and postoperative pain scores (see left)

Reviewer 1: GF Reviewer 2: JB			
Reference and design	Technology	Participants	Outcome measures
		Baseline measurements: Sex (male), *n* (%): group 1, 25 (100); group 2, 22 (88); difference $p = 0.52$ Age, years, median (range): group 1, 4 (2–12); group 2, 5 (2–11); difference $p = 0.73$ Ethnic groups, *n* (%): NR Weight, kg, median (range): group 1, 16 (8–28); group 2, 16 (9–40); difference $p = 0.07$ ASA grade: I and II (not reported separately by group) Risk factors for awareness: none reported Comorbidities: none reported **Losses to follow-up:** none reported (all patients included in analysis) **Place of anaesthetic administration:** operating room	

NR, not reported.

Outcome	Group 1	Group 2	*p*-value (mean difference for parameter; 95% CI)
Intraoperative awareness/recall	NR	NR	NR
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time (minutes) to emergence from anaesthesia			
Recovery time (time to awakening), median (range)	7 (3–18)	10 (5–21)	0.017
Recovery time (time to awakening), mean ± SD	8.2 ± 4.49	10.96 ± 3.86	(2.72; 0.34 to 5.1)
Time to reach Steward score of 6, median (range)	6 (1–15)	8 (2–24)	0.464
Time to reach Steward score of 6, mean ± SD	7.08 ± 3.78	8.36 ± 4.8	(1.3; –1.2 to 3.7)
Time to extubation	Not applicable	Not applicable	Not applicable
Time (minutes) to discharge to/from the recovery room			
Time to discharge for PACU, median (range)	15 (5–31)	19 (10–40)	0.045
Time to discharge for PACU, mean ± SD	15.32 ± 6.6	19.32 ± 7.12	(4.0; 0.07 to 7.9)
Anaesthetic (isoflurane) consumption (%) mean[a]			
Immediately before laryngeal mask airway Laryngeal mask airway insertion	0.81	1.24	<0.05
15 seconds after LMA insertion	0.78	1.24	<0.05
15 seconds after caudal analgesia	0.69	0.84	<0.05
15 seconds after skin incision	0.68	0.78	<0.05
5 minutes after skin incision	0.68	0.79	<0.05
Immediately before removal	0.35	0.38	≥0.05
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs	NR	NR	NR
Pain/pain relieving drugs			
Postoperative pain scores, mean (standard error)			
30 minutes after admission to PACU	4.88 (0.319)	4.76 (0.09)	0.71 (0.12; –0.53 to 0.77)
60 minutes	4.48 (0.10)	4.76 (0.08)	0.01 (–0.28; 4.59 to 4.92)[b]
90 minutes	4.56 (0.10)	4.76 (0.08)	0.01 (–0.2; 4.59 to 4.92)[b]
120 minutes	4.88 (0.21)	5.44 (0.33)	0.01 (–0.56; 4.77 to 6.09)[b]
Required additional fentanyl intraoperatively, *n*	5	5	NR
Required additional fentanyl post surgery (CHEOPS >6), *n*	4	4	NR
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

CHEOPS, Children's Hospital of Eastern Ontario Pain Score; NR, not reported.

Additional results/comments (e.g. early response factors, QoL)

Surgical procedures (*n*, group 1/group 2) were: herniotomy (9/3), urethroplasty (6/8), orchidopexy (6/7), urethural fistula closure/cystoscopy (4/6), not reported (0/1)

Mean state entropy and response entropy values were higher in group 1 than group 2 throughout the procedure; however, the difference was statistically significant only at the moment the child awoke (pre awakening) ($p = 0.03$) and at 1 minute post awakening ($p = 0.01$)

Methodological comments

Allocation to treatment groups: allocation to groups was according to computer-generated random numbers in a sealed envelope (not stated whether or not opaque)

Allocation concealment: an anaesthesiologist not involved in the anaesthetic management of the patient opened the envelope and either obscured or kept the entropy values visible on the monitor (not stated how data were obscured)

Blinding: stated only that the anaesthesiologist in group 2 was blinded to state and response entropy values (method of blinding not stated). Times to awakening and recovery were assessed by a resident anaesthesiologist who was blinded to the treatment allocation (i.e. unaware to which study group a patient belonged)

Analysis by ITT: stated that the data were analysed by intention to treat (data from all 50 randomised patients were analysed)

Comparability of treatment groups at baseline: age and weight were not statistically significantly different in the two groups. Group 2 included two girls, otherwise all participants were boys. Ethnicity was not reported. The surgical procedures performed, and the duration of surgery and anaesthesia were comparable between the two groups

Method of data analysis: Baseline data compared between study groups using chi-squared test or Wilcoxon rank-sum test as appropriate. Heart rate, mean arterial pressure, end-tidal isoflurane concentration, state entropy and response entropy were compared between groups over time using a generalised estimating equation as the observations were correlated

Sample size/power analysis: stated that a pilot study on 15 patients in a 'conventional' group gave a recovery time (assumed by reviewers to refer to time to awakening) of 7 ± 4 minutes. Anticipating a 5-minute difference in recovery time between the study groups, with an error of 0.05 and 90% power, a sample size of 15 in each group was calculated

Attrition/dropout: none reported (all patients included in analysis)

General comments

Generalisability: predominantly (88–100%) male; children of mean age 4–5 years (range 2–12 years); of presumably Indian ethnicity (not stated); with ASA health status grade I-II; undergoing lower abdominal or urological day care surgery with induction under i.v. propofol or inhaled sevoflurane, followed by maintenance under inhaled isoflurane. No specific risk factors for intraoperative awareness identified

Intercentre variability: NA (one centre)

Conflict of interests: stated none

NA, not applicable.

a Mean estimated from graph by reviewer (95% CI was reported but has not been extracted by the reviewer as it was not stated to which group(s) or difference the CI applies).

b As reported: CI does not include the stated mean difference (interpretation unclear).

Domain	Author's judgement (state: low/high/unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Low	Computer-generated sequence
Allocation concealment	Unclear	Allocation sequence was in a sealed envelope but not reported whether or not envelope was opaque nor whom was responsible for entering the sequence from computer to envelope
Performance bias		
Blinding of participants and personnel	Unclear	No information on blinding of anaesthetists or patients was provided, except that anaesthetists were blinded to entropy values in group 2, which would not have concealed intervention assignment
Detection bias		
Blinding of outcome assessment	Low	Times to awakening and recovery were assessed by a resident anaesthesiologist who was blinded to the treatment allocation. Method of blinding not reported. Not stated whether or not assessment of other outcomes was blinded
Attrition bias		
Incomplete outcome data	Low	Analysis by ITT with no discernible attrition
Reporting bias		
Selective reporting	Low	No evidence to suggest selective reporting

Vakkuri *et al.*

Reviewer 1: GF Reviewer 2: JB			
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Vakkuri *et al.*[57] **Year:** 2005 **Study design:** RCT **Number of centres:** six **Countries:** Finland (three), Sweden (two), Norway (one) **Sponsor:** technical assistance, financial support, and equipment for data collection and analysis for this study were provided by Datex-Ohmeda, Helsinki, Finland	**Group 1:** E-Entropy and haemodynamic parameters (Entropy module of S/5 Anaesthesia Monitor with S/5 Collect software [GE Healthcare (formerly Datex-Ohmeda), Helsinki, Finland] Target device/index value: State entropy between 45 and 65 until last 15 minutes of anaesthesia then ideally 65 (not exceeding 70) during last 15 minutes. Response–state entropy difference (response entropy–state entropy) < 10. Heart rate and blood pressure to be kept within ± 20% of baseline (preoperative visit) values Commencement of monitoring: in operating room while patient was awake, before induction of anaesthesia **Group 2:** control: haemodynamic parameters only (heart rate and blood pressure to be kept within ± 20% of baseline values; entropy values recorded on a laptop computer but not displayed) Length of experience/ training of anaesthetist: anaesthetists were allowed to accustom themselves to the use of entropy monitoring for 3 weeks. All participants in the current study had substantial previous experience with electroencephalogram-based depth of anaesthesia monitors	**Total numbers involved:** 335 randomised (number randomised per group not reported). Numbers after attrition: group 1, 160; group 2, 160 Premedication used: oral diazepam 0.1–0.5 mg/kg 60 minutes before induction, except at Norwegian study site (where no premedication was used) General anaesthetic used: Induction: alfentanil bolus ≤ 30 µg/kg and propofol bolus 1.0–2.5 mg/kg Maintenance: continuous infusions of alfentanil ≤ 30 µg/kg/hour and propofol ≤ 9 mg/kg/hour. Lungs were normoventilated with a mixture of O_2 (35–50%) and N_2O (50–65%). In group 1, propofol was titrated to maintain the target state entropy; alfentanil or propofol boluses were permitted if state entropy suddenly increased; and alfentanil infusion was adjusted if the response entropy–state entropy difference > 10 or if haemodynamic parameters exceeded ± 20% of baseline values. In group 2, propofol and alfentanil were given to maintain heart rate and blood pressure within ± 20% of baseline values; propofol and alfentanil infusions were also adjusted depending on signs of unnecessarily deep or inadequate anaesthesia Recovery: infusions were closed down and N_2O was discontinued after skin closure Regional anaesthesia used: NR (implied that patients who underwent shoulder operations may have received inter-scalene plexus blocks post operatively) Muscle relaxants used: according to the anaesthetist's choice, when considered appropriate Antinausea drugs used: none reported Type of surgery: different types of gynaecological, abdominal, urological, orthopaedic, breast, thyroid and inguinal hernia operations Duration of surgery: NR Duration of GA (minutes) mean ± SD: group 1, 106 ± 48; group 2, 107 ± 49; difference NS	**Primary (powered) outcome:** • Time to awakening **Secondary outcomes:** • Device values • Anaesthetic consumption • Other drugs consumed (during surgery and in the PACU) • Durations of anaesthesia and surgery • Intraoperative reactions (movements, coughing, grimacing, eye opening) • Haemodynamic parameters (hypotension, hypertension, bradycardia, tachycardia) • Recovery times (to spontaneous breathing and extubation, eye opening, squeezing of the anaesthesiologist's hand on command, and orientation to time and place) • Time of discharge from operating room to PACU • Postoperative pain • Postoperative nausea and vomiting • Intraoperative awareness • Nurse estimation of postoperative variables (time needed in PACU, patient's need for care, patient's general recovery, patient's satisfaction with the anaesthesia, and actual time spent in the PACU) **Length of follow-up:** longest follow up appears to be the first postoperative day (for intraoperative awareness assessment) **Methods of assessing outcomes:** time to awakening: defined as the time to response to a verbal command Time to orientation to time and place: method of assessment not reported Anaesthetic consumption: infusion rates of anaesthetics were noted manually in the anaesthetic record Drug consumption: noted manually in the anaesthetic record

Reviewer 1: GF Reviewer 2: JB			
Reference and design	**Technology**	**Participants**	**Outcome measures**
		Inclusion criteria: either sex; age 18–80 years; ASA physical status I, II or III; ability to read and understand the consent form; elective surgery procedures expected to last 45–150 minutes **Exclusion criteria:** known psychiatric or neurological disorders; history of major head injury; substance abuse; medication affecting the central nervous system; acquired scalp or skull abnormalities; uncontrolled hypertension (baseline systolic pressure >160 mmHg or baseline diastolic pressure >105 mmHg); baseline systolic blood pressure <90 mmHg; baseline heart rate <55 beats/minute; insulin-dependent diabetes; renal or hepatic disease; pregnancy; BMI >33 kg/m^2; any serious medical condition that would interfere with cardiovascular response assessment; cardiac, vascular or cranial neurosurgery; intraoperatively activated epidural analgesia; emergency or other non-elective surgery Baseline measurements (reported only for analysed population after attrition; $N = 320$); all differences stated NS: Sex (male), *n* (%): group 1, 44 (28); group 2, 39 (24) Age, years, mean ± SD: group 1, 45 ± 14; group 2, 47 ± 13 Ethnic groups, *n* (%): NR Weight (kg) mean ± SD: group 1, 71 ± 12; group 2, 71 ± 12 ASA grade I/II/III (*n*): group 1, 113/42/5; group 2, 101/57/2 Risk factors for awareness: stated none Comorbidities: none reported (note extensive exclusion criteria for comorbid patients) **Losses to follow-up:** reported with reasons but not separable by study group **Place of anaesthetic administration:** operating room	Pain scores: measured with a VAS (no details given) Nausea and vomiting: measured with a VAS 'on the day after anaesthesia was studied' (meaning seems ambiguous); no details of the VAS given) Intraoperative awareness: assessed by modified Brice interview (reference cited) first in the PACU and again during the first postoperative day

NR, not reported; NS, not statistically significant.

Outcome	Group 1	Group 2	*p*-value
Intraoperative awareness/recall	0	0	NR
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time (minutes) to emergence from anaesthesia			
Time to spontaneous breathing, median (range)	4.74 (0.00–18.0)	7.07 (–1.00–28.5)	<0.001
Time to eyes open, median (range)	6.08 (0.15–37.5)	10.8 (2.23–43.2)	<0.001
Time to squeezes hand on command, median (range)	8.60 (1.17–47.4)	12.7 (2.43–48.1)	<0.001
Time to orientation to time and place, median (range)	10.3 (1.17–48.7)	15.1 (4.08–113)	<0.001
Time (minutes) to extubation, median (range)	5.80 (3.00–27.3)	9.16 (1.67–32.3)	<0.001
Time (minutes) to discharge to/from the recovery room			
Time to discharge from operating room to PACU, median (range)	10.3 (3.83–42.4)	13.0 (5.00–49.8)	<0.001
Time to discharge from PACU, median (range)	134 (50–1293)	150 (7–1020)	0.21
Anaesthetic consumption[a]			
Propofol (mg/kg/minute), median (range)	0.10 (0.04–0.23)	0.11 (0.03–0.21)	<0.001
Alfentanil (μg/kg/minute), median (range)	0.60 (0.12–2.2)	0.57 (0.16–1.6)	0.54
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs			
Patient-reported VAS score	NR	NR	Stated no difference between groups
Pain/pain relieving drugs			
Patient-reported pain VAS score 1 day after anaesthesia	NR	NR	Both outcomes: stated no difference between groups
Opioid analgesic requirements in the PACU	NR	NR	
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

NR, not reported.

Additional results/comments (e.g. early response factors, QoL)

Stated that the aim in all patients was to provide smooth, haemodynamically stable anaesthesia with the shortest possible emergence time and without intraoperative awareness

The initial eight to nine patients at each study site (total 50 patients) were assigned to a historical control group and their data were used to establish standard clinical practice of the participating anaesthetists before entropy monitoring started. The purpose of the historical control group was to get all of the study sites adjusted to the research protocol rather than to compare practices with and without central nervous system monitoring

Stated there were only minor differences between group 2 and the historical control group, with no differences statistically significant except higher values in the historical control group for: blood pressure at 1 minute after intubation ($p = 0.037$); propofol consumption during the last 15 minutes ($p = 0.001$); and alfentanil consumption during the last 15 minutes ($p = 0.02$)

Both group 1 and group 2 had more women than men because many of the participating centres included mainly gynaecological surgery patients in this study (patient numbers not reported by surgery type)

Stated that the incidence of untoward intraoperative reactions (movement or increased muscle tension, tearing, coughing, frowning, eye-opening, and episodes of hypertension, tachycardia or bradycardia) did not differ between study groups (no quantitative data reported)

Stated haemodynamic data were similar between groups; heart rates and blood pressures did not differ between groups until skin closure, where the entropy group had higher heart rate (mean ± SD: 63 ± 11 vs 60 ± 10 beats/minute; $p = 0.029$) and blood pressure (83 ± 10 vs 79 ± 12 mmHg; $p = 0.008$) (no other haemodynamic data reported)

Stated that recovery in the PACU was similar between groups. The incidence of postoperative nausea and vomiting, the nurse's estimation of time needed in the PACU, the nurse's estimation of the patient's need for care, the nurse's estimation of the patient's general recovery, and the patient's satisfaction with the anaesthesia, and the actual time spent in the PACU were similar between the two study groups (no quantitative data reported)

Cumulative percentages of patients not responding to verbal command, not yet discharged from the PACU, and not oriented to time and place after anaesthesia as a function of time were presented graphically (data not extracted by reviewer). Each of these outcomes was significantly smaller in group 1 than in group 2 ($p < 0.001$)

Stated that similar haemodynamic profiles in group 1 and group 2 are to be expected because haemodynamic responses guided the alfentanil dose in the study protocol in both groups, not only in group 2

Methodological comments

Allocation to treatment groups: random assignment according to computer-generated random numbers

Allocation concealment: each study site was provided with a sufficient number of closed randomisation envelopes (not stated whether or not opaque). With sequential coding, the subjects were treated in blocks of 10 (five patients per group). The envelopes were opened in the operating room immediately before the induction of anaesthesia

Blinding: not reported, other than entropy values recorded for patients in group 2 were not displayed

Analysis by ITT: no; 15 patients excluded after randomisation were omitted from the analysis

Comparability of treatment groups at baseline: ethnicity was not reported but age, sex, weight, and ASA health status did not differ significantly between group 1 and group 2. Height (data not extracted) also did not differ significantly between groups. (Note that baseline data were reported only for patients included in the analysis, not the full randomised population)

Method of data analysis: data normality was tested by Kolmogorov–Smirnov test and visual estimation of histograms. Unpaired *t*-test was used to test differences in haemodynamic variables, age, weight, height and the duration of anaesthesia. Mann–Whitney *U*-test was used to test differences in all other variables. Kaplan–Meier analysis was performed to test differences in cumulative recovery as a function of time after anaesthesia

Sample size/power analysis: sample size estimate was based a priori on time to awakening after propofol anaesthesia in another study (which specifically focused on clonidine premedication effects on awakening time) (reference cited). A minimum of 147 patients in each group was calculated to detect a 20% difference in patients' responses to a verbal command with a power of 0.8 and an α of 0.05

Attrition/dropout: 385 patients were initially recruited, of which 50 were used as historical controls to determine pre-existing anaesthesia practice. Stated that 17/385 patients were excluded, of which two were from the historical control group. The remaining 335 patients were randomised. The final analysis was on 320 patients (160 per group), with 15 patients excluded after randomisation. Reasons for exclusion were reported [most exclusions (14/17) were a result of 'lack of registered data'] but the origin of the excluded patients (historical control group, group 1 or group 2 was not reported)

General comments

Generalisability: adult population (mean age mid-40s), 72–76% female, assumed Scandinavian, with ASA health status predominantly I/II, undergoing varied types of surgery under inhaled GA with alfentanil and propofol. Population noted not to be at particular risk of intraoperative awareness

Intercentre variability: not reported. Stated that there may have been differences in the recovery protocols between study sites but the study protocol did not override the hospital policy for discharge from PACU to ward

Conflict of interests: study supported by the device manufacturer (formerly Datex-Ohmeda, then GE Healthcare, Finland); authors included a research engineer, research scientist and chief scientist of GE Healthcare and two medical advisors to GE Healthcare. One author was an employee of VTT Information Technology, Finland

a Reported that for propofol the significant difference ($p < 0.001$) applied both during the whole operation and especially during the last 15 minutes, but not stated to which of these time periods the numeric data refer.

Domain	Author's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Low	Computer-generated random assignment
Allocation concealment	Unclear	Steps were taken to conceal allocation using envelopes that were opened only in the operating room immediately before anaesthesia. However, it was not stated whether envelopes were opaque or how codes were transferred from computer to envelopes
Performance bias		
Blinding of participants and personnel	Unclear	No information on blinding of anaesthetists or patients was provided, except that anaesthetists were blinded to entropy values in group 2, which would not have concealed intervention assignment
Detection bias		
Blinding of outcome assessment	Unclear	No information provided
Attrition bias		
Incomplete outcome data	Unclear	Attrition numbers and reasons reported but not separately by study group. Analysis was conducted only on the population after attrition (number randomised per group not discernible)
Reporting bias		
Selective reporting	Unclear	For several outcomes only a brief narrative statement that there was no difference between groups was provided, without any quantitative data or indication of variability
Other bias		
Other sources of bias	High	Notable conflict of interests discernible

Wu *et al.*

Reviewer 1: JB Reviewer 2: GF			
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Wu *et al.*[58] **Year:** 2008 **Study design:** RCT **Number of centres:** one **Country:** Taiwan **Sponsor:** supported in part by the National Science Council	**Group 1:** E-Entropy response entropy and state entropy values shown on GE Datex-Ohmeda S/5™ Anaesthesia Monitor Target device/index value: response entropy and state entropy target values 35–45, corresponding to stable 2% EtSevo in the absence of major surgical stimulation. Gradient between response entropy and state entropy within 5–10. Anaesthesia monitored by entropy unless haemodynamic changes of 30% persisted for more than 5 minutes **Group 2:** conventional group using haemodynamic variables and physical signs (sweating, lacrimation, flushing, wrinkling of frontal facial muscles). If mean arterial pressure or heart rate fluctuated more than 30% of baseline value, EtSevo adjusted in steps of 0.2% until fluctuation < 30% Commencement of monitoring: in the operation room (appears to be before induction, although not explicitly stated so) Length of experience/ training of anaesthetist: NR	**Total numbers involved:** 68 patients enrolled and randomised; data for 65; group 1 = 34; group 2 = 31 Premedication used: none reported General anaesthetic used: Sevoflurane as sole inhalational anaesthetic Induction: fentanyl 2 µg/kg, propofol 2 mg/kg and 2 ml of 2% lidocaine Maintenance: after intubation sevoflurane delivered in a mixed flow of 0.3 l/minute air and 0.7 l/minute oxygen throughout operative period In maintenance period end-tidal CO_2 was kept between 35 and 40 mmHg Sevoflurane turned off once surgeon started to close skin layer Regional anaesthesia used: none used Analgesia used: fentanyl as above Muscle relaxants used: 0.30 mg/kg cis-atracurium Antinausea drugs used: NR Other drugs used: hypertension treated with nicardipine 0.25 mg (heart rate < 90/minute) or labetolol 2.5 mg (heart rate > 90/minute). Ephedrine 4 mg to treat hypotension (MAP < 70% of baseline). Atropine 0.5 mg i.v. bolus for bradycardia (heart rate < 45/minute) Type of surgery: total knee replacement Duration of surgery: approximately 1.5 hours Duration of GA (minutes) mean ± SD: group 1 = 133.74 ± 30; group 2 = 144.84 ± 30 **Inclusion criteria:** ASA I or II scheduled to undergo total knee replacement **Exclusion criteria:** history of cerebrovascular disease, treatment with psychoactive medication, existing cardiac dysrhythmia or weight < 70% or > 130% of ideal body weight **Baseline measurements:** Sex (male), *n* (%): group 1 = 28 (82%); group 2 = 25 (81%) Age (years), mean (SD): group 1 = 68.03 (6.1); group 2 = 68.90 (6.5) Ethnic groups, *n* (%): NR Weight (kg). mean (SD): group 1 = 64.8 (10.2); group 2 = 65.5 (12) ASA grade I/II: group 1 = 11/23; group 2 = 8/23 Risk factors for awareness: NR **Losses to follow-up:** reported with reasons, group 1 = 0, group 2 = 3 **Place of anaesthetic administration:** operation room	**Primary outcome:** • Consumption of sevoflurane **Secondary outcomes:** • Tourniquet-induced hyperdynamic responses • Pain status in the PACU • Postoperative nausea and vomiting • Level of awareness • Subjective complaints • Postoperative analgesic needs • Device values • Haemodynamic parameters **Length of follow-up:** 72 hours postoperatively for postoperative nausea and vomiting (follow-up for level of awareness and other outcomes unclear) **Methods of assessing outcomes:** consumption of sevoflurane determined by GE Datex Ohemeda S/5™ Anaesthetic Delivery Unit System Physiological changes at five major events recorded: intubation, tourniquet inflation, skin incision, tourniquet deflation, extubation For each event data collected at following time points: prior to commencement of event; 1 minute into event; 3 and 5 minutes into event Method of assessing level of awareness not reported

EtSevo, end-expiratory concentration of sevoflurane; MAP, mean arterial pressure; NR, not reported.

Outcome	Group 1, Entropy ($n = 34$)	Group 2, Conventional ($n = 31$)	p-value
Intraoperative awareness/recall	All 65 patients had no explicit recollection of procedure		NR
Patient distress and sequelae resulting from perioperative awareness	NR	NR	
Time to emergence from anaesthesia	NR	NR	
Time to extubation	NR	NR	
Time to discharge to/from the recovery room	NR	NR	
Anaesthetic consumption (ml), sevoflurane, mean (SD)	27.79 (7.4)	31.42 (6.9)	$p = 0.023$
HRQoL	NR	NR	
Nausea/vomiting/antisickness drugs			
Postoperative nausea and vomiting	No statistically significant difference between groups		NR
Pain/pain-relieving drugs			
Postoperative pain status and analgesic use	No statistically significant difference between groups		NR
Mortality	NR	NR	NR

NR, not reported.

Additional results/comments

No cardiovascular or cerebrovascular complication in any patient of either group postoperative

Height, hypertension diabetes reported for baseline but did not differ significantly between group 1 and group 2; same for heart rate and MAP

Treatment for hypertension, mean (SD): group 1 = 0.94 (1.15), group 2 = 1.48 (1.41), $p = 0.043$

Treatment for hypertension 45–60 minutes after tourniquet inflation: group 1 = 1, group 2 = 7, $p = 0.012$

Treatment for hypotension and bradycardia, no statistically significant difference between groups

Methodological comments

Allocation to treatment groups: randomised (no details)

Allocation concealment: no details reported

Blinding: study described as single blind but no details

Analysis by ITT: no (not all randomised patients analysed)

Comparability of treatment groups at baseline: stated no statistically significant differences in age, sex, ASA physical status, height, and weight

Method of data analysis: for nominal data, statistical analysis performed using chi-squared test. Age, sex, weight, height, duration of anaesthesia, heart rate, mean arterial pressure, consumption of sevoflurane statistically compared using independent sample *t*-test. RE and SE values were compared using Mann–Whitney *U*-test. Incidence of treatment of intraoperative adverse events (hypertension, hypotension, bradycardia) compared using Wilcoxon's ranked-sum test. A *p-value* <0.05 was considered significant

Sample size/power analysis: NR

Attrition/dropout: three patients from group 2 not included in results because of missing data (reasons not stated)

General comments

Generalisability: opioids only briefly given during induction phase but not sustained during the operative period. This approach might result in a higher incidence of increased blood pressure in both groups compared with other studies. The ranges of RE and SE were set arbitrarily and different results in consumption of sevoflurane, intraoperative haemodynamics and need for antihypertensive drugs could result with other entropy values. Results applicable to Chinese elderly adults, ASA status I/II undergoing total knee replacement surgery with sevoflurane anaesthesia with the stated entropy values. No specific risk factors for intraoperative awareness identified

Intercentre variability: NA, assumed single centre

Conflict of interests: NR

MAP, mean arterial pressure; NA, not applicable; NR, not reported.

Domain	Reviewer's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation.	Unclear	No methods described
Allocation concealment	Unclear	No methods described
Performance bias		
Blinding of participants and personnel	Unclear	Single blind (no details)
Detection bias		
Blinding of outcome assessment	Unclear	No details
Attrition bias		
Incomplete outcome data	Unclear	Three patients from group 2 excluded from analysis, reasons not stated
Reporting bias		
Selective reporting	Low	No evidence of selective reporting (but some results reported narratively only)

Zhang *et al.*

Reviewer 1: GF Reviewer 2: JS			
Reference and design	**Technology**	**Participants**	**Outcome measures**
Author: Zhang *et al.*[40] **Year:** 2011 (enrolment November 2008–November 2010) **Study design:** RCT **Number of centres:** 13 **Country:** China **Sponsor:** NR (device manufacturer provided BIS electrodes)	**Group 1:** BIS-guided A-2000 BIS Monitor (Aspect Medical Systems, USA) Target device/ index value: 40–60 **Group 2:** routine TIVA (no details – possible variation among centres) BIS monitored but screen covered Commencement of monitoring: NR Length of experience/ training of anaesthetist: NR	**Total numbers involved:** number randomised not reported. Stated 5309 provided outcome data but only 5228 were analysed (group 1 = 2919; group 2 = 2309) Premedication used: none used General anaesthetic used: Induction: midazolam and propofol (doses at the discretion of the anaesthetist) Maintenance: propofol (dose at the discretion of the anaesthetist) Regional anaesthesia used: NR Analgesia used: drugs and doses at the discretion of the anaesthetist Muscle relaxants used: drugs and doses at the discretion of the anaesthetist Antinausea drugs used: NR Other drugs used: NR Type of surgery, group 1/group 2, (%): chest and abdominal 42.8/35.3; craniofacial and cervical 27.2/32.8; gynaecological and obstetric 14.1/12.5; neurosurgery 0.9/0.8; urinary 7.5/8.3; spine and limb (orthopaedic) 5.2/7.8; cardiac 0.8/0.9; other 1.3/1.4; overall difference between groups in surgery type: $p < 0.01$ Duration of surgery (≤1 hour/1–2 hours/>2 hours) (%): group 1: 18.7/43.4/37.9; group 2, 16.3/44.2/39.5; $p = 0.083$ Duration of GA: NR **Inclusion criteria:** age ≥18 years; without any apparent mental defect; scheduled for TIVA; and gave informed consent **Exclusion criteria:** patients unable to be interviewed after surgery (decision criteria not stated); unable to communicate in Mandarin Chinese; under awake intubation; or undergoing intraoperative arousal test **Baseline measurements:** Sex (male), *n* (%): group 1, 1237 (42.8);[a] group 2, 971 (42.6); $p = 0.902$ Age, mean ± SD, years: group 1, 46.95 ± 14.86; group 2, 46.06 ± 14.59; $p = 0.054$ Ethnic groups, *n* (%): NR; assumed majority were Chinese Weight, mean ± SD, kg: group 1, 63.80 ± 11.21; group 2, 63.39 ± 14.59; $p = 0.113$ ASA grade (1/2/>3),%:[b] group 1, 52.3/42.5/5.2; group 2, 59.5/37.5/2.9; $p < 0.01$ Risk factors for awareness: none reported; mentioned in discussion that the types of surgery that could influence awareness risk (cardiac, obstetric) did not differ between the study groups. Mentioned in the introduction that TIVA patients are at increased risk of awareness Comorbidities: NR	**Primary outcome:** • Intraoperative awareness **Secondary outcome:** • None reported **Length of follow-up:** 1 day and 4 days post surgery (awareness) **Methods of assessing outcomes:** awareness was assessed by a blinded observer using a structured questionnaire based on the Brice Interview on the first and fourth days post surgery. The research staff classified awareness as no awareness, possible awareness or awareness (criteria specified). An independent committee assessed the interview results and identified confirmed or possible awareness cases (committee membership not reported)

Reviewer 1: GF Reviewer 2: JS			
Reference and design	**Technology**	**Participants**	**Outcome measures**
		Losses to follow-up: of 5309 patients who provided outcome data, 81 (1.5%) were excluded from analysis (reasons reported, but not in all cases separately by study group). Unclear whether or not 5309 was the total number randomised **Place of anaesthetic administration:** NR	

NA, not reported; TIVA, total intravenous anaesthesia.

Outcome	Group 1	Group 2	*p*-value; OR (95% CI)
Intraoperative awareness/recall, *n* (%)			
Confirmed awareness	4/2919 (0.14)	15/2309 (0.65)	0.002; OR 0.21 (0.07 to 0.63)
Possible awareness	4/2919 (0.14)	6/2309 (0.26)	0.485
Confirmed or possible awareness	8/2919 (0.27)	21/2309 (0.9)	<0.01
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time to emergence from anaesthesia	NR	NR	NR
Time to extubation	NR	NR	NR
Time to discharge to/from the recovery room	NR	NR	NR
Anaesthetic consumption	NR	NR	NR
HRQoL	NR	NR	NR
Nausea/vomiting/antisickness drugs	NR	NR	NR
Pain/pain-relieving drugs	NR	NR	NR
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

NR, not reported.

Additional results/comments (e.g. early response factors, QoL)

Anaesthesia history differed significantly between study groups at baseline ($p = 0.017$). The proportion with anaesthesia history was 18.1% in group 1 and 15.5% in group 2

BIS values were obtained for only six of the total 19 confirmed awareness cases (attributed to poor data collecting and recording). Of these, five cases showed light anaesthesia (BIS >60), with most (four) of these light anaesthesia cases occurring in group 2. BIS data from one patient with intraoperative awareness in group 1 indicated that BIS exceeded the target value (BIS>60 for 21 minutes, with a maximum BIS value of 75), giving light anaesthesia

Anaesthetic consumption was not specified as an outcome but the authors mention that intraoperative records showed that in some patients with awareness insufficient anaesthetic had been applied

Methodological comments

Allocation to treatment groups: carried out at each individual centre through computer-generated random numbers. Details not specified

Allocation concealment: NR

Blinding: anaesthetist was blinded to BIS values in group 2 (monitor screen was covered); stated that interviewers and patients were blinded to the group allocation (details not specified)

Analysis by ITT: not an ITT analysis: number randomised unclear and analyses excluded attrition

Comparability of treatment groups at baseline: the groups differed statistically significantly in terms of patients' ASA status (a higher proportion with worse grades in group 1); anaesthesia history (a higher proportion in group 1 had previous anaesthesia); and the type of surgery received (details above). These variables were tested in univariate analyses (details not specified) to exclude a confounding effect on intraoperative awareness ($p > 0.05$). The groups were otherwise well balanced for age, weight, sex, type of airway (tracheal intubation or laryngeal mask), proportion with a difficult airway and proportion with stable/unstable circulation status

Method of data analysis: independent-samples *t*-tests for intergroup comparisons and also chi-squared tests (no other details given)

Sample size/power analysis: stated (without citing a source) that the required sample size in each group was from 2000 to 2800 to achieve 90% power at 5% two-sided type I error. To allow for missing data, 5000–6000 patients were recruited

Attrition/dropout: number randomised not reported. Stated that outcome data were collected from 5309 patients but only 5228 (i.e. 81 fewer) were analysed. Reasons for attrition were lack of information on group allocation ($n = 54$; not reported separately by group; stated that this attrition was without awareness cases); age <18 years ($n = 11$ in group 1; $n = 10$ in group 2); failure to participate in either of the postoperative interviews ($n = 2$ in group 1; $n = 2$ in group 2); postoperative death ($n = 1$; group not specified); and surgery cancelled after anaesthesia induction ($n = 1$; group not specified)

General comments

Generalisability: Chinese adult population receiving TIVA for a wide range of surgical procedures in 13 centres; no specific risk factors for intraoperative awareness identified

Intercentre variability: NR

Conflict of interests: device manufacturer (Aspect Medical Systems) provided BIS electrodes

NR, not reported.

a Reported percentage differs slightly from actual value (<1%).

b The reported percentages imply that the data are based on fewer patients than were allocated to the study groups (approximately 2650–2654 patients in group 1 and approximately 2224–2241 patients in group 2) (back-calculated numbers are approximate because of rounding errors).

Domain	Reviewer's judgement (state: low/high/ unclear risk)	Support for judgement
Selection bias		
Random sequence generation	Low	Computer-generated random numbers
Allocation concealment	Unclear	No information provided
Performance bias		
Blinding of participants and personnel	Low	Stated that anaesthetists and patients were blinded to group allocation
Detection bias		
Blinding of outcome assessment	Low	Stated that interviewers were blinded to group allocation
Attrition bias		
Incomplete outcome data	Unclear	Attrition not included in analysis; not an ITT analysis; attrition incompletely reported and unclear whether or not balanced across groups
Reporting bias		
Selective reporting	Low	Study focused on one outcome (awareness)

Appendix 6 Data extraction and critical appraisal forms used in the systematic review of cost-effectiveness

Study characteristics

Reference

Abenstein, 2009[97]

Health technology

BIS

Interventions and comparators

What interventions/strategies were included?

GA with BIS

Was a no treatment/supportive care strategy included?

GA without BIS

Research question

What are the stated objectives of the evaluation?

Are the changes in patient outcomes clinically relevant and if so are they cost-effective?

Study type: *cost-effectiveness/cost–utility/cost–benefit analysis?*

Cost-effectiveness

Study population

What definition was used for [condition]? What are the characteristics of the baseline cohort for the evaluation?

Not stated

Institutional setting: *where is/are the intervention(s) being evaluated usually provided?*

Not stated

Country/currency

Has a country setting been provided for the evaluation? What currency are costs expressed in and does the publication give the base year to which those costs relate?

USA, $. Base year not stated

Funding source

Not stated

Analytical perspective

What is the perspective adopted for the evaluation (health service, health and Personal Social Services, third-party payer, societal (i.e. including costs borne by individuals and lost productivity)?

Not stated

Effectiveness

Were the effectiveness data derived from: a single study, a review/synthesis of previous studies or expert opinion? Give the definition of treatment effect used in the evaluation. Give the size of the treatment effect used in the evaluation

Effectiveness data derived from several studies All patients: Incidence of awareness episodes (Ekman): 18/10,000 procedures (GA); 4/10,000 procedures (GA with BIS) High-risk patients: Incidence of awareness episodes (Myles/Avidan): 59/10,000 procedures (GA); 18/10,000 procedures (GA with BIS)

Intervention costs

Were the cost data derived from: a single (observational) study, a review/synthesis of previous studies expert opinion? Were the methods for deriving these data adequately described (give sources if using data from other published studies)? List the direct intervention costs and other direct costs used in the evaluation – include resource estimates (and sources for these estimates, if appropriate) as well as sources for unit costs used.

Sources of intervention costs not stated BIS monitor US$9000 Cost of each BIS electrode sensor was US$17

indicate the source for individual cost values (if appropriate)

Indirect costs (costs due to lost productivity, unpaid inputs to patient care)

Were indirect costs included?

Not applicable

Health state valuations/utilities (if study uses quality-of-life adjustments to outcomes)

Were the utility data derived from: a single (observational) study, a review/synthesis of previous studies expert opinion? Were the methods for deriving these data adequately described (give sources if using data from other published studies)?

Not applicable

List the utility values used in the evaluation

Not applicable

Indicate the source for individual cost values (if appropriate)

Modelling

If a model was used, describe the type of model used (e.g. Markov state transition model, discrete event simulation). Was this a newly developed model or was it adapted from a previously reported model? If an adaptation, give the source of the original. What was the purpose of the model (i.e. why was a model required in this evaluation)? What are the main components of the model (e.g. health states within a Markov model)? Are sources for assumptions over model structure (e.g. allowable transitions) reported – list them if reported.

Simple calculation

Extract transition probabilities for [natural history/disease progression] model and show sources (or refer to table in text).

Not applicable

What is the model time horizon?

Not applicable

What, if any, discount rates have been applied in the model? Same rate for costs and outcomes?

Not applicable

Results/analysis

What measure(s) of benefit were reported in the evaluation?

Cost per awareness episode avoided

Provide a summary of the clinical outcome/benefits estimated for each intervention/strategy assessed in the evaluation.

See above section on intervention costs

Provide a summary of the costs estimated for each intervention/strategy assessed in the evaluation.

Cost of monitor estimated by assuming 7 years use, monitor will be used on four patients per day, 300 days per year, i.e. US$1.07 per patient

Thus cost of BIS monitoring is US$18.07 per patient

Synthesis of costs and benefits – are the costs and outcomes reported together (e.g. as cost-effectiveness ratios)? If so, provide a summary of the results.

The associated cost of preventing each episode of awareness is US$11,294 for all patients. The associated cost of preventing each episode of awareness is US$4410 for high-risk patients.

Give results of any statistical analysis of the results of the evaluation.

None

Was any sensitivity analysis performed – if yes, what type(s) (i.e. deterministic (one-way, two-way, etc. or probabilistic).

No

What scenarios were tested in the sensitivity analysis? How do these relate to structural uncertainty (testing assumptions over model structure such as relationships between health states), methodological uncertainty (such as choices of discount rate or inclusion of indirect costs) or parameter uncertainty (assumptions over values of parameters in the model, such as costs, QoL or disease progression rates)?

None

Give a summary of the results of the sensitivity analysis – did they differ substantially from the base-case analysis. If so, what were the suggested causes?

Not applicable

Conclusions/implications

Give a brief summary of the author's conclusions from their analysis

General use of BIS monitoring does not seem warranted and appears not to be cost-effective

What are the implications of the evaluation for practice?

Not stated

SHTAC commentary

This study is a simple calculation and may not contain all relevant parameters. As such the economic evaluation is of poor quality.

Critical appraisal checklist of economic evaluation (Questions in this checklist based on Philips *et al.*[37])

	Item	Abenstein	Comments
1	Is there a clear statement of the decision problem?	Y	Are the clinical advantages of BIS monitoring . . . clinically relevant and cost-effective?
2	Is the comparator routinely used in UK NHS?	Y	
3	Is the patient group in the study similar to those of interest in UK NHS?	Y	
4	Is the health-care system comparable to UK?	Y	
5	Is the setting comparable to the UK?	Y	
6	Is the perspective of the model clearly stated?	N	
7	Is the study type appropriate?	Y	
8	Is the modelling methodology appropriate?	Y	
9	Is the model structure described and does it reflect the disease process?	Y	
10	Are assumptions about model structure listed and justified?	N	
11	Are the data inputs for the model described and justified?	?	Unclear where the costs are from
12	Is the effectiveness of the intervention established based on a systematic review?	N	
13	Are health benefits measured in QALYs?	N	
14	Are health benefits measured using a standardised and validated generic instrument?	N	
15	Are the resource costs described and justified?	?	Unclear where the costs are from
16	Have the costs and outcomes been discounted?	N	
17	Has uncertainty been assessed?	N	
18	Has the model been validated?	N	

Y, Yes; N, No; ?; unclear.

Appendix 7 Studies excluded from the review of economic evaluations

Reference	Reason for exclusion
Medical Advisory Secretariat. Bispectral index monitor: an evidence-based analysis. *Ont Health Technol Assess Ser* 2004;**4**(9)	Not full economic evaluation
Hayes Inc. *Bispectral index monitoring for anaesthesia awareness*. 2005	Unobtainable
Bard JW. The BIS monitor: a review and technology assessment. *AANA J* 2001;**69**:477–83	Review
Lehmann A, Karzau J, Boldt J, Thaler E, Lang J, Isgro F. Bispectral index-guided anaesthesia in patients undergoing aortocoronary bypass grafting. *Anaesth Analg* 2003;**96**:336–43	Wrong comparator
Liu SS. Effects of Bispectral Index monitoring on ambulatory anaesthesia: a meta-analysis of randomised controlled trials and a cost analysis. *Anaesthesiology* 2004;**101**:311–15	Cost analysis
Mayer J, Boldt J, Schellhaass A, Hiller B, Suttner SW. Bispectral index-guided general anaesthesia in combination with thoracic epidural analgesia reduces recovery time in fast-track colon surgery. *Anaesth Analg* 2007;**104**: 1145–9	Not full economic evaluation
Myles PS, Hunt JO, Fletcher H, Watts J, Bain D, Silvers A *et al.* Remifentanil, fentanyl, and cardiac surgery: a double-blinded, randomised, controlled trial of costs and outcomes. *Anaesth Analg* 2002; **95**:805–12	Not full economic evaluation
Danish Centre for Health Technology Assessment. *Monitoring depth of anaesthesia – a health technology assessment*. Copenhagen: Danish Centre for Evaluation and Health Technology Assessment (DACEHTA). 2007	Not English language (Danish)
Penuelas-Acuna J, Oriol-Lopez SA, Castelazo-Arredondo JA, Hernandez-Bernal CE. Usefulness of bispectral index in pharmaceutical cost reduction for anaesthesia. [Utilidad del indice biespectral (BIS) en la reduccion del costo de farmacos para la anestesia.] *Cirugia y Cirujanos* 2003;**71**:300–3	Not English language (Spanish)
White PF, Tang J, Ma H, Wender RH, Sloninsky A, Kariger R. Is the patient state analyser with the PSArray2 a cost-effective alternative to the bispectral index monitor during the perioperative period? *Anaesth Analg* 2004;**99**:1429–35	Not full economic evaluation
Windisch PA, Worsham GM. The effect of the bispectral index on medication utilisation in the operating room and time to discharge from the postanesthesia care unit. *Hosp Pharm* 2002;**37**: 386–90	Not full economic evaluation
Yli-Hankala A, Vakkuri A, Annila P, Korttila K. EEG bispectral index monitoring in sevoflurane or propofol anaesthesia: analysis of direct costs and immediate recovery. *Acta Anaesthesiol Scand* 1999; **43**:545–9	Not full economic evaluation
Satisha M, Sanders GM, Badrinath MR, Ringer JM, Morley AP. Introduction of bispectral index monitoring in a district general hospital operating suite: a prospective audit of clinical and economic effects. *Eur J Anaesthesiol* 2010;**27**:196–201	Not full economic evaluation

Appendix 8 Pooled intravenous anaesthetic consumption for Narcotrend randomised controlled trials

The mean normalised consumption for propofol and for remifentanil reported in two trials (one in patients undergoing minor orthopaedic surgery[63] and one in all kinds of elective surgery[60]) using Narcotrend depth of anaesthesia monitoring were pooled. *Table 121* reports the normalised propofol consumption (mg/kg/hour) and mean difference in each of the included trials. Pooled estimates for the mean difference are reported in *Table 122* (*Figure 10* presents a forest plot for the analysis).

Table 123 reports the normalised remifentanil consumption (μg/kg/hour) and mean difference in each of the included trials. Pooled estimates for the mean difference are reported in *Table 124* (*Figure 11* presents a forest plot for the analysis).

TABLE 121 Propofol consumption in RCTs using Narcotrend depth of anaesthesia monitoring

Trial	Narcotrend			Standard clinical monitoring			Mean difference	Standard error	95% CI	
	Mean	SD	*n*	Mean	SD	*n*			Lower	Upper
Rundshagen *et al.*[60]	5.58	2.52	20	6.84	2.10	24	–1.26	0.7080	–2.65	0.13
Kreuer *et al.*[63]	4.50	1.10	40	6.80	1.20	40	–2.30	0.2574	–2.80	–1.80

TABLE 122 Pooled estimates for reduction in propofol consumption in RCTs using Narcotrend depth of anaesthesia monitoring

Analysis	Pooled estimate	Standard error	95% CI		Q	I^2	τ^2
Fixed effect	–2.18	0.2419	–2.65	–1.70	1.91	47.53	0.26
Random effect	–1.99	0.4761	–2.92	–1.06			

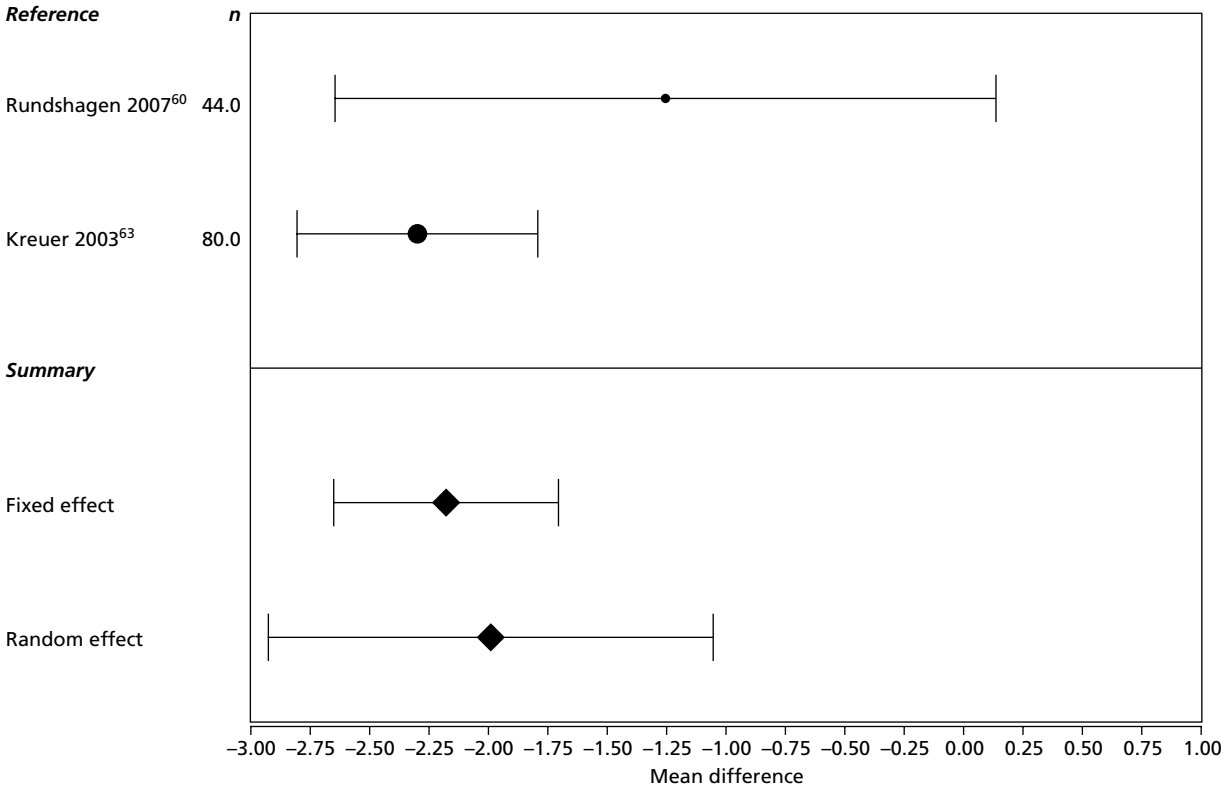


FIGURE 10 Forest plot for the pooled estimate of the mean difference in propofol consumption using Narcotrend depth of anaesthesia monitoring compared with standard clinical monitoring.

TABLE 123 Remifentanil consumption in RCTs using Narcotrend depth of anaesthesia monitoring

Trial	Narcotrend			Standard clinical monitoring			Mean difference	Standard error	95% CI	
	Mean	SD	*n*	Mean	SD	*n*			Lower	Upper
Rundshagen *et al.*[60]	0.31	0.10	20	0.34	0.11	24	–0.03	0.0317	–0.09	0.03
Kreuer *et al.*[63]	0.22	0.06	40	0.23	0.07	40	–0.01	0.0146	–0.04	0.02

TABLE 124 Pooled estimates for reduction in remifentanil consumption in RCTs using Narcotrend depth of anaesthesia monitoring

Analysis	Pooled estimate	Standard error	95% CI		Q	I^2	τ^2
Fixed effect	–0.01	0.0132	–0.04	0.01	0.33	0.00	0.00
Random effect	–0.02	0.3589	–0.72	0.68			

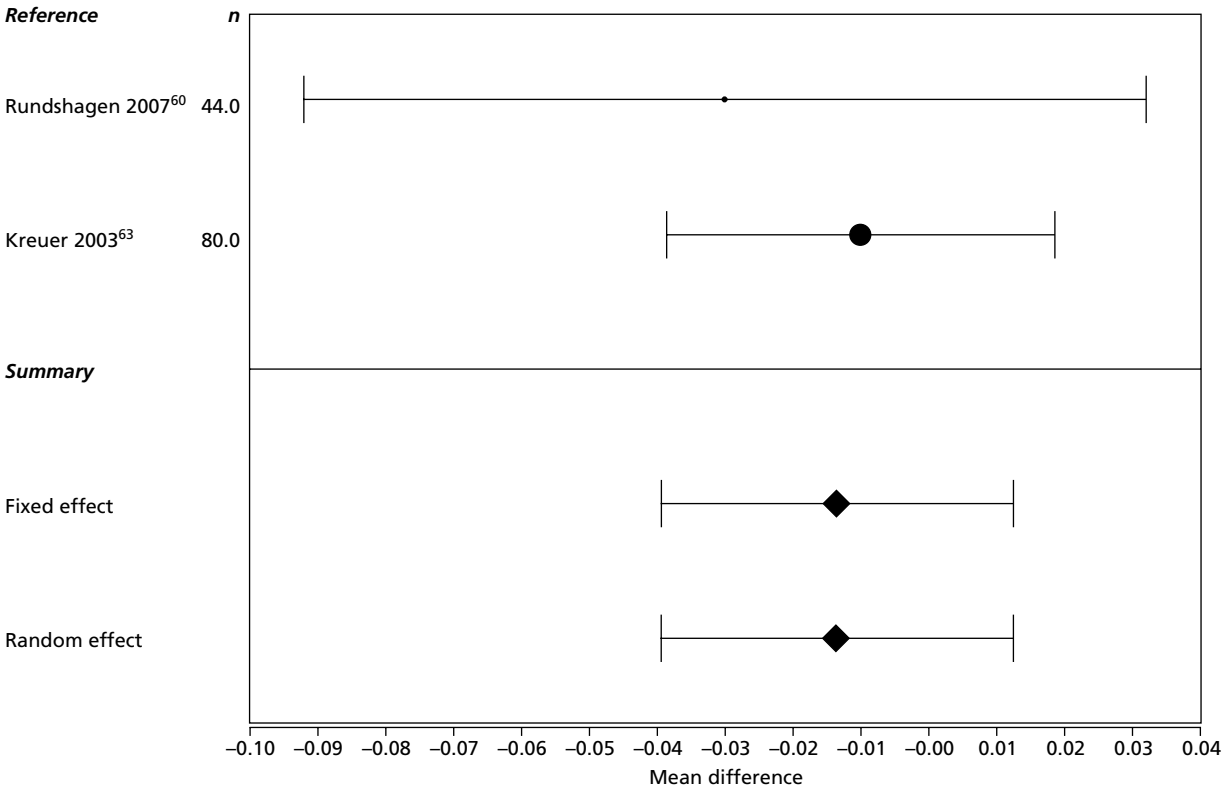


FIGURE 11 Forest plot for the pooled estimate of the mean difference in remifentanil consumption using Narcotrend depth of anaesthesia monitoring compared with standard clinical monitoring.

Appendix 9 Derivation of the pooled estimates of cumulative incidence of awareness used in the model

Table 39 in *Model parameters* of this report presents the cumulative incidence of awareness in studies identified by our targeted searches, for general surgical populations and for patients deemed as being at high risk of awareness. The proportion of patients identified as experiencing awareness in each study were pooled by first transforming the proportions to the Freeman–Tukey variant of the arcsine square root transformed proportion, which is suitable for calculating fixed or random-effect summaries. The pooled proportion is calculated as the back-transform of the weighted mean of the transformed proportions, using inverse arcsine variance weights for the fixed-effect model and DerSimonian–Laird weights for the random-effects model.

Figure 12 shows the forest plot for all identified studies in general surgical populations. A pooled estimate from all these studies gives a cumulative incidence of awareness of 0.21% (95% CI 0.06% to 0.45%) assuming random effects [Cochran's Q = 212.55 (df = 5), $p < 0.0001$, $I^2 = 97.6\%$ for fixed-effect model].

Excluding the two outlying studies (Pollard and colleagues[14] and Errando and colleagues[18]) yields a slightly lower estimate of 0.16% [95% CI 0.10% to 0.23%] assuming random effects [Cochran's Q = 7.85 (df = 3), $p = 0.0493$, $I^2 = 61.8\%$ for fixed-effect model] (*Figure 13*).

Figure 14 shows the forest plot for studies in high-risk surgical populations. A pooled estimate from all these studies gives a cumulative incidence of awareness of 0.45% (95% CI 0.06% to 1.19%) assuming random effects [Cochran's Q = 19.97 (df = 4), $p = 0.0005$, $I^2 = 80.0\%$ for fixed-effect model].

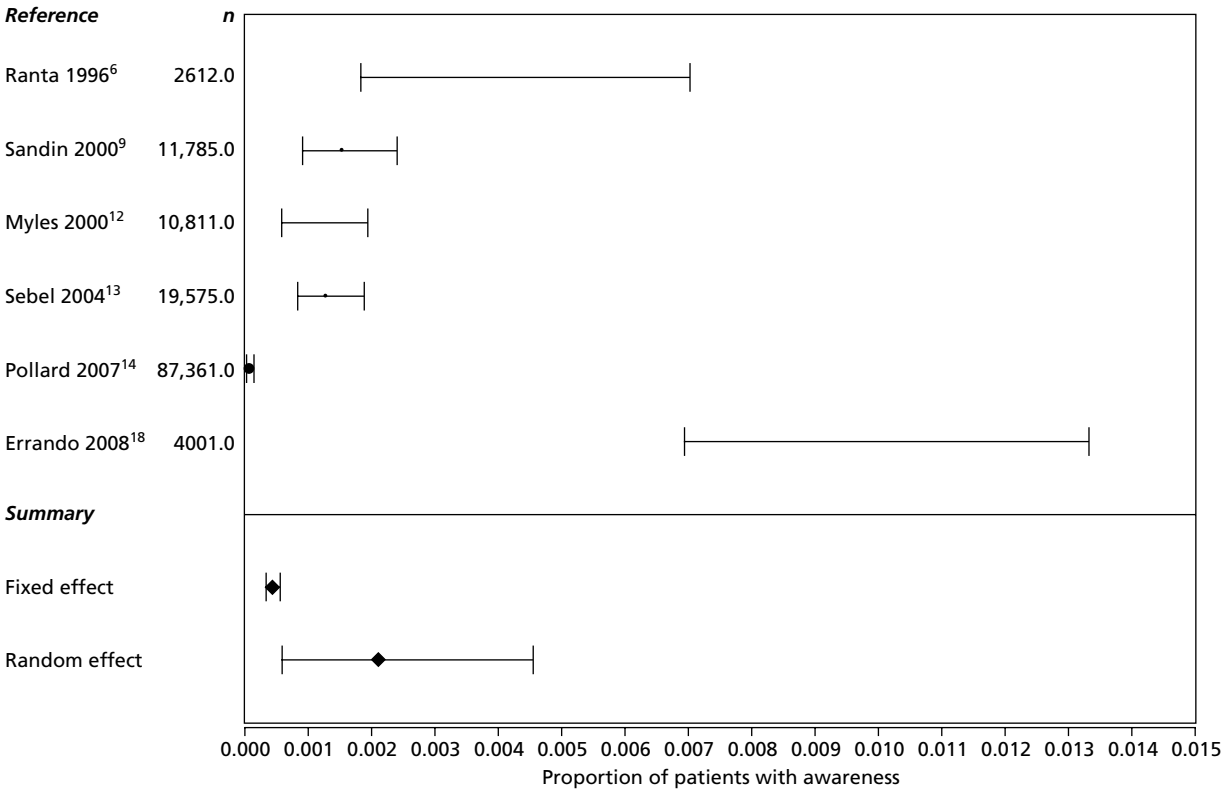


FIGURE 12 Forest plot for pooled estimate of proportion of general surgical patients experiencing awareness.

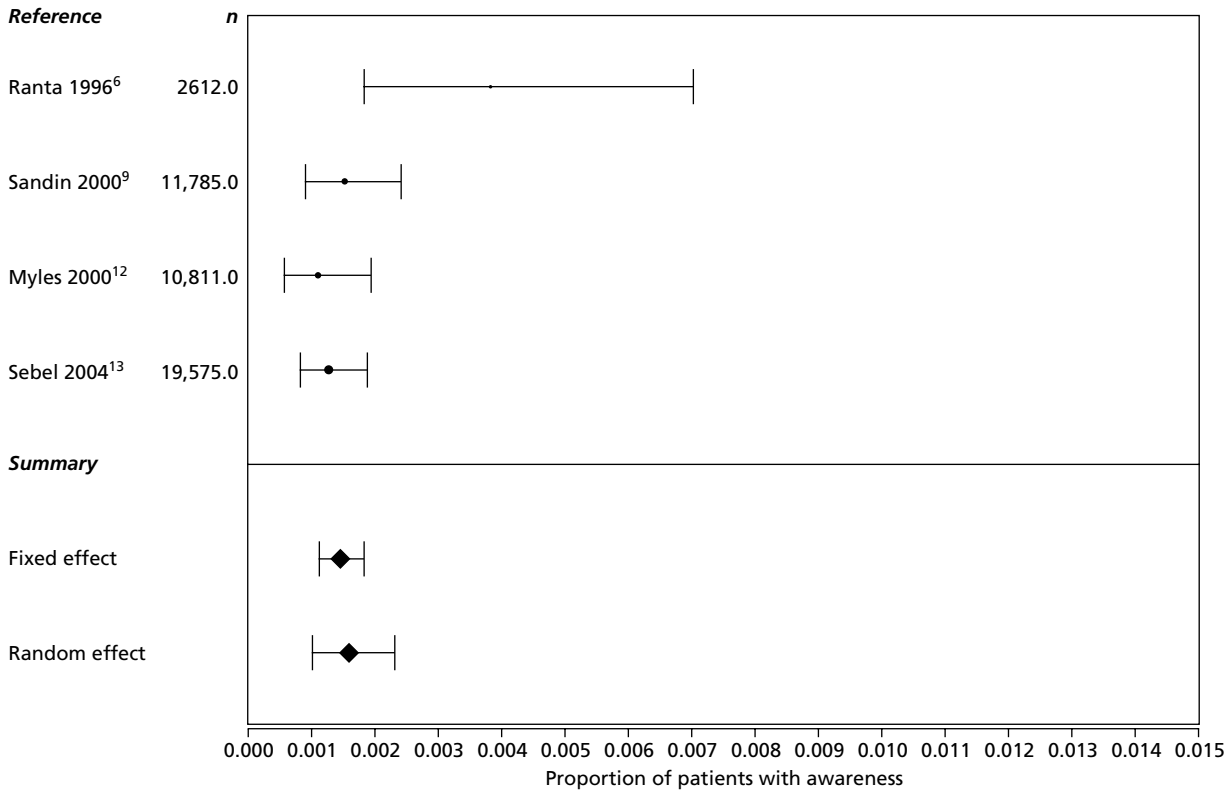


FIGURE 13 Forest plot for pooled estimate of proportion of general surgical patients experiencing awareness (excluding outliers).

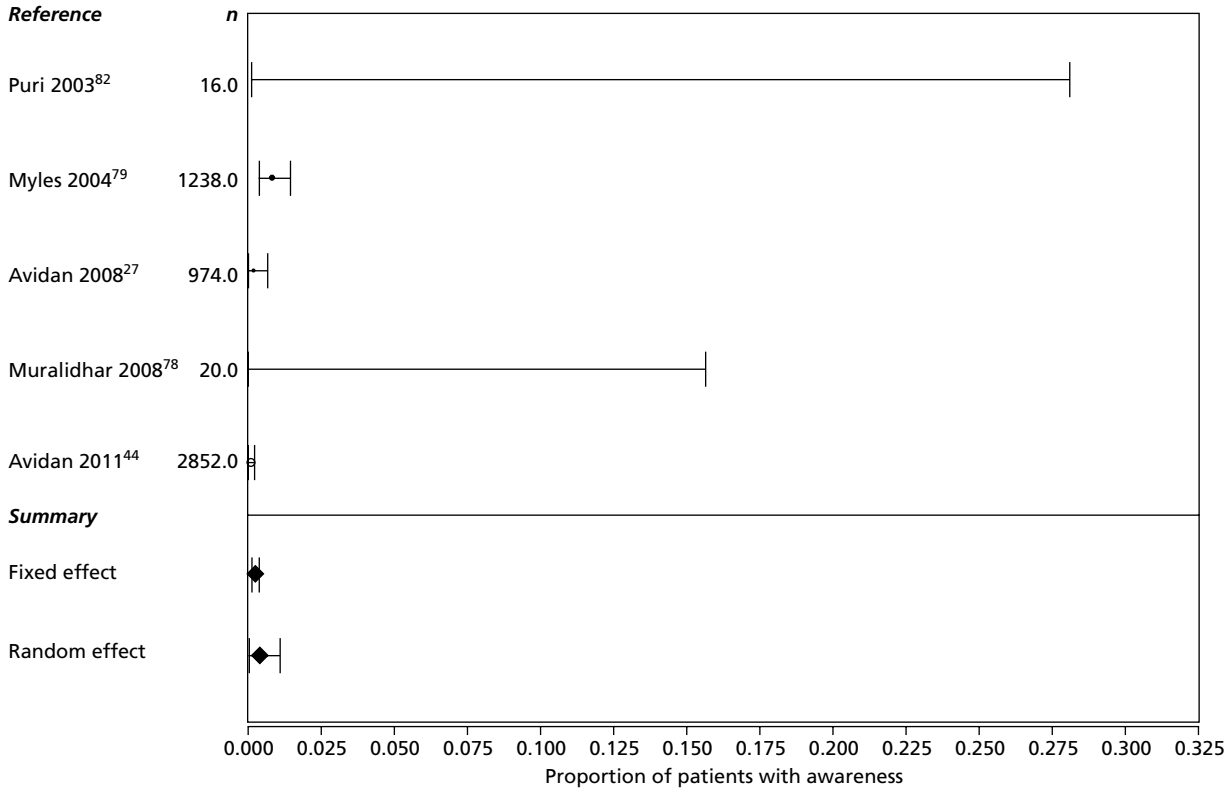


FIGURE 14 Forest plot for pooled estimate of proportion of high-risk surgical patients experiencing awareness.

Appendix 10 Survival modelling methodology

The survival model adopted for this report, to derive the mean duration of PTSD from published survival curves, was developed using linear regression to estimate the parameters of a linear transformation of the observed Kaplan–Meier estimates for duration of PTSD symptoms in identified studies. A parametric survival function (Weibull) was estimated and assessed for goodness of fit to the observed data by visual inspection.

For a Weibull distribution the survival function is given by

$$S(t) = \exp\left(-\lambda t^{\gamma}\right) \tag{3}$$

with scale parameter λ and shape γ. Taking the log of both sides gives

$$\log\left(S(t)\right) = -\lambda t^{\gamma} \tag{4}$$

Taking the log of both sides again, gives

$$\log\left(-\log\left(S(t)\right)\right) = \log(\lambda) + \gamma \log(t) \tag{5}$$

which is a linear function and can be fit using least squares methods to provide estimates of $\log(\lambda)$ and γ.

General method for extracting data from published curves

Figures presenting Kaplan–Meier estimates for duration of PTSD symptoms in identified studies were scanned from the original publications and imported into Enguage software (http://digitizer.sourceforge.net). The process of extracting data from a chart usually begins with the user identifying key reference points on the chart (e.g. indicating the location of the origin and points along the *x*- and *y*-axes). Enguage software will indicate what appear to be data points in the imported image or the user can select individual data points to be extracted using the mouse. Points along the curve were selected at approximately 3-month intervals and the raw data (without any interpolation) were extracted to a text file and imported in Excel (Microsoft Corporation, Redmond, WA, USA).

The following table reports the parameter estimates for linear regression for the Weibull survival function.

	log(λ)	γ
Weibull	–2.82786	0.61006

The mean duration of symptoms can be estimated using the following equation:[150]

$$(1/\lambda)^{(1/\gamma)} \times \Gamma[1 + (1/\gamma)] \tag{6}$$

where Γ is the mathematical gamma function. Therefore, mean duration of PTSD symptoms is estimated as $(1/\exp(-2.82786)^{(1/0.61006)} \times \Gamma[1 + (1/0.61006)] = 151.80$ months, or 12.7 years.

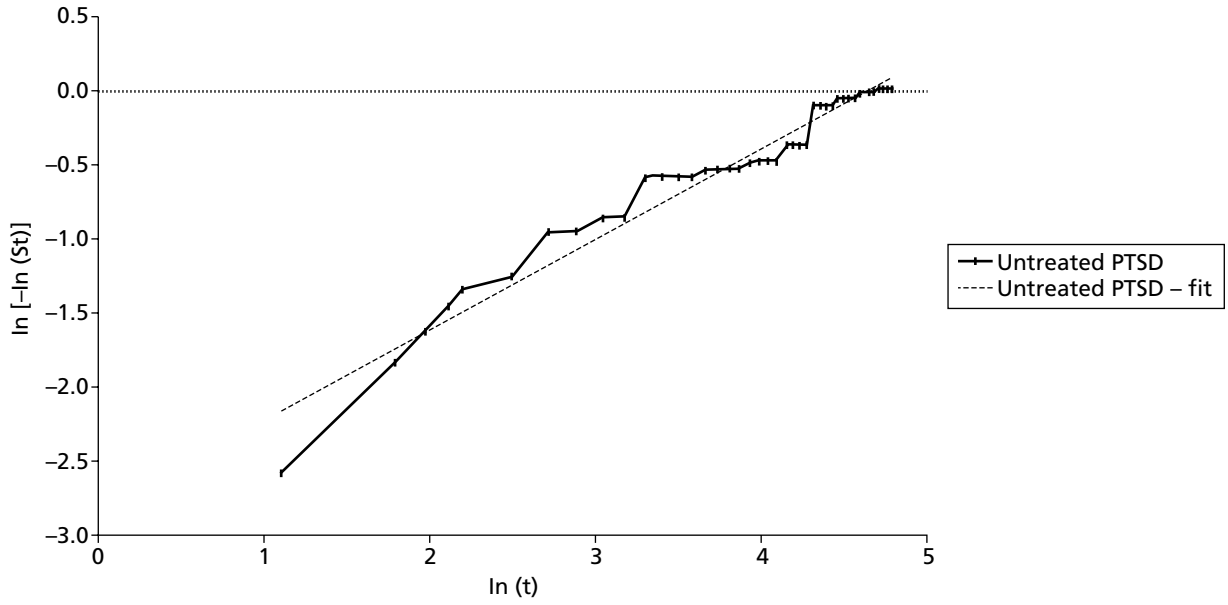


FIGURE 15 Transformed survival curve for duration of PTSD symptoms and linear fit.

Appendix 11 Search strategy to identify utility values for post-traumatic stress disorder

Specific post-traumatic stress disorder and quality-of-life search

Database: Ovid MEDLINE(R) <1948 to November Week 3 2011> 6 December 2011

Also run on MEIP, Science Direct searched for HRQoL terms linked to PTSD terms.

Search strategy

1. value of life/ (5202)
2. quality adjusted life year/ (5364)
3. quality adjusted life.ti,ab. (4269)
4. (qaly$ or qald$ or qale$ or qtime$).ti,ab. (3568)
5. disability adjusted life.ti,ab. (789)
6. daly$.ti,ab. (817)
7. health status indicators/ (17,509)
8. (sf36 or sf 36 or short form 36 or shortform 36 or sf thirtysix or sf thirty six or shortform thirstysix or shortform thirty six or short form thirty six or short form thirtysix or short form thirty six).ti,ab. (11861)
9. (sf6 or sf 6 or short form 6 or shortform 6 or sf six or sfsix or shortform six or short form six).ti,ab. (881)
10. (sf12 or sf 12 or short form 12 or shortform 12 or sf twelve of sftwelve or shortform twelve or short form twelve).ti,ab. (1805)
11. (sf16 or sf 16 or short form 16 or shortform 16 or sf sixteen or sfsixteen or shortform sixteen or short form sixteen).ti,ab. (19)
12. (sf20 or sf 20 or short form 20 or shortform 20 or sf twenty of sftwenty or shortform twenty of short form twenty).ti,ab. (299)
13. (euroqol or euro qol or eq5d or eq 5d).ti,ab. (2429)
14. (hql or hqol or h qol or hrqol or hr qol).ti,ab. (5279)
15. (hye or hyes).ti,ab. (50)
16. health$ year$ equivalent$.ti,ab. (36)
17. health utilit$.ab. (731)
18. hui or hui1 or hui2 or hui3).ti,ab. (677)
19. disutil$.ti,ab. (156)
20. rosser.ti,ab. (69)
21. quality of well being.ti,ab. (285)
22. quality of wellbeing.ti,ab. (6)
23. qwb.ti,ab. (144)
24. willingness to pay.ti,ab. (1562)
25. standard gamble$.ti,ab. (577)
26. time trade off.ti,ab. (568)
27. time tradeoff.ti,ab. (186)
28. tto.ti,ab. (433)
29. (index adj2 well being).mp. (404)
30. (quality adj2 well being).mp. (712)
31. (health adj3 utilit$ ind$).mp. (516)
32. ((multiattribute$ or multi attribute$) adj3 (health ind$ or theor$ or health state$ or utilit$ or analys$)).mp. (201)
33. quality adjusted life year$.mp. (7057)

34. (15D or 15 dimension$).mp. (1002)
35. (12D or 12 dimension$).mp. (304)
36. rating scale$.mp. (73,060)
37. linear scal$.mp. (463)
38. linear analog$.mp. (776)
39. visual analog$.mp. (23,714)
40. (categor$ adj2 scal$).mp. (1028)
41. or/1-40 (145653)
42. (letter or editorial or comment).pt. (1,103,139)
43. 41 not 42 (141,638)
44. Stress Disorders, Post-Traumatic/ (17,609)
45. "posttraumatic stress".tw. (8436)
46. "post traumatic stress".tw. (4553)
47. PTSD.tw. (8895)
48. or/44-47 (20,455)
49. 43 and 48 (2564)
50. or/2-35 (47,102)
51. 48 and 50 (253)
52. HRQOL.tw. (4851)
53. "health related quality of life".tw. (15,264)
54. (health adj2 utility).tw. (573)
55. (health adj2 utilities).tw. (661)
56. ("quality of life" adj5 (predict* or estimat*)).tw. (2279)
57. (model* adj5 "quality of life").tw. (683)
58. ("quality of life" and utility).tw. (3272)
59. qualy*2.tw. (18)
60. ("sf 36" or "SF36" or "short form 36").tw. (11,857)
61. standard gamble*.tw. (577)
62. or/13-30 (11,699)
63. or/52-61 (28,623)
64. 62 or 63 (32,460)
65. 48 and 64 (222)
66. 51 or 65 (316)
67. (visual adj analogue adj scale*1).tw. (11,051)
68. ("linear analogue" adj5 (assessment*1 or scale*1)).tw. (329)
69. 48 and (67 or 68) (11)
70. 66 or 69 (326)

Appendix 12 Ongoing trials identified

Title (country); trial number	Study dates	Population	Intervention	Comparator	Outcomes
Use of Bispectral Index (BIS) for Monitoring of Total Intravenous Anaesthesia in Paediatric Patients (Denmark); NCT01043952	January 2010–September 2012 (ongoing)	Children undergoing ear, nose and throat surgery (aged 1–65 years; stratified by age and surgery type)	BIS-guided anaesthesia with propofol and remifentanil	Standard clinical practice anaesthesia with propofol and remifentanil	**Primary**: anaesthetic consumption; time to extubation **Secondary**: analgesia consumption; device values
Intraoperative depth of anaesthesia and influence on the incidence of postoperative cognitive deficits: a prospective, randomised, controlled, two-armed single-centre pilot trial (Germany); ISRCTN36437985	March 2009–February 2012 (record indicates completed but no publications referenced)	Adults aged ≥60 years undergoing elective GA with a planned duration of procedure ≥1 hour	Unblinded BIS monitoring (anaesthetic not specified)	Blinded BIS monitoring (anaesthetic not specified)	**Primary**: postoperative delirium incidence (DSM-IV) **Secondary**: device values; postoperative delirium (alternative delirium scores); postoperative cognitive dysfunction; time to discharge (recovery room; hospital); length of stay (recovery room; hospital); QoL (EQ-5D); organ dysfunction at hospital discharge; postoperative pain
Bispectral Index (BIS) Monitoring in Abdominal Surgery (Croatia); NCT01470898	February 2011–February 2012 (ongoing)	Adults aged ≥18 years undergoing major abdominal surgery	BIS-guided anaesthesia with sevoflurane and muscle relaxant	Routine anaesthesia care with sevoflurane and muscle relaxant	**Primary**: device values **Secondary**: effect of BIS monitoring on faster recovery time in abdominal surgery patients; time to extubation

DSM-IV, *Diagnostic and Statistical Manual of Mental Disorders*-Fourth Edition.

EME
HS&DR
HTA
PGfAR
PHR

Part of the NIHR Journals Library
www.journalslibrary.nihr.ac.uk

Published by the NIHR Journals Library