Technology Assessment Report commissioned by the NIHR HTA Programme on behalf of the National Institute for Health and Clinical Excellence – Diagnostics Assessment Report

Depth of anaesthesia monitoring (E-Entropy, Bispectral Index and Narcotrend)

Produced by	Southampton Health Technology Assessments Centre (SHTAC), University of Southampton
Authors	Jonathan Shepherd Jeremy Jones Geoff Frampton Jackie Bryant Louise Baxter Keith Cooper
Correspondence to	Jonathan Shepherd
Date completed	2nd April 2012

Source of funding: This report was commissioned by the NIHR HTA Programme as project number 11/57

Declared competing interests of the authors

None

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.

Rider on responsibility for report
The views expressed in this report are those of the authors and not necessarily those of the NIHR HTA Programme. Any errors are the responsibility of the authors.

This report should be referenced as follows:

J Shepherd, J Jones, G Frampton, J Bryant, L Baxter, K Cooper. Depth of anaesthesia monitoring (E-Entropy, Bispectral Index and Narcotrend) *Health Technology Assessment* 2012

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.

G 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.

Commercial in confidence information highlighted and underline in blue

Table of Contents

Tables

Figures

1 DEFINITION OF TERMS AND 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.

Term	Definition
ASA	American Society of Anesthesiologists
BAG-RECALL	BIS or Anesthetic Gas to Reduce Explicit Recall
BIS	Bispectral Index
BNF	British National Formulary
CI	Confidence interval
ECG	Electrocardiogram
EEG	Electroencephalography
ETAC	End-tidal anaesthetic concentration
FGF	Fresh Gas Flow
GA	General anaesthesia
HRQoL	Health Related Quality of Life
ICER	Incremental Cost-effectiveness Ratio
ITT	Intention to treat
LPS	Late psychological symptoms
MAC	Minimum Alveolar Concentration (MAC)
NICE	National Institute for Health and Clinical Excellence
NR	Not Reported
NS	Not statistically significant
PACU	Post Anaesthesia Care Unit (PACU)
POCD	Post-operative cognitive dysfunction
PONV	Post-Operative Nausea and Vomiting
PTSD	Post-Traumatic Stress Disorder
QoL	Quality of Life
QALY	Quality Adjusted Life Year
RCT	Randomised Controlled Trial
TIVA	Total intravenous anaesthesia
TTO	Time Trade Off
SG	Standard Gamble

2 EXECUTIVE SUMMARY

Background

It is important that the level of general anaesthesia 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 post-operative nausea, vomiting and varying degrees of 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. Awareness can be assessed by interviewing the patient after general anaesthesia and asking them whether they recall any experiences between going to sleep and waking up ('explicit' awareness). It may also be assessed during general anaesthesia by asking the patient to respond verbally or physically to commands, or after general anaesthesia via memory tests ('implicit' awareness). Intraoperative awareness is a relatively rare event with an incidence typically around 1-2 patients per 1000. However, over time awareness may cause depression, anxiety and post-traumatic stress disorder (PTSD).

During general anaesthesia 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). However, clinical observation alone may not be a reliable surrogate marker of anaesthetic depth. 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 which 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 Bispectral index (BIS), E-Entropy and Narcotrend technologies to monitor the depth of anaesthesia in surgical patients undergoing general anaesthesia.

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. Due to the relatively large volume of evidence for BIS we only included 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 on the cost effectiveness of depth of anaesthesia monitoring compared to 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.

Targeted literature searches were undertaken for: studies reporting costs of anaesthetics or estimates of anaesthetic consumption against duration of anaesthesia; studies reporting incidence of intraoperative awareness in general surgical populations and in those populations identified as being at greatest risk of awareness; and studies describing symptoms of patients who had reported intraoperative awareness in order to understand the health-related consequences.

It was assumed that a proportion of patients who experience awareness will suffer psychological symptoms and that a proportion of those will develop post-traumatic stress disorder (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 which are attached to the patient. Equivalent annual costs for each monitor were calculated for an effective equipment life of five years. Unit costs of anaesthetic drugs were taken from the British National Formulary (BNF). 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. Costs of treating PTSD have been estimated based on assumptions contained in the national cost impact report associated with the NICE Clinical Guideline on the management of PTSD in adults and children in primary and secondary care.

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 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 to standard clinical monitoring (NB. 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 due to limitations in reporting of methodological details. Uncertainty was greatest in relation to concealment of the random

allocation process, and blinding of outcome assessment. The risk of bias associated with random sequence generation was generally low, though unclear in some trials. There didn't appear to be a high risk of bias associated with selective reporting of outcomes.

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. less 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.

Surgical procedures undertaken included open heart surgery, major orthopaedic procedures, abdominal surgery, and microwave coagulation for liver cancer. In children procedures included tonsillectomy and/or adenoidectomy, urogenital / urologic surgery, and dental rehabilitation. 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.

Commonly reported outcomes included anaesthetic consumption, recovery outcomes (e.g. time to extubation, time to eye opening), and time from the end of surgery to the recovery room (post-anaesthesia care unit, PACU). Adverse outcomes associated with general anaesthesia were less commonly reported.

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 Odds Ratio was 0.45 (95% CI 0.25, 0.81) in favour of BIS. Caution is advised in the interpretation of this result as, overall, there was statistically significant heterogeneity ($p=0.009$, $I^2 = 79\%$). The sub-group of trials which included a trial of mixed inhaled and intravenous anaesthesia, and the sub-group which included trials of total intravenous anaesthesia, both statistically favoured BIS monitoring. However, in the sub-group of trials which used only inhaled anaesthesia the Peto Odds Ratio was 1.79 (95% CI 0.63, 5.11) favouring standard clinical monitoring, though not statistically significant.

None of the trials reported the longer-term detrimental impact of awareness, though one did report patient distress and sequelae as a *post hoc* secondary outcome. There was a higher percentage of distress reported in the BIS monitored group, but no statistically significant difference between groups. Implicit awareness was reported in only one trial, and there was no statistically significant difference between EEG monitoring and standard clinical monitoring.

There were mixed findings for changes in anaesthetic consumption: statistically significant reductions favouring all three types of EEG monitoring were reported, notably trials where inhaled anaesthetic was used for maintenance; in other trials EEG monitoring was associated with reductions in consumption but these were not statistically significant. In general EEG monitoring with all three technologies was associated with statistically significantly shorter times from end of surgery to admission to / discharge from the PACU. Reductions in the time needed to recover from general anaesthesia such as time to tracheal extubation and time to eye opening were statistically significant for all three technologies in the majority of trials. Post-operative nausea and vomiting (PONV) was reported in a handful of trials, and there were generally no statistically significant differences between groups. Not all of the trials were statistically powered for all of these outcome measures so caution is advised.

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 $4,410. The authors of the study concluded that BIS monitoring did not appear cost effective. However the results and conclusions should be viewed with caution due to poor methodological and reporting quality.

Economic evaluation

For each technology we presented a base case analysis for two modes of anaesthetic administration (total intravenous anaesthesia (TIVA) and mixed anaesthesia (induction with IV anaesthesia and maintenance with inhaled anaesthesia or a combination of inhaled and IV 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).

BIS compared with standard clinical monitoring

In cohorts of 10,000 patients, at high risk of intraoperative awareness, undergoing general anaesthesia 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 6 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. By reducing the incidence of awareness and longer-term effects of post-operative cognitive dysfunction BIS monitoring was associated with improved outcomes. The ICER, for BIS compared with standard clinical monitoring in this population was £27,345. Deterministic sensitivity analyses indicated 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 indicated 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).

For the population of general surgical patients, undergoing general anaesthesia 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 4 cases of LPS (from 5.2 to 1.3), which included a reduction of 2 cases of PTSD (from 2.8 to 0.7). While 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). 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 (£45,033) despite the lower incremental cost estimated for this population, arising from reduced anaesthetic consumption. Deterministic sensitivity analyses indicated the ICER was sensitive to the same input parameters as for the population at high risk of awareness. In all cases the ICER remained above £30,000 per QALY gained – the most favourable ICER was associated with a reduction in the cost of sensors.

The baseline estimates of awareness, LPS and PTSD for patients undergoing mixed general anaesthesia were the same as for high risk patients undergoing TIVA (45, 14.7 and 8 per 10,000 patients, respectively). However, given the odds ratio 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 £36,126. Deterministic sensitivity analyses indicated the ICER was sensitive to the same parameters as for high risk patients undergoing TIVA.

The baseline estimates of awareness, LPS and PTSD in the population of general surgical patients, undergoing mixed general anaesthesia, were the same as for TIVA (16, 5.2 and 2.8 per 10,000 patients, respectively), while 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 (£16.23 compared with £14.20). 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 (£61,869). Deterministic sensitivity analyses indicated 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.

Entropy compared with standard clinical monitoring

Insufficient evidence was identified to estimate the effectiveness of depth of anaesthesia monitoring with Entropy on the incidence of intraoperative awareness or on post-operative cognitive dysfunction. In the absence of evidence specific to Entropy we have applied the effectiveness estimates derived for BIS, described above. This meant that the modelled clinical effectiveness of Entropy was identical to that reported for BIS – this is an untested assumption and must be considered a weakness in the evidence base for Entropy.

In patients at high risk of awareness, undergoing general anaesthesia with TIVA, the modelled cost per patient with 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. Entropy monitoring was associated with improved outcomes, based on applying clinical effectiveness evidence reported for BIS. The ICER for 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 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.

In the population of general surgical patients, undergoing general anaesthesia with TIVA, Entropy monitoring had a higher cost per patient than standard clinical monitoring. There was no reduction in anaesthetic drug costs (based on evidence reported from two clinical trials) to offset the additional costs of 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 (£3,131 - £31,430). Deterministic sensitivity analyses indicated the ICER was sensitive to the same variables as for high risk patients.

In patients considered at high risk of awareness, undergoing mixed general anaesthesia, 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 Entropy compared with standard clinical monitoring in this population was £19,367. Deterministic sensitivity analyses indicated the ICER was sensitive to the same variables as for high risk patients, undergoing TIVA.

In the population of general surgical patients, undergoing mixed general anaesthesia, Entropy monitoring had higher costs than standard clinical monitoring. In contrast to 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 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 Entropy monitoring than for high risk patients, the lower incremental cost resulted in an equivalent ICER (£19,000). Deterministic sensitivity analyses indicated the ICER was sensitive to the same input parameters as for the population at high risk of awareness.

Narcotrend compared with standard clinical monitoring

Insufficient evidence was identified to estimate the effectiveness of depth of anaesthesia monitoring with Narcotrend on the incidence of intraoperative awareness or on post-operative cognitive dysfunction. 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 at high risk of awareness, undergoing general anaesthesia 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. In contrast to BIS and 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 £5,681. 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.

In the general surgical population, undergoing general anaesthesia with TIVA, 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 all the deterministic sensitivity analyses.

In patients at high risk of awareness, undergoing mixed general anaesthesia, 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 £8,033. Deterministic sensitivity analyses indicated the ICER was sensitive to the same parameters as for high risk patients undergoing TIVA.

In the population of general surgical patients, undergoing mixed general anaesthesia, Narcotrend monitoring had higher costs than standard clinical monitoring. The reduction in cost of anaesthetic was sufficient to offset the additional cost of Narcotrend monitoring. 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 in the base case and remained dominant in the majority of deterministic sensitivity analyses. Conclusions from the scenario analyses were similar to those undertaken for high risk patients.

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 to 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 not associated with a statistically significant reduction in intraoperative awareness in patients classified as at higher risk, though there is uncertainty in effect estimates due to significant heterogeneity. Caution is advised due to 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.

3. BACKGROUND AND DEFINITION OF THE DECISION PROBLEM

3.1 Condition and aetiology

3.1.1 Background

When patients undergo surgical procedures under general anaesthesia 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 post-operative 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 post-operative 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 over-dosing the anaesthetic, since 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 post-operatively (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 intra-operatively 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 general anaesthesia without recollection of this in the post-operative 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]

3.1.2 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 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 two years[11] or five 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.

3.1.3 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 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 1-2 patients per 1000, although a considerably lower incidence of 0.07 per 1000 patients was found in the largest study which included over 87,000 patients whilst a higher incidence of 4.1 per 1000 patients was found in a Chinese study (Table 1).

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	Not reported	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 (1-2 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)

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]

3.1.4 Risk factors for intraoperative awareness

Some groups of patients undergoing general anaesthesia are at increased risk of intraoperative awareness because they cannot tolerate adequate doses of anaesthetic; because signs of inadequate anaesthesia are masked; or due 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 done under regional anaesthesia (epidural or spinal) rather than under general anaesthesia. 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 whilst 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 whilst experiencing intraoperative awareness, and have suffered long-term psychological ill-health. Around half of all operations under general anaesthesia involve 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]

3.1.5 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 healthcare 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 under-estimated. 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.

3.1.6 Measurement of intraoperative awareness

Basic signs of intraoperative awareness during anaesthesia include tachycardia (rapid heart rate), hypertension, sweating, lacrimation (tear production), movement/grimacing, 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 they recall having been aware during the period of anaesthesia. During surgery the isolated forearm technique (IFT) 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 inflated above systolic blood pressure to isolate a 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 IFT has not been widely used in practice, though has been used as a research tool in a number of studies.[21;23]

The most popular approach for post-operative 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 whilst 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 have looked into the psychometric properties of the interview questionnaires that they 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 post-operative 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 changes in behaviour as a result of implicit awareness are associated with longer-term morbidity.

3.1.7 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 post-operative nausea and vomiting 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 to attain a specified recovery score); consumption of general anaesthetic or other drugs (such as analgesics and anti-nausea agents); and assessment of cognitive or neurological function.

3.2 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 section 3.3). 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 which 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 which 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, 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 Section 3.3). 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 NICE with relevant stakeholders, three of the technologies currently available were prioritised for the current assessment: Bispectral index, E-entropy, and Narcotrend.

3.2.1 Bispectral Index (BIS) (Covidien)

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 general anaesthesia is 40-60 which indicates a low probability of consciousness.

BIS 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, Philips, Dräger). This works via the BISx or BISx4 plug-in connector which allows integration with existing anaesthesia systems.

3.2.2 E-Entropy module (GE Healthcare)

Entropy monitoring in anaesthesia has been studied over the last ten years. E-Entropy (previously known as M-Entropy) is designed to aid the management of general anaesthesia 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 (FEMG) data to produce two values that indicate the depth of anaesthesia. The first value, response entropy (RE), is based on both EEG and FEMG signals and provides an indication of the patient's responses to external stimuli and may signal early awakening. The second value, state entropy (SE), 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 RE-SE 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 vary 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 RE scale ranges from 0 (no brain activity) to 100 (fully awake) and the SE scale ranges from 0 (no brain activity) to 91 (fully awake). The clinically relevant target range for entropy values is 40-60. RE and SE 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 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.

3.2.3 Narcotrend (Narcotrend)

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 a 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 to F scale, an EEG index (100 = awake, 0 = very deep hypnosis) is also calculated.

The Narcotrend-Compact M is a stand-alone monitor which 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 Nacrotrend algorithms are revised continually.

3.2.4 Subgroups of patients

Unsuitable patient populations include those undergoing specific surgical procedures where 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 only validated for patients over the age of 2 years, it is not for patients undergoing procedural or conscious sedation, and seizure activity may cause interference. Also, E-Entropy readings may be inconsistent when monitoring patients with neurological disorders, or with 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.

3.2.5 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 source (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 on screen.

3.2.6 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. [redacted]

3.2.7 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 (though 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 on-line multimedia courses. For E-Entropy a 30 minute introductory training is suggested for healthcare staff before using E-Entropy with particular attention being paid to sensor application. A one day visit from staff to give a lecture and to demonstrate the use of the Nacrotrend in the operating theatre is judged sufficient training by the manufacturer for the majority of Nacrotrend users.

3.3 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 rhythm); 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 (tear production), and sweating.

End-tidal anaesthetic gas concentrations (ETAC) 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 Section 3.4), and anaesthesia can be titrated to keep within a certain MAC range.

Of all the signs and variables, the key things to observe are end-tidal anaesthetic gas concentrations (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]

3.4 Care pathways

In UK health care settings general anaesthesia 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 general anaesthesia, and continues until surgery is complete and the patient is moved from the theatre to the recovery room (also referred to as the Post-Anaesthesia 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 or inhalational. Propofol is a commonly used intravenous anaesthetic and can be used for induction and / or maintenance of anaesthesia. Use of an intravenous anaesthetic for induction and maintenance is sometimes referred to as total intravenous anaesthesia (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 intravenous or volatile agents, in a concentration of 50 to 66% in oxygen.[33] Volatile anaesthetics can be used for induction and maintenance of anaesthesia, and also following induction with an intravenous 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 the potency of isoflurane. It is not recommended for induction of general anaesthesia. 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 to 60 years; and 5.2, 1.5 and 1.0 for people older than 65 years, respectively.[34]

3.5 Summary of the decision problem

As has been described, the purpose of anaesthesia monitoring is to ensure adequate sedation of the patient under general anaesthesia. 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-term 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 electroencephalography (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 general anaesthesia.

4. ASSESSMENT METHODS

4.1 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 in Appendix 1.

4.1.1 Identification of studies

A search strategy was developed for Medline and pilot tested by an experienced information scientist. The Medline strategy (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); Cochrane Central; Cochrane Library (Cochrane Database of Systematic Reviews; Cochrane Central Register of Controlled Trials); Database of Abstracts of Reviews of Effectiveness (DARE); and Health Technology Assessment Database (HTA). For 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 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 upon a Cochrane systematic review,[34] which contained 31 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 section 4.1.4 for further information about how results from the Cochrane review are integrated into the current review.) For studies of E-entropy and Narcotrend we searched from 1995 to November 2011 (and then updated in February 2012).

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, Entropy and Narcotrend as supplied by the device manufacturers (via NICE);
- Reference lists of included studies;
- Databases of research in progress, searched on 07 December 2011: 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 (section 4.1.2 below). 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.

4.1.2 Inclusion/exclusion criteria

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 (Appendix 3). In addition to the language and publication type restrictions, the following selection criteria were applied:

Population

- Included: Patients who received general anaesthesia for surgery, including adults and children (over the age of two years) in whom the technology is licensed.
- Excluded: studies of patients receiving sedation in intensive care or high dependency units; studies in healthy volunteers; studies in non-surgical anaesthesia.

Diagnostic technologies:

- Included: E-Entropy; BIS; Narcotrend

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, anti-sickness drugs and muscle relaxants.

Study design: 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: section 4.1.3).

4.1.3 Data extraction and critical appraisal methods

A standardised data extraction and quality appraisal template (Appendix 5) was used to extract information on the relevant study characteristics for assessing the impact of the interventions on the outcomes listed above (section 4.1.2) 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 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 (section 5.1).

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.

4.1.4 Method of data synthesis

Analyses of the three monitoring devices are presented in respective separate sub-sections of this report (Section 5.1). 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 only included 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 version 5.1.6 to conduct the meta-analyses.

4.2 Systematic review of cost-effectiveness

4.2.1 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 (MEIP); EMBASE; The Cochrane Library including Central Register of Controlled Trials (CENTRAL) and Cochrane Database of Systematic Reviews (CDSR); Centre for Reviews and Dissemination (CRD) including Health Technology Assessment Database, Database of Abstracts of Reviews of Effects (DARE) and National Health Service Economic Evaluation Database (NHSEED); ECONLIT. Bibliographies of retrieved articles were checked for any additional references, and the expert advisory group were contacted to identify additional published and unpublished studies.

4.2.2 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 using criteria in Table 2.

Table 2 Inclusion/ exclusion criteria for screening titles and abstracts

Criterion	
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

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, post-operative morbidity or mortality, health-related quality of life) were not included in the review, but were retained to inform the development and population of the decision analytic model.

4.2.3 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 upon that proposed by Drummond and colleagues[36] and Philips and colleagues[37], see Appendix 6 .

4.2.4 Method of data synthesis

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

4.3 Economic evaluation

We developed a decision analytic model was developed 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, anti-sickness 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]

4.3.1 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 maybe associated with depth of anaesthesia monitoring are included in the base case analysis based upon 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 been 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 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.

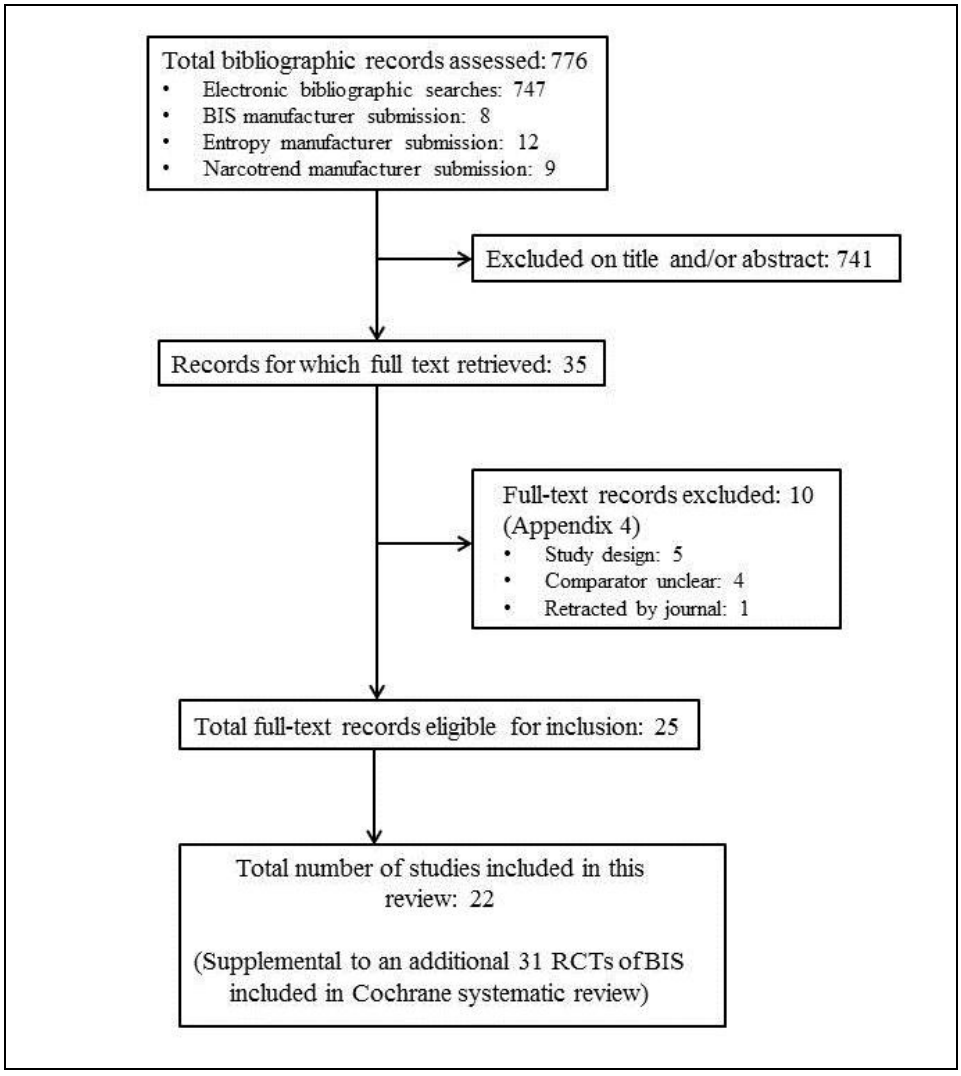
5. ASSESSMENT RESULTS

5.1 Results of systematic review of patient outcomes

5.1.1 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, Entropy and Narcotrend monitors (see Figure 1).

Figure 1 - PRISMA flow chart showing the study selection process for bibliographic records (excluding those already identified in a Cochrane systematic review of BIS studies)



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 (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 evaluating BIS. Due to 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, Entropy or Narcotrend) and a comparator that reflected standard clinical monitoring.

The 22 included studies were 2-arm or 3-arm RCTs that compared the following technologies against standard clinical monitoring:

- BIS alone: 11 studies[40;44-53]
- Entropy alone: 5 studies[54-58]
- Narcotrend alone: 2 studies[59;60]
- BIS and entropy: 2 studies[61;62]
- BIS and Narcotrend: 2 studies[63;64]

These 22 studies provide 15 comparisons of BIS against standard clinical monitoring, seven comparisons of entropy against standard monitorng, and four comparisons of Narcotrend against standard monitoring (Table 3).

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 *et al*[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]	●		

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 sub-sections for the following reasons:

- One of the trials of BIS and Entropy compared to standard clinical monitoring was included in the Cochrane BIS review[61], and therefore is only described within the Entropy sub-sections of this report (i.e. for the comparison of Entropy with standard clinical monitoring).
- Two of the trials of BIS and Narcotrend compared to standard clinical monitoring were included in the Cochrane BIS review[63;64] and are therefore only described within the Narcotrend sub-sections 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 section 5.1.3) 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 BIS 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 judgments for the 31 RCTs in the Cochrane BIS review are not tabulated in this report, but are summarised in the text below).

Table 4- Summary of Risk of Bias - BIS

	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 *et al*[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

In many cases the risk of bias in the trials was unclear due to 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 unclear.

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 low risk of bias due to 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 low risk of bias due to adequate random sequence generation. Most of the remainder were unclear due to 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 due to 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 due to the study being funded in part by the BIS module manufacturer.[45]

The trials varied in terms of their sample sizes, from as low as 20 patients to over 6000. There were six (46%) that included under 100 patients; and five (39%) that had between 101 and 200 patients. One trial included 921 patients[47], another included 5309,[40] and another, the largest, that included 6041 patients.[44] In the Cochrane BIS review[34] the majority of trials included less 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 2007 - included 1961 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] whilst a second publication focuses on recovery outcomes from the trial[74]).

Six (55%) of the 11 BIS trials reported a statistical sample size calculation based on a nominated primary outcome, though 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%) 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 due to 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 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. The trial by Banister and colleagues[45] stated that Aspect Medical Systems (AMS) supplied the BIS monitor, and that one author was employed by AMS and another author was a paid consultant to AMS. This funding therefore represents a conflict of interest. The trial by Kerssens and colleagues[49] reported that AMS did not financially support the study, but that the lead author had received an educational grant in support of her salary from AMS, and one co-author was a paid consultant to AMS. In the trial by Zhang and colleagues[40] AMS 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 entropy trials

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

Table 5 - Summary of Risk of Bias - 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 in the entropy trials was unclear in many cases due to 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 entropy trials.

Risk of bias due to 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 due to 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 because sealed envelopes were used for allocation codes but it was not stated whether the envelopes were opaque.[54-57]

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

In three of the seven entropy trials, the risk of attrition bias due to 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 intention to treat 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 due to incomplete outcome data, either because attrition was not reported at all[55] or because it was not reported separately by study arm.[57]

The risk of bias due to selective reporting of outcomes was judged to be low for six of the seven 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 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 entropy trials were supported by the 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 due to conflict of interests. The study which involved the most extensive links with the manufacturer[57] was deemed by the reviewers to be at high risk of bias due to a high likelihood of conflicting interests. In the four entropy trials that were not supported by the 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 entropy studies were published during 2005 to 2010 and ranged in their total sample size from 50 to 335 patients. Five of the trials involved a two-arm comparison of entropy against standard clinical monitoring.[54-58] One trial involved a three-arm comparison of BIS, entropy and standard clinical monitoring.[61] The remaining trial was a three-arm comparison of entropy, entropy and BIS, and standard practice.[62] The number of patients randomised per arm ranged from 25 to 40 in six trials. In the seventh (largest) trial, only the number per arm after attrition (160 patients) was reported.[57]

Only one of the 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% to 11% of the total population per trial, or 0% to 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 due to ‘technical problems’ in one trial;[54] and were not specified in the remaining trial.[58]

An analysis by intention to treat (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, since it was stated that all patients were included into 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. As noted above, two of these trials[92;93] were considered at high risk of bias due to their incomplete outcome data.

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.

In many cases the risk of bias in the trials was unclear due to 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.

Table 6 - Summary of Risk of Bias - Narcotrend

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

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 due to 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 only reported for one study where the paper was translated from Chinese to English prior to publication and it is unclear whether 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 attrition happened pre-or post-randomisation. All four trials 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 study[60] reported that the study was supported by a pharmaceutical company and a university institutional research grant.

5.1.2 Characteristics of included studies - BIS

The following sub-sections 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 between four to six years, and age ranges from two to 18 years.[45;46;51-53] The remaining six studies were conducted in adults, 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 to 91kg in the adult studies, and 17 to 28 kg in the children studies. In addition to reporting weight in kg, 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 only reported in one trial, in which the population was predominantly (>80%) classified as White.[44] The countries in which the trials were conducted included: 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, with one trial taking place in two centres, another one taking place in three centres, and a third study not reporting the number of centres.

The type of surgery reported in the adult trials varied, from open heart[44] or major non-cardiac[47] or major orthopaedic[49] to orthopaedic[62], and elective moderate abdominal surgery.[48] The surgical procedures in the trials of children included: tonsillectomy and/or adenoidectomy[45]; urogenital / urologic 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; American Society of Anesthesiologists (ASA) status 4; end-stage lung disease; history of intraoperative awareness; history of or anticipated difficult intubation; cardiac ejection fraction <40%; and marginal exercise tolerance. Thetrial by Zhang and colleagues[40] included patients receiving total intravenous anaesthesia (TIVA) which they cited as a risk factor for intraoperative awareness. The Cochrane BIS review[34] included four trials which 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 co-morbidities, 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 to 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] (though 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 were classified as III (severe systemic disease). 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 a BIS module was used.[44;47]. Most commonly reported was the BIS Monitor Model A-2000 as mentioned in four trials. 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)'.[48] One trial used BIS (version 3.3, Aspect Medical Systems) using an A-1050 EEG monitor, [45] whilst 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] Whilst 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]

Whilst all of the trials compared BIS against standard clinical monitoring, the parameters monitored varied. Only one trial measured end-tidal anaesthetic agent concentration (ETAC) in order to detect possible intraoperative awareness.[44] Audible alarms sounded if the age-

adjusted minimum alveolar concentration (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 was guided by the presence of clinical signs in relation to the BIS value. If the patient exhibited hypertension or tachycardia and the BIS was greater than 60 then sevoflurane was increased. If BIS was in the target range of 50-60 then fentanyl was given. If BIS was less than 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 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, e.g. 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 trials mentioned that 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. 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 intravenous and inhalational general anaesthetic was used. 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, though did not provide information on which agent was used.[62] One trial mentioned that regional anaesthesia was used for post-operative pain management.[49] The remaining nine trials 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 use of analgesia was not stated.[44;47;48] Pre-medication with midazolam was used in seven trials.[40;44-46;52;53;62] In two of these trials ketamine was also used as pre-medication.[52;53]

Muscle relaxants were used in seven of the trials, including atracurium,[46;48] cisatracurium, [62] 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 post-operative nausea and vomiting (n=3 trials); and emergence delirium (n=1 trial[59]). One trial, by Leslie and colleagues[95], reported stroke, myocardial infarction, mortality for all surviving and available patients after 30 days post-operation.[95] 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: intra-operative awareness) was included in the Cochrane BIS review[34]).

Table 7 - BIS study outcomes

Study	Avidan[44]	Bannister[45]*	Bhardwaj[46]*	Chan[47]	Ellerkmann[62]	Leslie[50]	Kamal[48]	Kerssens[49]	Liao[51]*	Messieha 2004[52]*	Messieha 2005[53]*	Zhang:[40]*	*Total this review*	*Total Cochrane BIS Review*[34]	*Grand total*
Outcomes															
Anaesthetic consumption		X	**P**	X	**P**			X	X				6	24	30
Intraoperative awareness	**P**				X		X	**P**	X			**P**	6	4	10
Distressing experience of awareness	X												1	0	1
Analgesic consumption		X	X		X		X		X				5	6	11
Muscle relaxant requirement													0	2	2
Time to response to commands			X										1	12	13
Time to eye opening			X				X		X				3	19	22
Time to extubation		X	X				X			X	X		5	21	26
Time to laryngeal mask airway removal									X				1	0	1
Time to first movement response		X							**P**				1	0	1
Time to recovery of orientation									X				1	7	8
Time to phonation									X				1	0	1
PACU stay		X	X				X			X			4	12	16
Time to home readiness													0	7	7
Monitoring device values		X						X			X		3	0	3
Post-operative nausea and vomiting									X				1	2	3
Emergence delirium									X				1	0	1

Postoperative cognitive dysfunction (POCD)				X									1	1	2
Parental satisfaction									X				1	0	1
Treatment of haemodynamic events													0	0	0
Haemodynamic profiles		X	X						X				3	3	6
Stroke						X							1	0	1
Myocardial infarction						X							1	0	1
Mortality						X							1	0	1

P = primary outcome measure; X = stated secondary outcome measure / not stated whether primary or secondary outcome measure; * study of children

Six of the 11 BIS trials 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 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 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 the 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 1st and 4th 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 due to 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 due either to insufficient anaesthetic, or to 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 post-operative 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 analyses 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 three days post-operation. In the trial by Ellerkmann and colleagues[96] the interview took place on the first and third day post-operative days, in the trial by Kamal and colleagues[48] the interview took place on first, second and third day postoperatively, and in the trial by Liao and colleagues[51] the timing was not specified. Little detail of the interviews was given other than '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] which 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 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.

5.1.3 Assessment of outcomes – BIS

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.

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.* 2011[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.* 2010[62] n/N (%)	0/27	0/27	-
Kamal *et al.* 2009 [48] n/N (%)	0/28	0/29	-
Kerssens *et al.* 2009[49] n/N (%)	2/67 (3)	1/61 (2)	NR
Liao *et al.* 2011[51a] n/N (%)	0/52	0/54	-
Zhang *et al.* 2011[40] n/N (%)			
Confirmed awareness	4/2919 (0.14)	15/2309 (0.65)	p=0.002; OR=0.21 (0.07 – 0.63)
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

NR = Not reported
[a] study of children

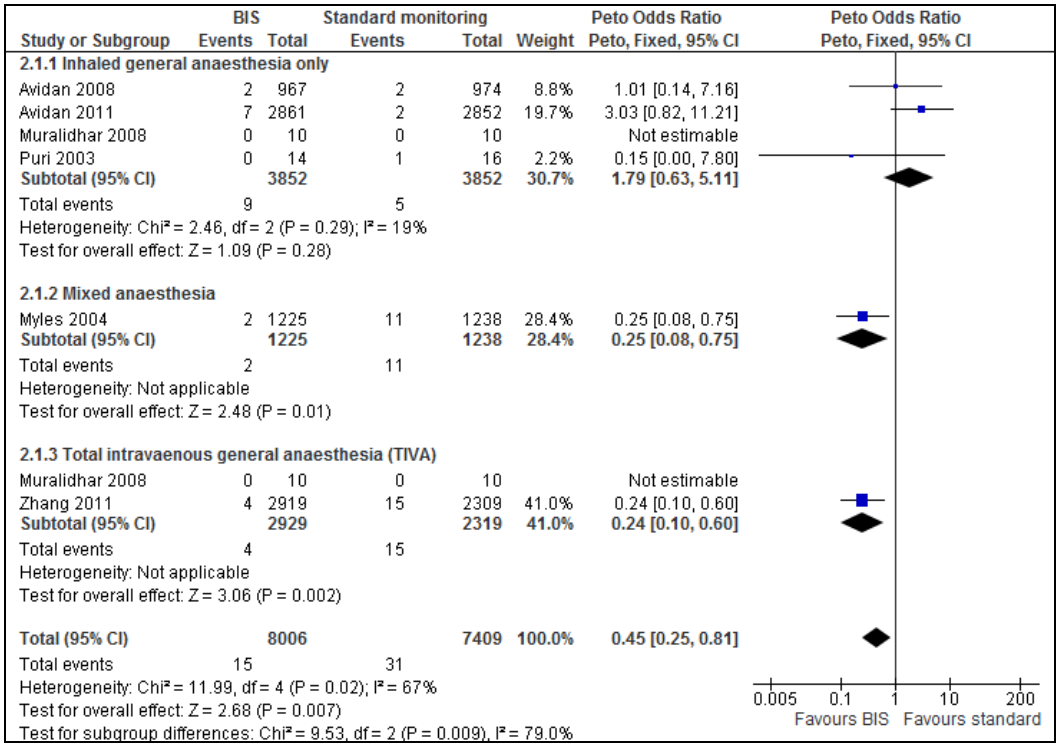
No cases of awareness were reported at all in three trials,[48;51;62] and a very low number of cases were reported in a fourth trial.[49] As stated earlier, these trials were not specifically designed to detect the effect of depth of anaesthesia monitoring on awareness, and therefore are unlikely to 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 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, compared to those who received standard clinical monitoring. Incidence of possible awareness was also lower for BIS monitored patients, though 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 only reported awareness outcomes 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 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 (Avidan and colleagues 2011[44] and Zhang and colleagues[40]). Figure 2 reports the results of this meta-analysis.

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



The meta-analysis included three sub-group analyses: trials which used inhaled general anaesthesia only; trials which used a mixture of inhaled and intravenous anaesthesia; and trials which used total intravenous general anaesthesia (TIVA). The original overall pooled Peto Odds Ratio from the

Cochrane review was 0.33 (95% CI 0.13, 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 odds ratio to 0.45 (95% CI 0.25, 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 sub-group of trials which used only inhaled anaesthesia the Peto Odds Ratio was 1.79 (95% CI 0.63, 5.11) in favour of standard clinical monitoring. This is in contrast to the other two sub-groups 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 post-operatively 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 post-operatively 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 differences between study groups were statistically significant. Intra-group and overall differences between postoperative target and distractor word recall suggest 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≥0.05).

Anaesthetic consumption

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

Table 9 – Consumption of anaesthetic during BIS monitoring

Study	BIS	Standard clinical monitoring	Mean difference (95% CI) P-value
Volatile anaesthetic consumption (sevoflurane) Mean ± SD end-tidal sevoflurane concentration (%)			
Bannister *et al* 2001[45a] Maintenance of GA Last 15 minutes of GA End of procedure	 1.8 ± 0.4 1.6 ± 0.6 1.1 ± 0.6	 2.4 ± 0.6 2.1 ± 0.7 1.5 ± 0.7	 <0.05 <0.05 NS
Kerssens *et al* 2009[49] Maintenance phase	 1.31 ± 0.29	 1.56 ± 0.29	 <0.001
Liao *et al* 2011[51a] Maintenance	 2.5 ± 0.4	 2.9 ± 0.5	 0.001[b] ; <0.01[c]
Propofol consumption			
Bhardwaj *et al* 2010 [46a] Maintenance phase µg/kg/min, Mean (SD)	 108.6 (37.8)	 106.6 (38.9)	 P value NR Mean difference 1.9 (–19.9 to 23.7)
Chan *et al* 2010[47] 25.3% reduction vs Standard clinical monitoring[d]			
Ellerkmann *et al* 2010[62] Maintenance phase µg/kg/min, Mean (SD)	 104 (20)	 101 (22)	 P=0.27 Entropy / BIS vs standard clinical monitoring

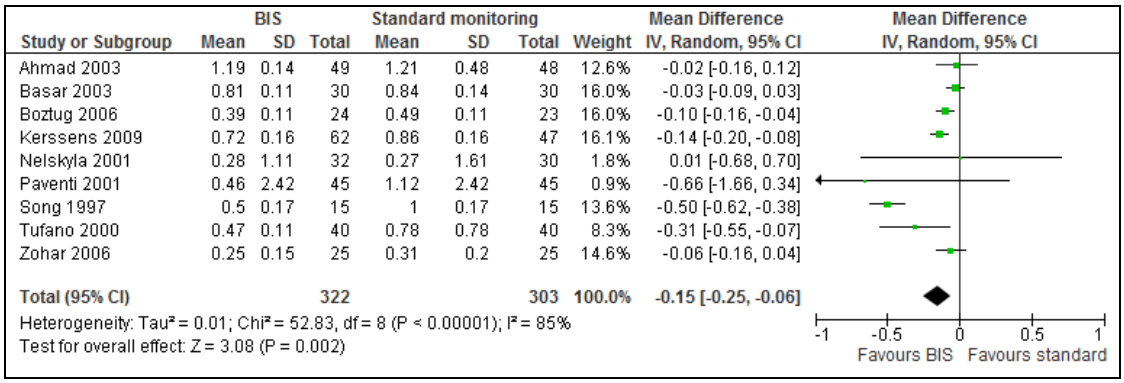
GA = General anaesthesia
[a] study of children
[b] for 3-group comparison (BIS; Auto-regressive index; standard clinical monitoring); [c] post-hoc comparison BIS v Standard clinical monitoring
[d] assumed that this comparison was between BIS and standard clinical monitoring; however the wording of the results does not rule out that the comparison may instead have been between BIS and a matched "control" group.

Six of the 11 BIS trials included in this systematic review reported this outcome measure, 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 general anaesthesia in each of these three trials was statistically significantly lower in the BIS monitored group compared to standard clinical monitoring. The other three trials reported intravenous 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 to 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).

Figure 3– Meta-analysis of volatile anaesthetic consumption (sevoflurane) during BIS monitoring, MAC equivalents

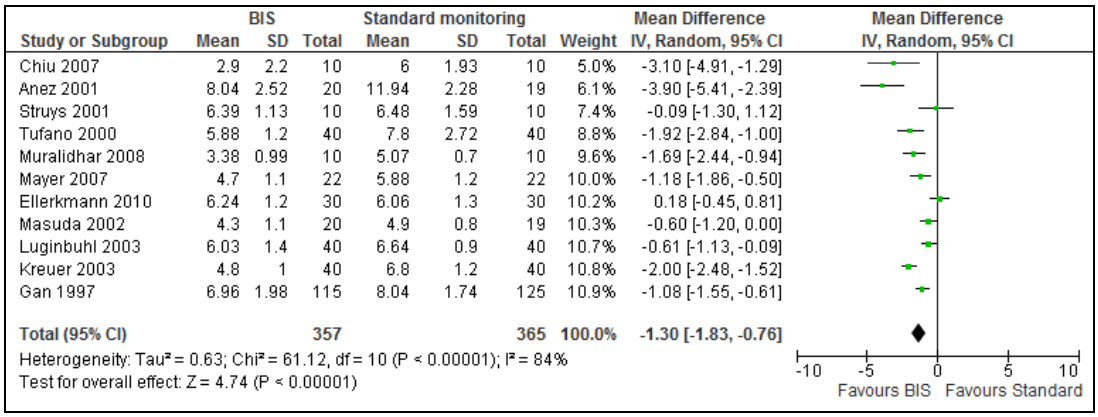


IV = inverse variance

As stated, two of the three of the 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 2009[49]). The original mean difference in MAC equivalents from the Cochrane review for sevoflurane consumption was -0.16 (-0.29, -0.04), indicating a statistically significant difference in favour of BIS. Updating the meta-analysis with the trial by Kerssens and colleagues 2009[49] reduced the mean difference slightly to -0.15 (95% CI -0.25, -0.06), but remained statistically significant. However, caution is advised due to 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.

Figure 4 – Meta-analysis of propofol consumption during BIS monitoring, mg/kg/min



IV = inverse variance

As stated, one of the three of the studies measuring propofol consumption in our systematic review was conducted in children. 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/min) in the Cochrane review was -1.44 (-1.95, -0.93), indicating a statistically significant difference in favour of BIS. Updating the meta-analysis with the trial by Ellerkmann and colleagues 2010[62] reduced the mean difference slightly to -1.33 (95% CI -1.82, -0.84), but remained statistically significant. Again, caution is required due to highly significant unexplained statistical heterogeneity ($P<0.00001$; $I^2 = 80\%$).

Outcomes related to post-anaesthesia care unit (PACU) stay

Five of the 11 BIS trials 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, whilst Kamal and colleagues[48] and both the trials by Messieha and colleagues [52;53] stated measuring the end of general anaesthetic to PACU discharge (though 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), whilst Bhardwaj and colleagues[46] used the Steward recovery scoring system (eligibility = score of 6). Messieha and colleagues[48;52] did not report use of discharge criteria.

Table 10 shows the results of the trials relating to stay in the post-anaesthesia care unit.

Table 10 - Post-Anaesthesia Care Unit (PACU) stay outcomes following BIS monitoring

Study	BIS	Standard clinical monitoring	Mean difference (95% CI) P-value
Bannister *et al* 2001[45a] Mean ± SD time to discharge from the PACU minutes	20.0 ± 7.9	26.7 ± 11.2	<0.05
Bhardwaj *et al* 2010 [46a]	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* 2009 [48] Arrival at PACU (min) PACU discharge (min)	 9.4 (1.9) 53.9 (14.7)	 14.1 (2.8) 78.6 (21.5)	 P<0.01 P<0.01
Messieha *et al* 2004 [52a] Time to PACU discharge, minutes. Mean (SD) Duration of PACU stay, minutes. Mean (SD)	 60 (± 13) 45 (± 8)	 90 (± 11) 71 (± 9)	 <0.001 <0.001
Messieha *et al* 2005[53a] Duration of PACU stay, minutes. Mean (SD)	 47 (± 17)	 63 (± 17)	 0.02

[a] study of children

In all trials time to discharge from the PACU was statistically significantly greater in the standard clinical monitoring group compared to the BIS monitoring group, with mean differences in the range of 6.7 minutes to 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 about how timings may differ according to these different outcomes. Given this lack of clarity, and the fact that the Cochrane review only included 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, -2.76) 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.

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			
Bannister *et al* 2001[45a]	7.1 (3.7)	11.3 (5.9)	<0.05
Bhardwaj *et al* 2010 [46a]	Time to extubation reported to be comparable in the two groups.		
Kamal *et al* 2009 [48]	4.3 (2.1)	4.8 (2.3)	>0.05
Messieha *et al* 2005[52a]	9 (5)	13 (5)	0.07
Messieha *et al* 2005[53a]	5 (2)	10 (7)	0.04

[a] study of children

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 compared to standard clinical monitoring by as much as five minutes, and 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 (SD) in minutes was 1.8 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 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 e was -2.87 minutes (95% CI, -3.74, -1.99), 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 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			
Bhardwaj *et al* 2010[46a]	Time to eye opening reported to be comparable in the two groups.		
Kamal *et al* 2009[48]	4.1 (1.6)	4.4 (1.9)	>0.05
Liao *et al* 2011[51a]	15.0 ± 16.4	16.1 ± 11.3	0.17[b]

[a] study of children

[b] for 3-group comparison (BIS; Auto-regressive index; standard clinical monitoring);

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, though by modest duration (up to one 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, -1.29), 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.

Table 13 -Time to other recovery outcomes

Study	BIS	Standard clinical monitoring	Mean difference (95% CI) P-value
Bannister *et al* 2001[45a] Mean ± SD time to first movement response, minutes	4.2 ± 3.7	7.0 ± 3.9	<0.05
Bhardwaj *et al* 2010[46a] Time to response commands	Time to response to commands reported to be comparable in the two groups.		
Liao *et al* 2011[51a] Time to emergence from anaesthesia, minutes, mean ±SD: Spontaneous movement Phonation	 3.6 ± 2.7 8.4 ± 5.2	 6.1 ± 5.7 12.9 ± 9.0	 0.02 [b]; <0.05 [c] 0.11 [b]

[a] study of children
[b] for 3-group comparison (BIS; Auto-regressive index; standard clinical monitoring); [c] post-hoc comparison BIS v Standard clinical monitoring

All three of the trials 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 reduction in time to first spontaneous movement of 2.5 minutes. This trial 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.

Post-operative nausea and vomiting

Post-operative nausea and vomiting was only reported by one of the trials included in the current systematic review, the trial by Liao and colleagues[51] There was no difference between BIS and standard clinical monitoring patients in 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 Pediatric Anesthetic 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 five 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).

Post-operative cognitive dysfunction

The only trial to report post-operative 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 one and three weeks after surgery (no information on the tests reported). POCD was confirmed when 2 or more test parameters or the combined Z score > 1.96 (no further information given). There was no statistically significant difference between BIS and standard clinical monitoring in rates of dysfunction at one week post-surgery (146 (32.5%); 177 (39.1%), respectively, p=0.07). However, the difference between groups become significant at three months post-surgery (36 (8.1%); 54 (12%), respectively, p=0.03; OR (95% CI) 1.6 (1.0, 2.4). Caution is advised as this trial 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[95], reported stroke, myocardial infarction, and mortality for all surviving and available patients after 30 days post-operation.[95] (Table 14). This is a long-term follow-up (median = 4.1 years) publication of the B-Aware trial in patients classified at higher risk of intraoperative awareness due to 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 2004[79] (primary outcome: intra-operative 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	p-value
Mortality rate per 1000 patient years (95% CI)	67 (60-76)	70 (62-79)	NR
Myocardial infarction, n (%)	105 (9)	111 (9)	NR
Stroke n (%)	53 (4)	62 (5)	NR

NR = Not reported

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

Summary of BIS 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 Odds Ratio was 0.45 (95% CI 0.25, 0.81), in favour of BIS. However, there was statistically significant heterogeneity and a non-significant difference in the sub-group of trials in which only inhaled general anaesthesia 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, -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 (95% CI -1.82, -0.84) mg/kg/min, 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 greater in BIS-monitored patients, with mean differences in the range of 6.7 minutes to 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, -2.76) 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 to standard clinical monitoring by as much as five 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, -1.99) 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, though by modest duration (up to one 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, -1.29), indicating a statistically significant difference in favour of BIS (with unexplained statistically significant heterogeneity).
- Post-operative nausea and vomiting was only reported by 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 incidence of post-operative cognitive dysfunction. There was no statistically significant difference between groups in rates of dysfunction at one week post-surgery. By three months post-surgery incidence had fallen to around 8 to 12%, with a significant difference in favour of BIS. This study was reported only as a conference abstract and it is not clear whether this outcome was adequately statistically powered.
- Longer-term post-operative 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, though 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 sub-group 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.

5.1.4 Characteristics of included studies – Entropy

Study populations

Two of the seven entropy trials were conducted with children, with median age 4–6 (range 3 –12) years.[54;56] The remaining five trials were on 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] whilst 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] whilst another trial included more middle-aged women than men (men : women ratio approximately 1 : 3). The remaining three 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 16kg to 22kg in the child studies and from 65kg to 82kg 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 2 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 entropy trials reported the ethnicity of their participants.

Four of the entropy trials were on 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 was specifically on children undergoing tonsillectomy or adenoidectomy.[54] Another trial was specifically on women undergoing laparoscopic gynaecological procedures.[55] The remaining trial focused on total knee replacement surgery.[58] Only one of the 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 patients had any co-morbidities that affect EEG monitoring. However, all the 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 entropy trials reported that they used the 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 entropy algorithm software used.

The target 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 entropy trials reported that 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 (pre-operative visit) values.[57] The second trial stated that 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 sulfentanil 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 entropy trials, the entropy monitoring was initiated in the operating theatre. Two trials stated[57] or implied[58] that entropy monitoring was started before anaesthesia induction and two trials stated that entropy monitoring began after anaesthesia induction.[55;56] The remaining three trials did not report whether 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 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, intravenous 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 IV propofol for anaesthesia induction.[55;61;62] One trial used IV propofol with alfentanil analgesic for induction.[57] A further trial employed propofol if patients had an IV 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 IV 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 IV agents (all included propofol) for both induction and maintenance;[55;57;62] and two trials used an IV anaesthetic for induction followed by an inhaled anaesthetic for maintenance.[56;61]

Regional anaesthesia was only clearly reported in one of the 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 entropy trials stated that analgesics were not used during induction or maintenance of anaesthesia, although kerotolac was used after anaesthetic cessation.[54] One trial used IV sufentanil during induction and maintenance, with morphine during the last 15 minutes of surgery, followed by paracetamol, nefopam or Non-Steroidal Anti-inflammatory Drugs post-operatively.[61] Two trials used fentanyl during anaesthesia maintenance. Of these, one also used fentanyl and lidocaine during induction,[58] whilst the other used fentanyl post-operatively, 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]

Pre-medication was reported in five of the entropy trials. The agents used were oral hydroxyzine,[61] oral midazolam alone,[62] oral midazolam with a benzodiazepine,[55] IV midazolam, [54] and oral diazepam (in 5/6 study centres).[57] The remaining two trials did not specify whether premedication was used or not.

All of the 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 to 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 to 48.1 minutes,[54] or a median of 29 to 30 minutes (range 15 to 95 minutes),[56] or was not reported.[57;62]

The training and experience of the anaesthetists in entropy module use was reported in four of the seven entropy trials. One trial stated that anaesthetists were allowed to accustom themselves to the use of entropy monitoring for three 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 entropy studies (Table 15).[54;58;61;62] The method of assessing anaesthetic consumption was by weighing the vaporizer,[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.

Table 15 - Entropy study outcomes

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

P = primary outcome measure; X = stated secondary outcome measure / not stated whether primary or secondary outcome measure
* = study of children

The most frequently reported outcomes overall for which quantitative results were reported were: anaesthetic consumption (a primary outcome in four trials 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 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); and postoperative pain (a secondary outcome in two trials). Other outcomes which were reported quantitatively in one trial each were: post-operative pain, analgesia consumption, post-operative nausea and vomiting, 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 since 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 post-operative nausea and vomiting did not differ between entropy and clinical practice groups but no quantitative results were reported for these outcomes and so these are not included in Table 15.

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 Nordstrom and colleagues[97](one study[62]). The three remaining trials stated only that intraoperative recall was assessed by independent nurses;[54] that patients were questioned about memory and awareness;[55] or that 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 post-operative day,[62] in the post-anaesthesia care unit 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 one day post-surgery (for intraoperative awareness) in three trials,[54;55;57] three days post-surgery (for intraoperative awareness) in three trials,[58;61;62] and only two 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 one week post-surgery.

5.1.5 Assessment of outcomes – 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 Ellerkmann 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.

Table 16 – Intraoperative awareness during 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)	Not reported
Choi *et al*[54] [a] n/N (%)	0/39 (0)	0/39 (0)	Not reported
Ellerkmann *et al* [62] n/N (%)	0/30 (0)	0/30 (0)	Not reported
Gruenewald *et al* [55] n/N (%)	0/37 (0)	1/35 (2.8)	Not reported
Vakkuri *et al* [57] n/N [b] (%)	0/160 (0)	0/160 (0)	Not reported
Wu *et al* [58] n/N (%)	0/34 (0)	0/31 (0)	Not reported

[a] study of children
[b] number reported only after attrition

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 entropy-guided anaesthesia group compared to 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 entropy group. In this trial 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 intravenous anaesthetics[55;57;62] showed mixed results (Table 17). Propofol consumption in the 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 a trial that assessed anaesthetic consumption as the primary outcome.[62] Remifentanil consumption was significantly higher in the 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).

Table 17 – Consumption of anaesthetic during entropy monitoring

Study	Entropy	Standard clinical monitoring	Mean difference (95% CI); p-value
Volatile anaesthetic consumption (sevoflurane) Mean ± SD vaporizer weight change			
Aime *et al*[61] Total, g Rate, g/h Rate normalised, g/kg/h [a]	 22.8 ± 14.4 7.8 ± 3.4 0.10 ± 0.05	 25.6 ± 17.2 9.4 ± 5.6 0.14 ± 0.09	 p=0.49 p=0.07 p=0.003
Volatile anaesthetic consumption (sevoflurane) Mean ± SD end-tidal sevoflurane concentration (%)			
Choi *et al*[54] [b]	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 (%)			
Talawar *et al*[56] [b] Immediately before LMAI [c] 15 s after LMAI [c] 15 s after caudal analgesia 15 s after skin incision 5 min after skin incision Immediately before LMAR [d]	 0.81 0.78 0.69 0.68 0.68 0.35	 1.24 1.24 0.84 0.78 0.79 0.38	 p<0.05 p<0.05 p<0.05 p<0.05 p<0.05 p≥0.05
Intravenous anaesthetic consumption (propofol & remifentanil) Mean ± SD consumption rate and number (%) requiring propofol bolus based on entropy			
Ellerkmann *et al*[62] Propofol, µg/kg/min Remifentanil, µg/kg/min Requiring bolus, n/N (%)	 106 ± 24 0.08 ± 0.02 12/30 (40)	 101 ± 22 0.09 ± 0.02 10/30 (33)	 p=0.27 p=0.56 Not reported
Gruenewald[55] Propofol, µg/kg/min Remifentanil, µg/kg/min	 81 ± 22 0.46 ± 0.08	 95 ± 14 0.39 ± 0.08	 p<0.01 p<0.001
Intravenous anaesthetic consumption (propofol & alfentanil) Median (range) consumption rate			
Vakkuri[57] [d] Propofol, mg/kg/min Alfentanil, µg/kg/min	 0.10 (0.04–0.23) 0.60 (0.12–2.2)	 0.11 (0.03–0.21) 0.57 (0.16–1.6)	 p<0.001 p=0.54

[a] normalised to patient body weight and anaesthetic duration
[b] study of children
[c] LMAI = laryngeal mask airway insertion; LMAD = laryngeal mask airway removal
[d] unclear whether data are for whole operation or last 15 minutes (p-value the same for both)

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]

Table 18 - Time to recovery from anaesthesia (before discharge to PACU)

Study	Entropy	Standard clinical monitoring	Mean difference (95% CI); p-value
Time to eye opening Mean ± SD or median (range) [interquartile range] time since cessation of anaesthetic (or time from last suture[62]), min			
Aime *et al*[61]	7.6 ± 4.1	7.2 ± 4.7	Not reported
Choi *et al*[54 a]	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
Gruenewald *et al* [55 b]	3 (0–9) [1–5]	4 (0–14) [3–6]	Stated not significant
Talawar *et al*[56 a, b]	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 to extubation Mean ± SD or median (range) time since cessation of anaesthetic (or start time not reported[57]), min			
Aime *et al*[61]	11.5 ± 5.8	14.2 ± 9.0	Not reported
Choi *et al*[54 a]	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 to spontaneous breathing Median (range) (start time not reported), min			
Vakkuri *et al*[57]	4.74 (0.00–18.0)	7.07 (-1.00–28.5)	p<0.001
Time to recovery of orientation Mean ± SD or median (range) time since cessation of anaesthetic (or start time not reported,[57]) min			
Choi *et al*[54 a]	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 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 to complete recovery (Aldrete score ≥9) Mean ± SD time since cessation of anaesthetic, min			
Choi *et al*[54 a]	24.3 ± 7.3	28.8 ± 5.7	p<0.05
Time to recovery (Steward score of 6) Mean ± SD time since cessation of anaesthetic, min			
Talawar *et al*[56 a]	7.08 ± 3.78 6 (1–15)	8.36 ± 4.8 8 (2–24)	1.3 (-1.2–3.7) p=0.464

[a] Study of children
[b] primary outcome

Time to eye opening was significantly shorter by approximately 2–4 minutes in the entropy group than the standard clinical practice group in two of six trials, one of which assessed this as a primary

outcome in children[56] and the other which assessed it as a secondary outcome in adults.[57] In the remaining four trials (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 entropy group than 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 but statistical significance was not reported in the remaining trial (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 entropy group than the standard clinical practice group in the two trials that reported these outcomes (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 post-anaesthesia care unit (PACU) stay

The time from discharge from the operating room to the post-anaesthesia care unit (PACU) was shorter by approximately 3–4 minutes in the 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 statistically significant, although only marginally so in one of the trials.[56]

The time to discharge from the PACU was shorter in the 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] (Table 19).

Table 19 - Time for discharge to/from PACU

Study	Entropy	Standard clinical monitoring	Mean difference (95% CI); p-value
Time 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], min			
Talawar *et al*[56] [a]	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 to discharge from PACU Median (range) – not stated whether time since discharge from operating room or since admission to PACU, min			
Vakkuri *et al* [57]	134 (50–1,293)	150 (7–1,020)	p=0.21

[a] study of children

Post-operative pain

Two trials reported post-operative 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 entropy group than standard clinical practice for the adult population.[55] In the paediatric population, the CHEOPS scores were significantly lower in the entropy group at 60, 90 and 120 minutes after arrival in the PACU but not at 30 minutes after arrival.[56]

Table 20 – Post-operative 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[55]	6 (2–10) [4–7]	4 (1–10) [3–5]	p=0.03
Pain intensity score based on CHEOPS scale Mean (SE)			
Talawar[56] [a]			
After 30 min in PACU	4.88 (0.319)	4.76 (0.09)	0.12 (-0.53–0.77); p=0.71
After 60 min in PACU	4.48 (0.10)	4.76 (0.08)	-0.28 (4.59–4.92); p=0.01
After 90 min in PACU	4.56 (0.10)	4.76 (0.08)	-0.2 (4.59–4.92); p=0.01
After 120 min in PACU	4.88 (0.21)	5.44 (0.33)	-0.56 (4.77–6.09); p=0.01

[a] study of children

Analgesic consumption

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

Table 21 – Analgesic consumption during 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/h	13.6 ± 6.1	14.9 ± 8.3	p=0.66
Maintenance consumption, µg/kg/h	0.18 ± 0.09	0.22 ± 12	p=0.26

Post-operative nausea and vomiting

One trial that assessed post-operative nausea and vomiting after arrival in the recovery room[55] reported similar frequencies in the entropy and standard clinical monitoring that did not differ significantly (Table 22).

Table 22 – Post-operative 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[55]	15/37 (41)	13/35 (37)	Stated not significant

In addition to the outcomes reported above, the entropy trials reported that the following outcomes did not differ between 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]

Summary of Entropy assessment

- Six trials monitored intraoperative awareness in adults and children receiving different volatile and intravenous 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 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 intravenous 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 entropy group was reported in trials that assessed these as secondary outcomes
- Time to eye opening was significantly shorter in the 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 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 entropy group than in the standard clinical practice group. Except for time to orientation (2 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 2 trials which assessed secondary outcomes only) suggests that entropy monitoring favours shorter time to discharge to and from the PACU, but it is unclear whether the time gains are clinically important.
- No firm conclusions can be drawn about effects of entropy monitoring on post-operative pain because the only two trials that assessed this used different rating scales, and the effect of entropy monitoring on pain scores was temporally variable in one of the trials. Analgesic consumption and frequency of post-operative nausea and vomiting were assessed in one trial each and did not differ between the entropy and standard clinical practice groups. Post-operative pain, nausea and vomiting, and analgesic consumption were only assessed as secondary outcomes in these trials.
- In summary, compared to standard clinical monitoring, entropy monitoring favoured: lower consumption of volatile anaesthetics and some, but not all, intravenous anaesthetics; and shorter times to recovery and discharge to and from the PACU, assessed by various measures. 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, due to the uniqueness of the individual studies (which included different populations in terms of age, gender 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 entropy and standard clinical practice groups.

5.1.6 Characteristics of included studies - Narcotrend

Study populations

In all trials of Narcotrend the study population was adults (mean age 40 to 50 years) and 33-50% of participants were males for the three studies reporting gender. Mean weight ranged from 60kg to about 84kg. 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. Co-morbidities 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 co-morbidities. 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] only reported 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]whilst 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 ECG, heart rate and mean arterial blood pressure were continuously monitored alongside Narcotrend scores. The other three studies did not explicitly state whether standard clinical monitoring took place in addition to Narcotrend or not. 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 intravenous 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 atrcurium[64], cisatrcurium[63] and rocuronium.[60]

Mean duration of anaesthesia ranged from 113 to 125 minutes[64], 108 to 127 minutes[63], 88 to 91 minutes[59], and 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 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 post-operative nausea and vomiting.[59;60]

Table 23 - Narcotrend study outcomes

Study	**Kreuer [64]**	**Kreuer [63]**	**Lai [59]**	**Rundshagen [60]**
Outcomes				
Anaesthetic consumption	X	X	X	X
Intraoperative awareness	X	X	X	X
Analgesic consumption			X	X
Time to response to commands			X	
Time to eye opening	**P**	**P**		
Time to extubation	X	X		**P**
Time to recovery of orientation			X	
Time to arrival at PACU	X	X		
Post-operative nausea and vomiting			X	X

P = primary outcome measure; X = stated secondary outcome measure / not stated whether primary or secondary outcome measure

5.1.7 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 whilst 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 anaesthestic practice, although desflurane consumption per patient per minute was statistically significantly lower in the Narcotrend group.[64]

Table 24 - Anaesthetic consumption

Study	Narcotrend	Standard clinical monitoring	P value
Propofol consumption per patient			
Kreuer et al[63] Mg, mean ± SD mg/kg/hr, mean ± SD	 721.3 ± 401.2 4.5 ± 1.1	 970.5 ± 384.4 6.8 ± 1.2	 <0.05 <0.001
Lai et al[59] mg, mean ± SD	 380 ± 35	 460 ± 30	 <0.01
Rundshagen et al[60] µg/kg/min, mean ± SD	 0.093 ± 0.042	 0.114 ± 0.035	 0.089
Remifentanil consumption per patient			
Kreuer et al[64] Normalised remifentanil infusion rate, µg/kg/min,mean ± SD	0.22 ± 0.06	0.23 ± 0.07	NS
Kreuer et al[63] Normalised remifentanil infusion rate, µg/kg/min,mean ± SD	0.21 ± 0.07	0.20 ± 0.07	NS
Rundshagen et al[60] Remifentanil dose, µg/kg/min, mean ± SD	0.31 ± 0.10	0.34 ± 0.11	NS
Desflurane consumption per patient			
Kreuer et al[64] mg, mean ± SD mg/min, mean ± SD	 4655.9 ± 2891.7 374.6 ± 124.2	 5547.3 ± 2396.4 443.6 ± 71.2	 NS <0.05

NS = not significant

Time to arrival at post-anaesthetic 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 25 - Time to arrival at post-anaesthetic care unit.

Study	Narcotrend	Standard clinical monitoring	P value
Kreuer et al[64] minutes, mean ± SD	8.0 ± 1.9	9.4 ± 2.4	<0.05
Kreuer et al[63] Minutes, mean ± SD	6.6 ± 2.8	12.4 ± 5.7	<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.

Table 26 - Time to eye opening

Study	Narcotrend	Standard clinical monitoring	P value
Kreuer et al[64] minutes, mean ± SD	3.7 ± 2.0	4.7 ± 2.2	Ns
Kreuer et al[63] Minutes, mean ± SD	3.4 ± 2.2	9.3 ± 5.2	<0.001

NS = not significant

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, two other studies which 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]

Table 27 - Time to extubation

Study	Narcotrend	Standard clinical monitoring	P value
Rundshagen et al[60] minutes, mean ± SD	10.6 ± 7.19	9.29 ± 6.23	NS
Kreuer et al[64] minutes, mean ± SD	3.6 ± 2.0	5.0 ± 2.4	<0.05
Kreuer et al[63] Minutes, mean ± SD	3.7 ± 2.2	9.7 ± 5.3	<0.001

NS = not significant

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.[59]

Table 28 - Time to emergence from anaesthesia

Study	Narcotrend	Standard clinical monitoring	P value
Lai et al[59]			
Time to arousal mins, mean ± SD	4.9 ± 2.2	9.5 ± 2.9	<0.01
Orientation recovery mins, mean ± SD	6.6 ± 3.2	12.2 ± 3.5	<0.01

Post-operative 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 versus 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 monitored intraoperative awareness in adults receiving different volatile and intravenous anaesthetics; no patients reported explicit memory during anaesthesia although two patients receiving Narcotrend monitoring recalled dreaming during anaesthesia.
- Three studies that measured consumption of propofol reported different results; significantly lower consumption was found in the Narcotrend group in two studies whilst no difference was reported between groups in the third study.
- Three studies found no significant difference between groups in remifentanil or desflurane consumption.
- Two studies 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 which yielded conflicting results; one reported a significantly lower time in the Nacrotrend group compared with standard care and the other reported no difference between groups.
- Time to extubation was the primary outcome in one study which found no difference between groups; two other studies which reported this measure as a secondary outcome found significantly shorter time to extubation with Narcotrend monitoring compared with standard care.
- Time to arousal and duration of orientation recovery were reported to be shorter with Nacrotrend monitoring compared with standard care in the one study reporting these outcomes.
- Results suggest that there are no differences between groups in post-operative nausea and vomiting 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 total intravenous anaesthesia 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.

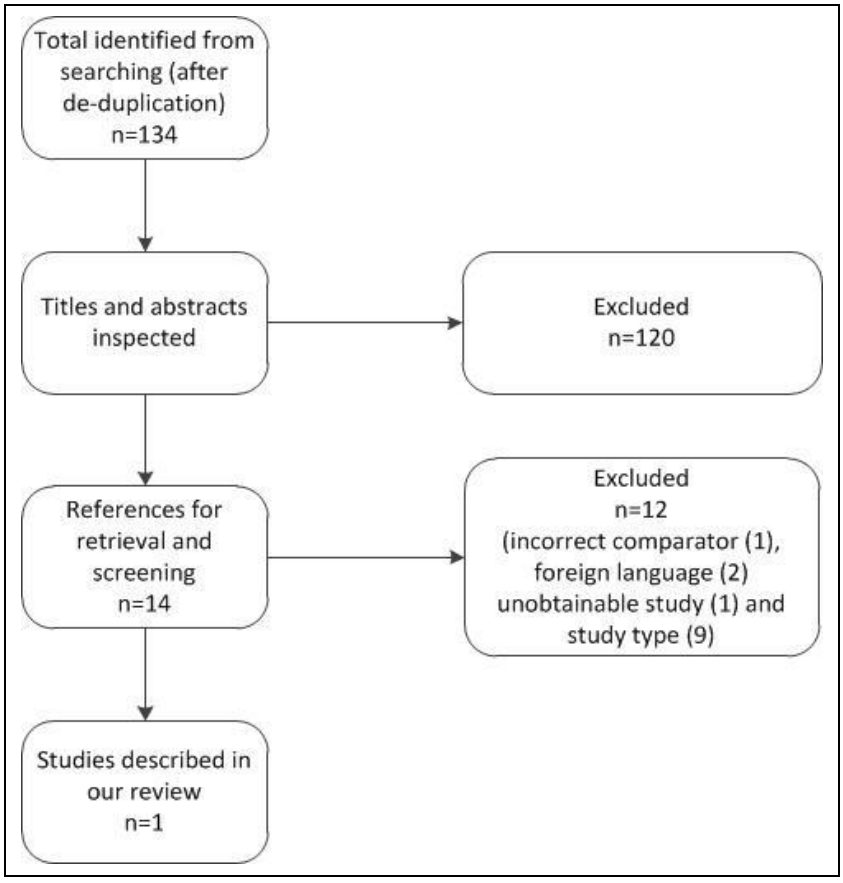
5.2 Results of systematic review of cost-effectiveness

The aim of this section is to assess the current state of evidence on the cost effectiveness of depth of anaesthesia monitoring compared to standard clinical monitoring through a systematic review of the literature. The methods used for the search strategy are described in section 4.2.1, and inclusion criteria are shown in section 4.2.2. 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.

5.2.1 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[98] 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.

Figure 5 Flow chart of identification of studies for inclusion in the review of cost effectiveness



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 studies were simple calculation models of BIS monitoring compared with standard treatment. The completed checklist for quality assessment of the included studies is shown in Table 29.

Table 29 - Critical appraisal checklist of economic evaluation (Questions in this checklist based on Philips and colleagues)

	Item	Abenstein[98]
1	Is there a clear statement of the decision problem?	Y
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?	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 QALYs?	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
Yes / No / ? (unclear)		

5.2.2 Characteristics and results of included studies

The included study was a simple calculation models 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 United Kingdom (UK) National Health Service (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 or investigate uncertainty through sensitivity analyses.

Table 30 - Characteristics of included economic evaluations

Author	**Abenstein [98]**
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 $11,294 for all patients.

Abenstein[98] used a simple calculation model to compare general anaesthesia (GA) with BIS monitoring to GA for high risk and general risk patients. The cost per avoided intraoperative recall (IR) is:

$$\frac{Cost\ per\ patient\ of\ BIS}{Incidence\ BIS - Incidence\ GA}$$

The cost per patient of BIS monitoring consisted of the cost of the sensors ($17 each) and the cost of the monitor. The monitor was assumed to cost $9000 and have a lifespan of seven years, and be used by four patients per day for 300 days per year ($1.07 per patient). The incidence for IR for patients of general risk was taken from a prospective study by Ekman and colleagues[99] who reported a recall rate of 0.04% (GA with BIS) compared to 0.18% (GA). The cost per avoided IR was $11,294. Abenstein[98] estimated the cost per avoided IR for high risk patients to be $4,410 per avoided IR. They used

estimates of the incidence of IR by averaging the difference between the Myles and colleagues[79] and Avidan and colleagues[27] studies which gave a reduction in incidence of IR 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.

5.2.3 Summary

One cost-effectiveness analysis[98] 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 due to the poor methodological and reporting quality.

5.3 Model structure, model parameterisation and results of economic evaluation

5.3.1 Description of decision analytic model

Overview

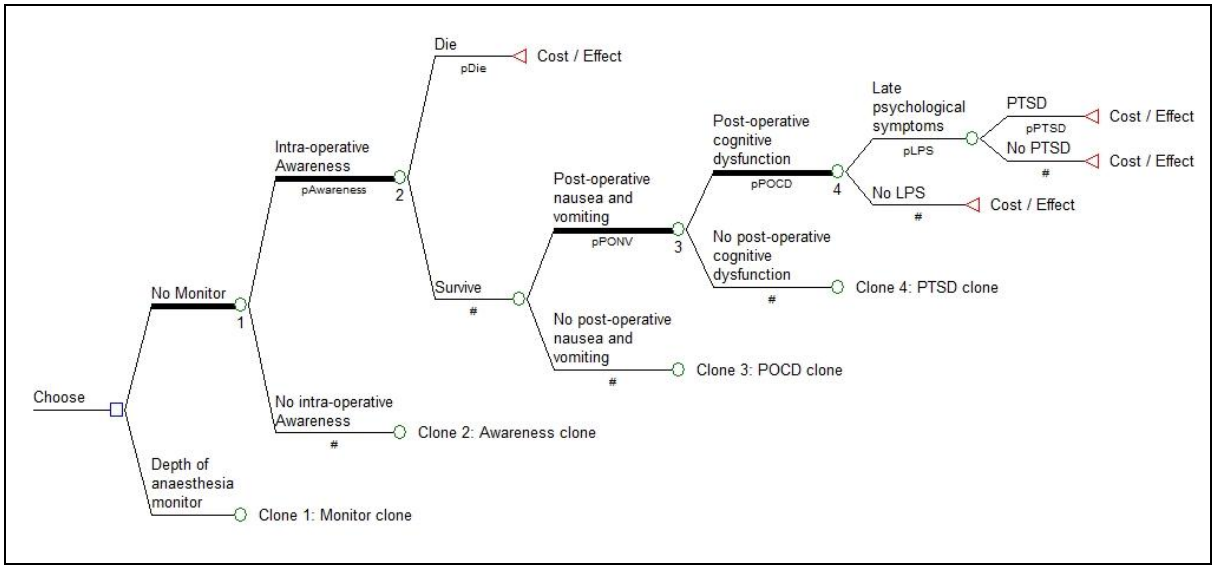
A decision analytic model was developed to assess the cost effectiveness of depth of anaesthesia monitoring, compared with standard clinical monitoring. Separate analyses are presented for each of the three included technologies (the included technologies are not compared with each other).

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 section 5.1). 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 quality adjusted life years (QALYs). 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 post-operative nausea and vomiting (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.

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



Each of the short term anaesthetic-related complications could be associated with additional treatment costs (such as anti-emetic medication for patients experiencing PONV, while for patients experiencing post-operative cognitive dysfunction (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 post-traumatic stress disorder (PTSD) and may seek treatment.

We assumed that monitoring of basic clinical signs, including blood pressure and heart rate, would be common components to standard clinical monitoring and to depth of anaesthesia monitoring using EEG devices (as discussed in section 3.2 of this report) and have therefore not been costed in the model. The key cost component 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 effect of EEG-based depth of anaesthesia monitoring (using the included technologies) compared with 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 odds ratios/ relative risks.

No quality of life impact (utility loss) is included in the model for short term anaesthesia-related complications (such as post-operative nausea and vomiting) as these are expected to be of limited duration. Similarly the model does not include an estimate of the quality of life impact (utility loss) for an intraoperative awareness episode. The most significant quality of life 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.

5.3.2 Model parameters

Cost of depth of anaesthesia monitoring

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

Table 31 - Costs of DoA modules

Depth of anaesthesia model	Manufacturer	Cost of depth of anaesthesia monitor (£)	Sensor cost, per patient (£)
E-entropy module	GE Healthcare	5,352	8.68[a]
Vista module (BIS)	Covidien	4,687.50[b]	17.75[c]
Compact M monitor	Narcotrend	8,572 – 11,998[d]	0.56[e]

Notes
[a] based on manufacturer's price of £217 for box of 25 sensors (1 sensor per patient)
[b] average across manufacturer's price of £4,350 (BIS Vista) and £5,025 (BIS Vista bilateral)
[c] average across manufacturer's price of £14.50 per patient (£362.50 for box of 25 sensors, 1 sensor per patient for Vista module) £21 per patient (£210 for box of 10 sensors, 1 sensor per patient for Vista bilateral module)
[d] range of prices quoted, dependent on model
[e] based on manufacturer's price of £0.14 per sensor (3 required for 1-channel recording and 5 required for 2-channel recording)

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

Table 32 - Equivalent annual costs of DoA modules

DoA module	Equivalent annual cost (£)
E-entropy module	1,185
BIS module[a]	1,038
Narcotrend monitor[b]	2,278
[a] based on an average cost across the Vista and Vista bilateral models of £4,687.50 [b] based on the mid-point of the range quoted by the manufacturer, £10,285	

The annual throughput of patients for each module is assumed to be 1,000 patients per year (equivalent to 4 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 2 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.

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 manufacturer of the E-entropy model state 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 time to other anaesthesia monitors (makes and coverage not specified)), while BIS and Entropy are modules designed to operate with other anaesthesia monitors. BIS is compatible with a range of monitoring platforms.

[redacted] E-entropy is compatible with GE Healthcare's most recent monitor range (CARESCAPE Monitors B850 and B650), but not older software levels (in HE 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 DoA 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 DoA modules

Unit costs for DoA modules include acquisition costs for the module (annualised, assuming a five 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[100] 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 paedriatric cases (general and 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), IV/ inhalational anaesthesia (propofol induction, isoflurane/N_20 maintenance or propofol induction, sevoflurane/N_20 maintenance) and total inhalational anaesthesia (sevoflurane induction, sevoflurane/N_20 maintenance). A total of 1,063 adult patients remained in the study until hospital discharge (265 propofol/propofol, 267 propofol/isoflurane, 280 propofol/sevoflurance, 251 sevoflurane/sevoflurane). The mean total and variable costs reported for the RCT are shown in Table 33.

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

	Propofol / propofol	Propofol / isoflurane	Propofol / sevoflurane	Sevoflurane / sevoflurane	Total
Mean total cost (£)	131.7	118.7	123.4	131.3	126.1
Mean variable cost (£)	21.1	7.1	13.8	15.3	14.4

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 depth of anaesthesia monitoring.

Baseline consumption of inhaled anaesthetic agents in the economic model was estimated using an equation reported by Chernin,[101] based on a formula originally presented by Dion.[102]

Cost per MAC unit time = (Concentration × FGF × duration × MW × cost/ml)/(2412 × D)
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 litres per minute for maintenance of anaesthesia.

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	litres/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 IV 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/hr or μg/kg/hr 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 Section 5.1.

Table 35 - Change in anaesthetic usage associated with depth of anaesthesia monitoring – mean difference and 95% confidence interval

Technology	Anaesthetic agent	Population	N Trials	Mean difference	Proportionate change
BIS vs standard clinical monitoring	Sevoflurane	General surgical	9	-0.15 (-0.25 to -0.06)	-0.202[a] (-0.330 to -0.074)
	Propofol	General surgical	11	-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 Remifentanil	Orthopaedic surgery	1[62]	5[d] (-7.54 to 17.54) -0.01[e] (-0.02 to 0.00)	0.050 (-0.075 to 0.174) -0.111 (-0.232 to 0.010)
	Propofol Remifentanil	Elective gynaecological laparoscopy	1[55]	-14[f] (-22.47 to -5.53) 0.07[g] (0.03 to 0.11)	-0.147 (-0.237 to -0.058) 0.179 (0.085 to 0.274)
Narcotrend vs standard clinical monitoring	Desflurane Remifentanil	Orthopaedic surgery	1[64]	-69[h] (-113.37 to -24.63) -0.01[i] (-0.04 to 0.02)	-0.156 (-0.256 to -0.056) -0.043 (-0.168 to 0.081)
	Propofol Remifentanil	Minor orthopaedic surgery	2[60;63]	-1.99[j] (-2.922 to -1.06) -0.01[k] (-0.04 to 0.01)	-0.292 (-0.429 to -0.155) -0.054 (-0.158 to 0.050)

Technology	Anaesthetic agent	Population	N Trials	Mean difference	Proportionate change
Notes [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/hr) in standard monitoring arm (meta-analysis weights). [c] mean difference in patient normalised consumption (g/kg/hr). Mean normalised consumption in standard monitoring arm of trial was 0.14 g/kg/hr [d] mean difference in patient normalised consumption (μg/kg/min). Mean normalised consumption in standard monitoring arm of trial was 101 μg/kg/min. [e] mean difference in patient normalised consumption (μg/kg/min). Mean normalised consumption in standard monitoring arm of trial was 0.09 μg/kg/min. [f] mean difference in patient normalised consumption (μg/kg/hr). Mean normalised consumption in standard monitoring arm of trial was 95 μg/kg/min. [g] mean difference in patient normalised consumption (μg/kg/min). Mean normalised consumption in standard monitoring arm of trial was 0.39 μg/kg/min. [h] mean difference in patient normalised consumption (mg/kg/min). Mean normalised consumption in standard monitoring arm of trial was 443.60 mg/kg/min. [i] mean difference in patient normalised consumption (μg/kg/min). Mean normalised consumption in standard monitoring arm of trial was 0.23 μg/kg/min. [j] mean difference in patient normalised consumption (mg/kg/hr), pooled across two trials (see Appendix 8). Mean normalised consumption pooled across the standard monitoring arms of the trials was 6.81 mg/kg/hr. [k] mean difference in patient normalised consumption (μg/kg/min), pooled across two trials (see Appendix 8). Mean normalised consumption pooled across the standard monitoring arms of the trial was 0.25 μg/kg/min.					

Consumption of anaesthetic drugs used in TIVA, for the comparison of Entropy and standard clinical monitoring is based on data reported in two clinical trials[55;62] which 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, no 62, September 2011[33]). Unit costs for volatile inhaled anaesthetic gases are not available in BNF. As a result these 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 – Unit costs of general anaesthetics

Anaesthetic agent	Unit	Cost (£)	Cost (£) / ml
Isoflurane	250ml bottle	57.00[a]	0.228
Desflurane	250ml bottle	76.00[a]	0.304
Sevoflurane	250ml bottle	148.00[a]	0.592
Propofol (1% injection, 10 mg/ mL)	50 mL bottle	10.10[b]	0.202
Source [a] University Hospital Southampton NHS Foundation Trust [b] BNF, No 62, September 2011[33]			

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 depth of anaesthesia monitoring, based on assumptions presented in Table 34, Table 35 and Table 36.

Table 37 - Estimated baseline cost, estimated change in consumption and cost of anaesthetic associated with depth of anaesthesia 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
	Ellerkman 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

Assumptions:
[a] anaesthetic duration of 1.6 hours
[b] normalised consumption of 6.73 mg/kg/h, 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/h, patient weight of 77 kg, anaesthetic duration of 2 hours[62]
[e] normalised consumption of 0.005 mg/kg/h, patient weight of 77 kg, anaesthetic duration of 2 hours[62]
[f] normalised consumption of 5.70 mg/kg/h, patient weight of 68 kg, anaesthetic duration of 1.8 hours[55]
[g] normalised consumption of 0.023 mg/kg/h, 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/h, patient weight of 79 kg, anaesthetic duration of 2.1 hours[64]

[j] normalised consumption of 6.81 mg/kg/h, patient weight of 79 kg, anaesthetic duration of 1.8 hours[60;63]
[k] normalised consumption of 0.015 mg/kg/h, patient weight of 79 kg, anaesthetic duration of 1.8 hours[60;63]

Post-operative nausea and vomiting

Our systematic review of patient outcomes identified limited evidence of the impact of depth of anaesthesia monitoring on the risk of PONV. A baseline risk of PONV (30%)[103-105] for standard clinical monitoring and depth of anaesthesia 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[106]. 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 intravenous injection (unit cost = £5.39, BNF no 62, September 2011[33]).

Post-operative 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 (over 60) available as an abstract, reported a reduction in POCD for BIS monitored patients at seven days and 3 months, although the difference at seven days was reported to not be statistically significant. There is disagreement over the true incidence of POCD with some authors arguing this may be under-estimated due to loss to follow up for the most severe cases,[107] while others argue that it may be over-estimated 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.[108] This study recruited people over the age of 60 who were presenting for major abdominal, non-cardiac thoracic or orthopaedic surgery under general anesthesia. Subjects with mini-mental state examination (MMSE) score less than 23 at baseline were excluded. Incidence of post-operative cognitive dysfunction at one week after surgery was 25.8% and was present in 9.9% of subjects at 3 months. This compared with 3.4% at one week and 2.8% at 3 months in non-surgical controls. Longer term follow up of subects in the ISPOCD study,[109] between one and two 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 one week and 7.1% at 3 months) represents cognitive dysfunction attributable to undergoing general anaesthesia, which will then gradually reduce to zero (at 18 months). Using these proportions (22.4% at one 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. We estimated the proportion of surgical patients experiencing POCD using data on the proportion of patients undergoing any procedure available from HES online,[110] which reported that 45% of patients were age 60 and above.

Change in post-operative cognitive dysfunction associated with depth of anaesthesia monitoring

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

Table 38 - Estimated odds ratios for POCD at seven days and three months estimated from Chan et al

	Routine care (n = 452)	BIS guided anaesthesia (n = 449)	Estimated odds ratio
POCD			
1 week	39.1%	32.5%	0.750
3 months	12.0%	8.1%	0.646

The odds ratios were applied to the baseline proportions with cognitive dysfunction at seven days and three months and mean duration of POCD associated with BIS monitoring was estimated at 21.10 days.

QoL impact of post-operative cognitive dysfunction associated with depth of anaesthesia monitoring

The quality of life impact of POCD was based on the utility decrement reported by Jonsson and colleagues[111] for the difference between a mini-mental state evaluation score greater than 25 (no dysfunction) which had a utility of 0.69 and a mini-mental state evaluation score between 21 and 25 (indicating mild cognitive impairment) which had a utility of 0.64.

Intraoperative awareness

Baseline value

Awareness (defined as post-operative recollection of events occurring during general anaesthesia) has generally been described as a rare occurrence, with an incidence of 0.1-0.2% in the general surgical

population. While 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;112;113]

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 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%). While 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 Q = 212.55 (df=5), p=0.0000, 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 confidence interval (0.16% [95% CI 0.10% to 0.23%] assuming random effects [Cochran 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 (section 5.1). Pooling these estimates gives a cumulative incidence of awareness of 0.45% (95% CI 0.06% to 1.19%) assuming random effects (Cochran 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.

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 (1991)[114]	One centre, prospective (2/1990 – 4/1990)	Patients aged 16 years or older undergoing GA[a]	1,000	Brice interview	Single interview between 20 and 36 hours after surgery	2 (0.2%)
Ranta et al (1998)[6]	One centre prospective (8/1994 – 8/1995)	Patients aged 12 years or older undergoing GA	2,612[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 program, 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 program, 8 centres (2002-2004)	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	4,001	Structured interview – does not appear include Brice questions	Three times: In PACU; 7 days and 30 days later	39 (0.99%)[f]
Lyons et al[112]	(1982 – 1989)	Patients undergoing GA for cesarian-section	3,000	Unclear	Unclear	8 (0.93%)
Ranta et al (1996)[113]		Cardiac surgery patients	204	Unclear	Unclear	3 (1.5%)
Puri et al [82g]	Multicentre RCT	High risk patients	16[h]	Not reported	Not reported	1 (6.25%)

Study	Study design (dates)	Population description	Sample size	Measure of awareness	Timing and frequency of measure	Incidence of awareness n (%)
Myles et al[79]	Multicentre RCT	High risk patients	1,238[i]	Structured questionnaire, not defined.	Three times: 2-6 hours, 24-36 hours and 30 days after surgery	11 (0.89%)
Avidan et al (2008)[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%)
Muralidhar et al[78g]		High risk	20[k]	Not reported	Not reported	0 (0.00%)
Avidan et al(2011)[44]	Multicentre RCT	High risk patients - at least one risk factor	2,852[l]	Modified Brice interview. Michigan awareness classification for assigning to possible or definite awareness	Twice: 72 h and 30 days after extubation	2 (0.07%)[m]

[a] excluded patients undergoing obstetric or intracranial surgery
[b] captured 54% (2612/ 4818 eligible cases)
[c] reported 10 definite and 9 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 3,921 (no explanation why this is lower than stated sample size of 4,001). If "high risk" patients (emergency surgery, intraoperative hypotension-shock and cesarian-section) were excluded the incidence reduced to 0.8% (28/3477). At the seven-day interview 6 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 6,041
[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)

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 Section 5.1. There are no entries for 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 odds ratios derived for BIS were used in the model to estimate the impact on intraoperative awareness of depth of anaesthesia monitoring with Entropy and Narcotrend.

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 odds ratios 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.

Table 40 - Effectiveness of depth of anaesthesia monitoring on risk of awareness – Peto Odds Ratio and 95% confidence interval from systematic review of patient outcomes

Model of General Anaesthetic	Population	N Trials	Peto Odds Ratio	95% CI
Mixed anaesthesia (includes both patients undergoing TIVA and patients undergoing inhaled general anaesthesia)[a]	High risk	1	0.25	0.08 to 0.75
Inhaled general anaesthesia 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%) while others had inhaled anaesthetics (with or without IV anaesthetic)				

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;115;116] see 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 [116] in patients with an LPS duration of less than 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 one month. [8]

Table 41 - Studies reporting incidence of late psychological symptoms and PTSD in patients who experienced awareness - summary of characteristics, methods and results

Study	Date	Method of recruitment	Identification/ classification of LPS & PTSD	Aware	LPS		PTSD	
				n	n	%	n	%
Evans[117]	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[115]	1998	Advertisements in four German papers and on internet (n= 21) or referral from 3 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 Anesthesiologists 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%

Study	Date	Method of recruitment	Identification/ classification of LPS & PTSD	Aware	LPS		PTSD	
				n	n	%	n	%
Osterman et al[10]	2001	Advertisment in newspapers, fliers in hospitals, self-referral following print and TV news stories or referral by anesthesiologist	PTSD defined using Clinician Administered PTSD scale (CAPS)	16	NR		9	56%
Lennmarken et al[11]	2002	18 patients identified as experiencing awareness during general anaesthesia 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 Diagnostic and statistical manual of mental disorders, 4th Edition (DSM-IV), American Psychiatry Association[118]	9	7	78%	4	44%
Samuelsson et al[116]	2007	Consecutive patients who had undergone general anaesthesia were interviewed regarding awareness during **previous** general anaesthesia	Late psychological symptoms 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%

Study	Date	Method of recruitment	Identification/ classification of LPS & PTSD	Aware	LPS		PTSD	
				n	n	%	n	%
Ghoneim et al[7]	2009	Data extracted from published case reports on "awareness" and "anesthesia" - from PubMed between 1950 and August 2005	No definition of late psychological symptoms	271	NR	22%	NR	
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 Clinician Administered PTSD scale (CAPS)	7	NR		5	71%

Notes
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, one had died.
[c] Samuelsson identified 8 (17%) patients as having a total symptom score (summed across seven symptoms) greater than 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 Lennmarken and colleagues[11] and Samuelsson and colleagues[116] estimate for severe psychological sequelae (n = 8, 17%) to derive an incidence of 26% (95% CI 15% to 37%)

Just two of the studies had a prospective design.[8;116] The study by Samuelsson and colleagues[116] 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 on-going, 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, due to retrospective design, participant recruitment methods or low recruitment levels.[5;7;10;11;19;115]

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

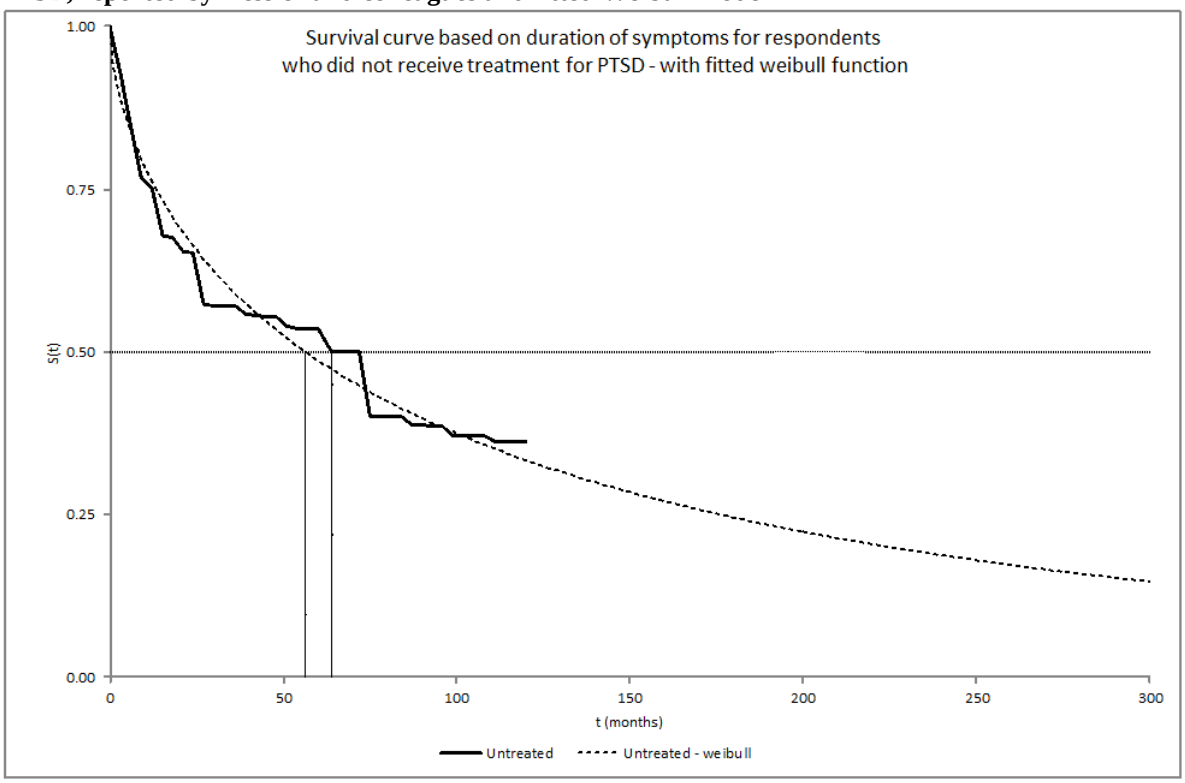
	Value	Method	Source
Late psychological symptoms			
Probability, given awareness	0.326 (0.195 -0.480)	15/ 46 patients with awareness	Samuelsson et al[116]
PTSD			
Probability, given awareness	0.177 (0.113 -0.230)	Pooled proportion of subjects with LPS having PTSD or severe symptoms, from (2) studies reporting this proportion, applied to probability of LPS Pooled estimate based on 0.571[11] (4/7) and 0.0533[116] (8/15) = 0.542 (95% CI 0.345-0.733) Probability PTSD = (15/46) * 0.542	Samuelsson et al[116] and Lennmarken et al[11]

Duration of PTSD

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, since these cases of PTSD reported were on-going, 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 non-institutionalised, civilian population aged 15-54 years, conducted in the United States.[119] These data were from the National Comorbidity Survey (a survey designed to study the distribution, correlates and consequences of psychiatric disorder in the United States) 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), see Figure 7.

Figure 7 - Survival curve based on duration of symptoms for respondents who did not seek treatment for PTSD, reported by Kessler and colleagues and fitted Weibull model



QoL impact of psychological sequelae

A review of the health-related quality of life (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.

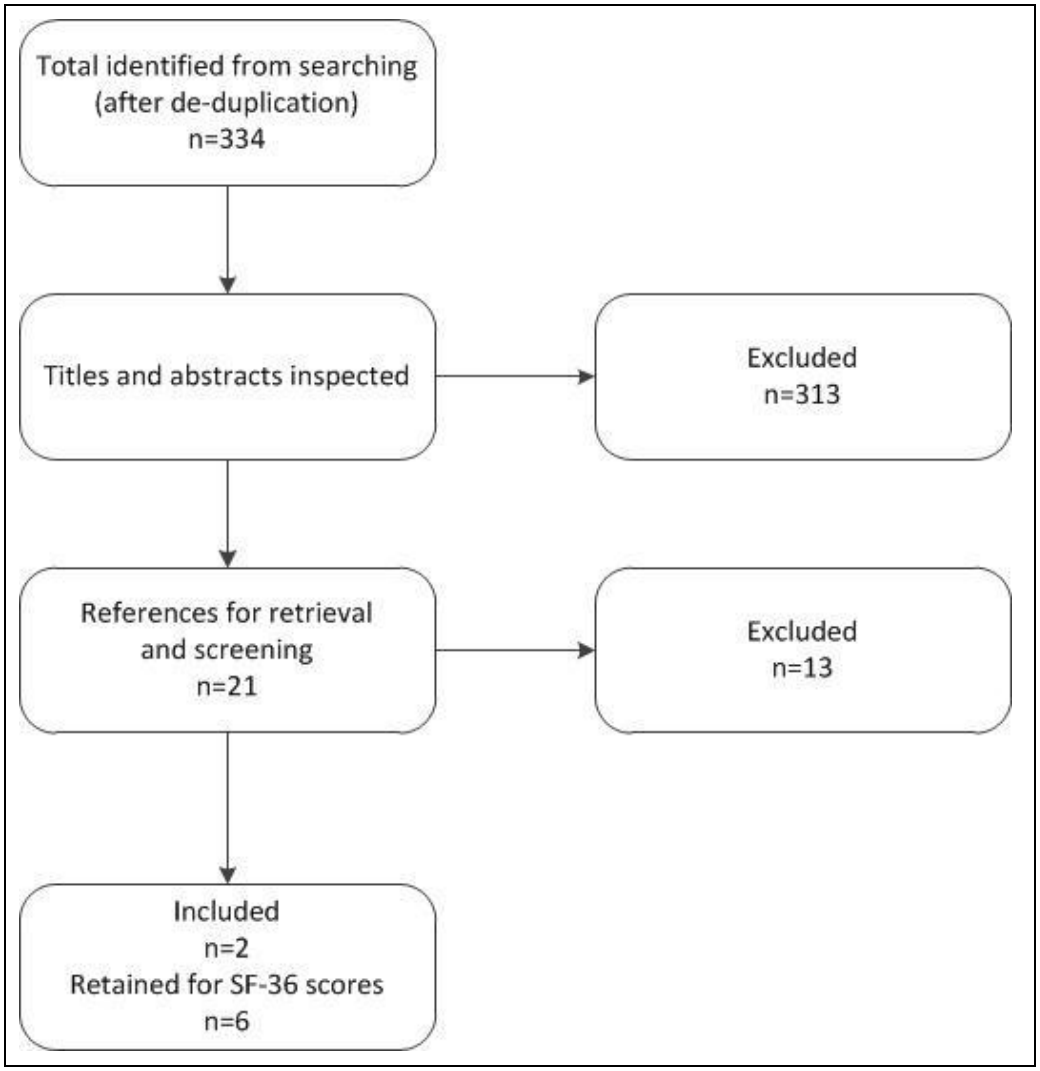
Methods

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. These were assessed against the inclusion criteria detailed in Table **43**.

Table 43 - Inclusion criteria for quality of life review

	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 quality of life, such as EQ-5D, SF6D, or other standard valuation technique such as standard gamble or time trade-off	
Interventions:	Any	
Other:	Articles published in English	Articles in languages other than English Conference abstracts

Figure 8 - Flow chart of identification of QoL studies for inclusion in the review



Characteristics of the included studies

Two papers[120;121] met the inclusion criteria for the review. The study design and population baseline characteristics are shown in Table 44 below.

Table 44 - Characteristics of included QoL studies

	Doctor et al[120]	**Freed et al[121]**
Patient Group	Patients with PTSD	Veterans with PTSD
Country and setting	US, multi-centre 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	Not reported
Age (mean ± SD)	37.31 ± 11.33	60 ± 12
Sex (F)	141 (76%)	176 (21%)

QoL instrument	SG/TTO/VAS[a]	SF-36[b]
[a]SG, standard gamble; TTO, time trade-off; VAS, visual analogue scale; [b]SF-36 scores transformed to utility scores using Brazier et al[122]		

The two included studies 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.[123-125] Freed and colleagues interviewed veterans with PTSD,[121] whereas Doctor and colleagues[126] interviewed a sample of patients taking part in a trial of treatments for chronic PTSD at baseline.

The two studies were considerably different in size with the Freed study[121] 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,[121] female patients comprised 21% of the sample, and the average age was 60 years. In the Doctor and colleagues study the sample was on average younger, with a mean age of 37, and 76% of the respondents were female.

In addition, the two studies generated the results using different valuation tools and methods. Neither of the included studies was based on the EQ-5D, as prescribed by the NICE reference case.[38] Doctor and colleagues[120] asked respondents to respond using standard gamble (SG), visual analogue scale (VAS), and time-trade off (TTO) techniques, the latter of which is recommended as an alternative. [38] Freed and colleagues[127] used the SF-36 responses from a previous study[128] and converted these to preference weighted health scores (PWHS) using the formula developed by Brazier and colleagues.[129]

Both studies included the results of statistical models generated in order to identify predictors for worsening or improvement of utility scores.[120;121]

Quality of the included studies

Doctor and colleagues[120] 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[121] 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 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 below.

Table 45 - Utility scores reported in the included QoL studies

	Doctor *et al*[120]			Freed *et al*[121]
	SG	**TTO**	**VAS**	
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]

SG, standard gamble; TTO, time trade-off; VAS, visual analogue scale

[a] Transformed from SF-36 using Brazier et al[122]

The scores for veterans in the Freed study 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[120] report three separate scores according to the valuation method. The scores for TTO and VAS are similar, (0.66 and 0.64 respectively,) while the score for SG appears high at 0.87. The authors argue that the mixed effects 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.[120] 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 SF-36 scores

A further six studies 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.[130] 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[121]. These were converted using the algorithm published by Ara and Brazier,[130] and are reported in Table 46.

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

Study	Patient Group	Utility		
		PTSD	No PTSD	Difference
Laffaye et al[123]	Women experiencing domestic abuse	0.634	0.748	0.114
Meeske et al[125]	Young adult survivors of childhood cancer	0.666	0.799	0.132
Berger et al[131]	Male ambulance workers	0.705	0.790	0.085
Shiner et al[132]	Veterans	0.508	-	-
Tsai et al[124]	Earthquake survivors (0.5 yrs post)[a]	0.649	0.783	0.134
Evren et al[133]	Alcohol dependent men with history of emotional abuse	0.592	0.659	0.068
[a] 3 year post earthquake and delayed PTSD and recovery scores also reported in Tsai et al[124]				

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 SF-6D[134] be employed in the 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 a similar difference in HRQoL between groups of patients who have similar experiences who go on to develop PTSD, and those who do not, and are consistent with the differences reported by Freed and colleagues.[121] 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,[121] with the exception of those reported by Shiner and colleagues [132] also elicited from veterans.

Summary

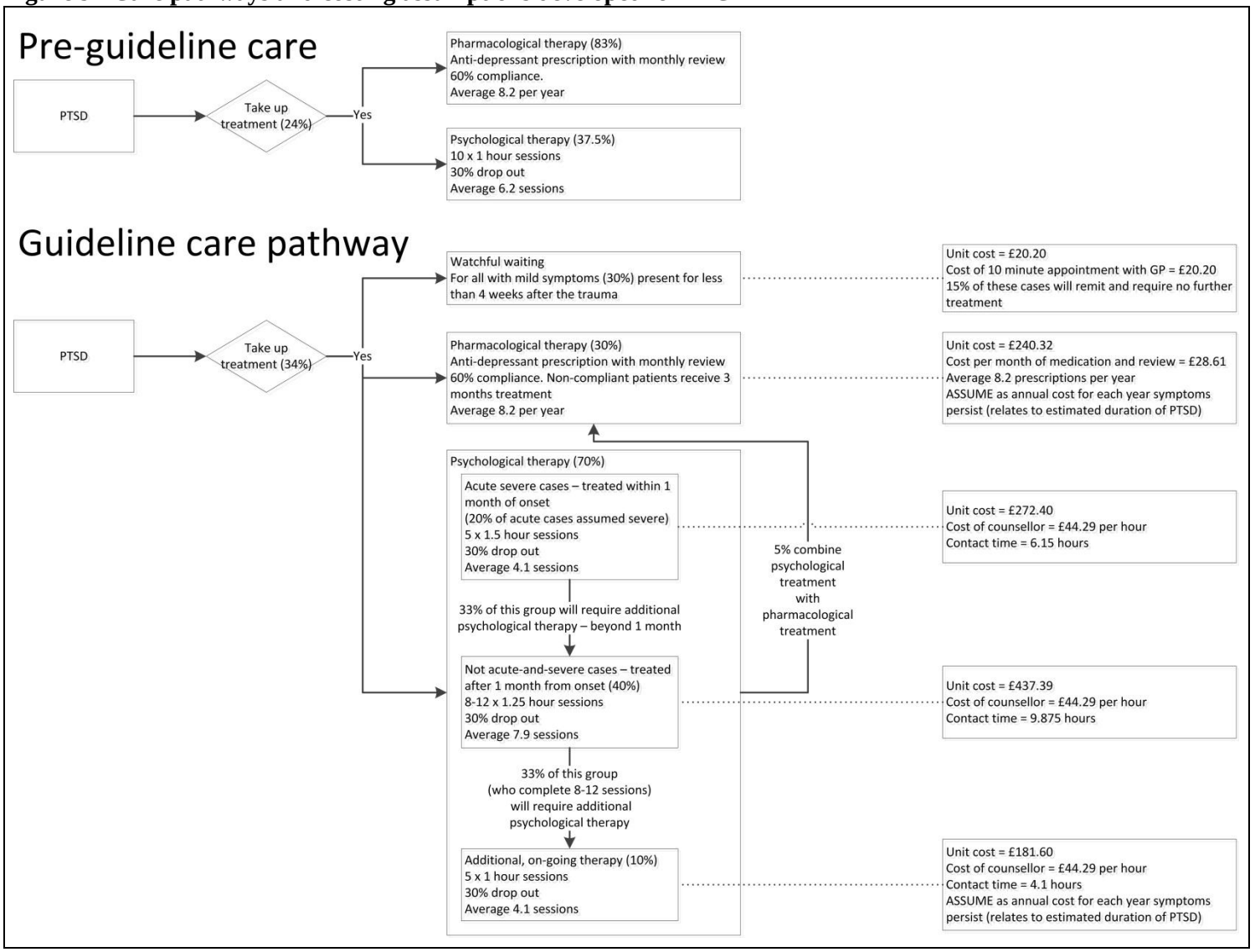
- Two papers met the inclusion criteria for this review of utility scores in PTSD. Six other papers reporting SF-36 scores for people with PTSD were also retained.
- Neither of these studies meeting the inclusion criteria (reporting a utility value based on a generic, preference-based measure) was based on the EQ-5D.

- One study reported a utility score for patients with PTSD based on TTO[120]. 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.[121]
- 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.

PTSD costs

The costs of treating PTSD have been estimated based on assumptions contained in the national cost impact report[135] associated with NICE Clinical Guideline no 26 on the management of PTSD in adults and children in primary and secondary care.[136] 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.

Figure 9 - Care pathways and costing assumptions developed for NICE



Data from the adult psychiatric morbidity survey,[137] 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 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, while 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-focussed 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, while 30% initially accept pharmacological treatment) and severity (30% patients have mild symptoms and are initially managed through watchful waiting, while 20% have severe symptoms and are offered trauma-focussed 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.

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 & 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, on-going psychological therapy	181.60	9.97%	229.86
Total			1528.26

The NICE Guideline 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 healthcare utilisation, derived from electronic medical records, for civilian primary care patients, including a proportion who had current PTSD.[138] This study reported an incidence rate ratio of 2.22 (adjusted for age, gender, income, substance dependence,

depression and comorbidity) for hospitalisation in subjects with PTSD compared to 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 (£2,590), based on 2010/2011 NHS Reference Costs[139] 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 £7,576 for hospitalisations among people with PTSD over the average duration of symptoms of 12.7 years.

The total cost associated with PTSD was £9,104 (undiscounted) or £6,128 (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 only be realised 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 under-dosed.

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 with depth of anaesthesia monitoring and a list of assumptions underlying the estimation of the cost and outcomes associated with the psychological sequelae of intra-operative awareness.

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

Table 48 - Model input parameters. Cost per patient of depth of anaesthesia modules

Parameter	Value	Source
BIS		
Cost per patient of depth monitoring device	£1.04	Equivalent annual cost for depth monitor (acquisition cost £4,687.50) assuming an effective life of five years and using a discount rate of 3.5%. Patient throughput assumed at 1,000 per year
Cost per patient of depth monitor sensors	£17.25	Average across manufacturer's price of £362.50 for a box of 25 sensors (for Vista monitor) and £210 for a box of 10 sensors (for Vista bilateral monitor)
Entropy		
Cost per patient of depth depth monitoring device	£1.19	Equivalent annual cost for depth monitor (acquisition cost £5,352) assuming an effective life of five years and using a discount rate of 3.5%. Patient throughput assumed at 1,000 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 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 five years and using a discount rate of 3.5%. Patient throughput assumed at 1,000 per year
Cost per patient of depth monitor sensors	£0.56	Average across manufacturer's price of £0.14 per sensor, using 3 for 1-channel recording and 5 for 2-channel; recording

Table 49 - Model input parameters. Anaesthetic drug consumption

Parameter	Value	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 litres per 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 IV anaesthetic cost	£20.92	Cost for 2 hours of propofol (at 6.77 mg/kg/hr [from control arms of RCTs in meta analysis] and patient average weight of 77kg). Unit cost of £0.0202 per mg.
Reduction in consumption of IV 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/hr
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 litres per 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/hr (in standard care arm, Aime et al[61]
Baseline IV anaesthetic cost	Propofol = £18.85 Remifentanil = £ 4.26	Ellerkman et al[62]
	Propofol = £14.35 Remifentanil = £14.94	Gruenewald et al[55]

Parameter	Value	Source
Reduction in consumption of IV anaesthetic using depth monitor	0.050 (-0.075 to 0.174) -0.111 (-0.232 to 0.010)	Propofol mean difference of 5 from baseline rate of 101 mg/kg/hr (Ellerkman et al)[62] Remifentanil mean difference of -0.01 from baseline rate of 0.09 mg/kg/hr (Ellerkman et al) [62]
	-0.147 (-0.237 to -0.058) 0.179 (0.085 to 0.274)	Propofol mean difference of -14 from baseline rate of 95 mg/kg/hr (Gruenewald et al)[55] Remifentanil mean difference of 0.07 from baseline rate of 0.39 mg/kg/hr (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 litres per 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/min from 443.6 mg/min (in standard care arm, Kreuer et al)[64]
Baseline IV anaesthetic cost	Propofol = £19.39 Remefentanil = £10.79	Cost for 108 minutes of propofol (at 6.81 mg/kg/hr [from control arms of RCTs] and patient average weight of 80kg). Unit cost of £0.0202 per mg. Cost for 108 minutes of remifentanil (at 0.120 mg/kg/hr [from control arms of RCTs] and patient average weight of 80kg). Unit cost of £5.12 per mg.
Reduction in consumption of IV anaesthetic using depth monitor	-0.292 (-0.429 to -0.155) -0.054 (-0.158 to 0.050)	Propofol mean difference of -1.99 from baseline rate of 6.8 mg/kg/hr[60;63] Remifentanil mean difference of -0.01 from baseline rate of 0.25 mg/kg/hr[60;63]

Table 50 - Model input parameters. Intra-operative awareness

Parameter	Value	Source
Intra-operative awareness		
Baseline awareness in surgical population at high risk of awareness	0.45% (0.06% to 1.19%)	Pooled estimate from control arms of RCTs in high risk patients
Reduction in awareness using depth monitor High risk patients undergoing TIVA (Peto odds ratio) High risk patients undergoing anaesthetic induction with IV and maintenance with inhaled anaesthetic (Peto odds ratio)	 0.24 (0.33 to 1.48) 0.45 (0.25 to 0.81)	Meta analysis of RCTs in high risk patients, undertaken as part of this review (see section 5.1.3)
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 odds ratio) General surgical population undergoing anaesthetic induction with IV and maintenance with inhaled anaesthetic (Peto odds ratio)	 0.24 (0.33 to 1.48) 0.45 (0.25 to 0.81)	Meta analysis of RCTs in high risk patients, undertaken as part of this review (see section 5.1.3). Effect assumed to be the same as for high risk patients
Psychological sequelae of intra-operative awareness		
Probability of LPS, given awareness	0.326 (0.195 to 0.480)	Samuelsson et al[116]
Duration of LPS	Six 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[116] and Lennmarken et al[11]
Duration of PTSD	12.7 years	Kessler et al[119]
Unit cost of PTSD	£9,104	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[138]. Average cost of in-patient stay. NHS Reference Costs 2010/11[139]
Utility reduction due to PTSD	0.12	Various

Table 51 - Model input parameters. Post-operative complication (PONV and POCD)

Parameter	Value	Source
Post-operative nausea and vomiting		
Baseline PONV	30%	Cohen M.M., Duncan P.G, DeBoer D.P., Tweed W.A. The postoperative interview: assessing risk factors for nausea and vomiting. Anesthetics and Analgesia 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 (4mg of ondanisetron)
Utility reduction due to PONV	0	
Post-operative cognitive dysfunction		
Baseline POCD	average duration of 29.65 days	ISPOCD study reported POCD in 25.8% (95% CI 23.1 to 28.5) patients at 1 week and in 9.9% (95% CI 8.1 to 12.0) 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 (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. In subjects Odds ratios 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[111] difference in utility between MMSE greater than 25 (0.69) and MMSE between 21 and 25 (0.64). Normal to mild cognitive dysfunction.

5.3.3 Model Results

The model results are presented in separate sub-sections for BIS, Entropy and Narcotrend respectively. Analyses are presented by mode of administration (TIVA and mixed anaesthesia (induction with IV 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 general anaesthesia only. While 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.

BIS compared with Standard clinical monitoring

Base case

Total intravenous anaesthesia (TIVA)

The costs, QALYs and incremental cost effectiveness ratio (ICER) modelled for patients considered at high risk patients of intraoperative awareness undergoing general anaesthesia with TIVA, comparing standard clinical monitoring with BIS are presented in Table 52.

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Standard clinical monitoring	24.19		-0.0011		
BIS	42.76	18.57	-0.0005	0.0007	27,345

BIS 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 6 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 cost difference of £18.57. The increased cost for BIS monitoring is primarily the result of the sensors attached to the patient (90% 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 (see Table 53).

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	19.83
Anaesthetic drugs	20.92	20.92
Post-operative nausea and vomiting	1.62	1.62
Post-operative cognitive dysfunction	0.00	0.00
PTSD	1.66	0.40

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 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 general anaesthesia with TIVA, BIS monitoring would be cost effective (at a threshold of £30,000 per QALY gained) where the odds ratio for awareness (BIS vs standard clinical monitoring) was less than 0.315.

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

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Standard clinical monitoring	23.13		-0.0007		
BIS	37.43	14.30	-0.0004	0.0003	45,033

While the cost of standard clinical monitoring in this group of patients was slightly lower than for the sub-group of patients at high risk of intraoperative awareness, the incremental cost of BIS monitoring is lower. This is attributable to the potential to off-set a reduction in

consumption of anaesthetic against the additional costs of depth of anaesthesia monitoring (see 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 £45,033 per QALY gained.

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	18.79
Anaesthetic drugs	20.92	16.88
Post-operative nausea and vomiting	1.62	1.62
Post-operative cognitive dysfunction	0.00	0.00
PTSD	0.59	0.14

Mixed anaesthesia (induction with IV anaesthetic (propofol) and maintenance with inhaled anaesthetic (sevoflurane))

The costs, QALYs and ICER modelled for patients considered at high risk of intraoperative awareness undergoing general anaesthesia, comparing standard clinical monitoring with monitoring by BIS are presented in Table 56.

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Standard clinical monitoring	14.31		-0.0011		
BIS	33.23	18.92	-0.0006	0.0005	36,126

BIS 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 BIS monitoring during anaesthesia in high risk patients was higher than for standard clinical monitoring, with an incremental cost of £18.92. As discussed previously, the majority of the cost increase with BIS 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 (see Table 57).

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	19.83
Anaesthetic drugs	11.04	11.04
Post-operative nausea and vomiting	1.62	1.62
Post-operative cognitive dysfunction	0.00	0.00
PTSD	1.66	0.75

BIS 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 3 cases of late psychological symptoms (from 5.2 to 2.33) 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 mixed general anaesthesia, comparing standard clinical monitoring with monitoring by BIS are presented in Table 58.

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Standard clinical monitoring	13.25		-0.0007		
BIS	29.48	16.23	-0.0004	0.0003	61,869

Costs of standard clinical monitoring and BIS monitoring in this group of patients are both lower than for the sub-group of patients at high risk of intraoperative awareness. The cost difference is lower, due to the potential to off-set a reduction in consumption of anaesthetic against the additional costs of depth of anaesthesia monitoring (see 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 odds ratio of 0.45) compared with TIVA (Peto odds ratio of 0.24). The QALY gain of 0.0003 was lower than in the high risk group and results in an increased ICER of £61,869 per QALY gained.

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	18.79
Anaesthetic drugs	11.04	8.81
Post-operative nausea and vomiting	1.62	1.62
Post-operative cognitive dysfunction	0.00	0.00
PTSD	0.59	0.27

Deterministic sensitivity analysis

Total intravenous anaesthesia (TIVA)

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

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 (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
Probability awareness	0.0006	22.76	-0.0005	42.42	-0.0003	19.66	0.0002	101,932
	0.0119	26.92	-0.0024	43.42	-0.0008	16.51	0.0016	10,322
Odds ratio awareness with DoA monitor	0.1	24.19	-0.0011	42.53	-0.0004	18.34	0.0008	23,423
	0.6	24.19	-0.0011	43.36	-0.0007	19.17	0.0004	46,428
Duration of LPS (years)	0.25	24.19	-0.0011	42.76	-0.0005	18.57	0.0007	27,975
	1	24.19	-0.0012	42.76	-0.0005	18.57	0.0007	26,165
Probability of LPS[a]	0.195	23.53	-0.0125	42.60	-0.0121	19.07	0.0004	46,126
	0.48	24.98	-0.0302	42.95	-0.0293	17.97	0.0008	21,260
Duration of PTSD (yrs)	5.6	24.19	-0.0010	42.76	-0.0004	18.57	0.0006	33,496
	9.6	24.19	-0.0014	42.76	-0.0005	18.57	0.0008	21,996
Proportion PTSD[b]	0.345	23.59	-0.0009	42.62	-0.0004	19.03	0.0005	38,096
	0.733	24.78	-0.0014	42.90	-0.0005	18.13	0.0009	21,243
LPS QoL decrement	-0.075	24.19	-0.0011	42.76	-0.0005	18.57	0.0007	27,815
	-0.05	24.19	-0.0011	42.76	-0.0005	18.57	0.0007	28,083
PTSD QoL decrement	-0.134	24.19	-0.0012	42.76	-0.0005	18.57	0.0007	25,061
	-0.068	24.19	-0.0008	42.76	-0.0004	18.57	0.0004	41,338
Probability people with PTSD seek treatment	0	22.54	-0.0011	42.36	-0.0005	19.83	0.0007	29,197
	1	27.41	-0.0011	43.54	-0.0005	16.13	0.0007	23,749
Cost of sensors	13.3125	24.19	-0.0011	38.32	-0.0005	14.13	0.0007	20,810
	22.1875	24.19	-0.0011	47.20	-0.0005	23.01	0.0007	33,880

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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 changes in the probability of awareness in the patients at high risk of intra-operative awareness, receiving TIVA resulted in a substantially altered ICER from the base case: £10, 322 per QALY gained and £101, 932 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.

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 (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
Proportional change in propofol use	-0.272	23.13	-0.0007	35.78	-0.0004	12.65	0.0003	39,830
	-0.113	23.13	-0.0007	39.10	-0.0004	15.98	0.0003	50,303
Probability awareness	0.001	22.91	-0.0006	37.38	-0.0003	14.47	0.0002	59,608
	0.0023	23.38	-0.0008	37.49	-0.0004	14.11	0.0004	34,842
Odds ratio awareness with DoA monitor	0.1	23.13	-0.0007	37.35	-0.0003	14.22	0.0004	40,121
	0.6	23.13	-0.0007	37.64	-0.0004	14.52	0.0002	65,094
Duration of LPS (years)	0.25	23.13	-0.0007	37.43	-0.0004	14.30	0.0003	45,819
	1	23.13	-0.0007	37.43	-0.0004	14.30	0.0003	43,540
Probability of LPS[a]	0.195	22.89	-0.0122	37.37	-0.0120	14.48	0.0002	64,906
	0.48	23.40	-0.0295	37.50	-0.0292	14.09	0.0004	37,396
Duration of PTSD (yrs)	5.6	23.13	-0.0006	37.43	-0.0003	14.30	0.0003	52,346
	9.6	23.13	-0.0007	37.43	-0.0004	14.30	0.0004	38,004
Proportion PTSD[b]	0.345	22.91	-0.0006	37.38	-0.0003	14.47	0.0003	57,020
	0.733	23.33	-0.0008	37.48	-0.0004	14.15	0.0004	37,266
LPS QoL decrement	-0.075	23.13	-0.0007	37.43	-0.0004	14.30	0.0003	45,620
	-0.05	23.13	-0.0007	37.43	-0.0004	14.30	0.0003	45,953
PTSD QoL decrement	-0.134	23.13	-0.0007	37.43	-0.0004	14.30	0.0003	42,114
	-0.068	23.13	-0.0006	37.43	-0.0003	14.30	0.0002	60,652
Probability people with PTSD seek treatment	0	22.54	-0.0007	37.29	-0.0004	14.75	0.0003	46,443
	1	24.27	-0.0007	37.70	-0.0004	13.43	0.0003	42,298
Cost of sensors	13.3125	23.13	-0.0007	32.99	-0.0004	9.87	0.0003	31,062
	22.1875	23.13	-0.0007	41.87	-0.0004	18.74	0.0003	59,005

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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

These results suggest that the ICER for the general surgical population is relatively robust to changes in the duration of LPS, changes in the quality of life 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 quality of life decrement applied to PTSD.

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 general anaesthesia. The results are shown in Table 62 and Table 63.

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

Parameter	Input value	Standard clinical monitoring		BIS		Incremental		ICER (£/ QALY gained)
		Cost (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
Probability awareness	0.0006	12.88	-0.0005	32.58	-0.0003	19.70	0.0002	114,456
	0.0119	17.04	-0.0024	34.47	-0.0012	17.43	0.0012	14,682
Odds ratio awareness with DoA monitor	0.25	14.31	-0.0011	32.90	-0.0005	18.58	0.0007	27,671
	0.81	14.31	-0.0011	33.83	-0.0009	19.51	0.0003	75,664
Duration of LPS (years)	0.25	14.31	-0.0011	33.23	-0.0006	18.92	0.0005	36,906
	1	14.31	-0.0012	33.23	-0.0006	18.92	0.0005	34,661
Probability of LPS[a]	0.195	13.65	-0.0125	32.93	-0.0122	19.28	0.0003	58,139
	0.48	15.10	-0.0302	33.58	-0.0295	18.49	0.0006	28,709
Duration of PTSD (yrs)	5.6	14.31	-0.0010	33.23	-0.0006	18.92	0.0004	43,641
	9.6	14.31	-0.0014	33.23	-0.0007	18.92	0.0006	29,419
Proportion PTSD[b]	0.345	13.71	-0.0009	32.96	-0.0005	19.25	0.0004	48,882
	0.733	14.90	-0.0014	33.49	-0.0007	18.60	0.0006	28,630
LPS QoL decrement	-0.075	14.31	-0.0011	33.23	-0.0006	18.92	0.0005	36,708
	-0.05	14.31	-0.0011	33.23	-0.0006	18.92	0.0005	37,040
PTSD QoL decrement	-0.134	14.31	-0.0012	33.23	-0.0007	18.92	0.0006	33,283
	-0.068	14.31	-0.0008	33.23	-0.0005	18.92	0.0004	52,923
Probability people with PTSD seek treatment	0	12.66	-0.0011	32.48	-0.0006	19.83	0.0005	37,863
	1	17.53	-0.0011	34.68	-0.0006	17.15	0.0005	32,755
Cost of sensors	13.3125	14.31	-0.0011	28.79	-0.0006	14.48	0.0005	27,652
	22.1875	14.31	-0.0011	37.67	-0.0006	23.35	0.0005	44,601

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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 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 ICERs of £14,682 and £114, 456 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 £23, 423 to £58, 139 per QALY gained.

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

Parameter	Input value	Standard clinical care		BIS		Incremental		ICER (£/ QALY gained)
		Cost (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
Proportional change in sevoflurane use	-0.330	13.25	-0.0007	28.07	-0.0004	14.82	0.0003	56,483
	-0.074	13.25	-0.0007	30.89	-0.0004	17.65	0.0003	67,254
Probability awareness	0.001	13.03	-0.0006	29.38	-0.0004	16.36	0.0002	78,532
	0.0023	13.50	-0.0008	29.60	-0.0005	16.09	0.0003	49,437
Odds ratio awareness with DoA monitor	0.25	13.25	-0.0007	29.36	-0.0004	16.12	0.0003	51,167
	0.81	13.25	-0.0007	29.69	-0.0005	16.45	0.0002	97,987
Duration of LPS (years)	0.25	13.25	-0.0007	29.48	-0.0004	16.23	0.0003	62,812
	1	13.25	-0.0007	29.48	-0.0004	16.23	0.0003	60,065
Probability of LPS[a]	0.195	13.01	-0.0122	29.37	-0.0121	16.36	0.0002	84,329
	0.48	13.52	-0.0295	29.61	-0.0292	16.08	0.0003	52,685
Duration of PTSD (years)	5.6	13.25	-0.0006	29.48	-0.0004	16.23	0.0002	70,492
	9.6	13.25	-0.0007	29.48	-0.0004	16.23	0.0003	53,244
Proportion PTSD[b]	0.345	13.03	-0.0006	29.38	-0.0004	16.35	0.0002	75,647
	0.733	13.45	-0.0008	29.57	-0.0004	16.12	0.0003	52,469
LPS QoL decrement	-0.075	13.25	-0.0007	29.48	-0.0004	16.23	0.0003	62,573
	-0.05	13.25	-0.0007	29.48	-0.0004	16.23	0.0003	62,972
PTSD QoL decrement	-0.134	13.25	-0.0007	29.48	-0.0004	16.23	0.0003	58,328
	-0.068	13.25	-0.0006	29.48	-0.0004	16.23	0.0002	79,881
Probability people with PTSD seek treatment	0	12.66	-0.0007	29.22	-0.0004	16.56	0.0003	63,103
	1	14.39	-0.0007	30.00	-0.0004	15.61	0.0003	59,473
Unit cost of sensors	13.3125	13.25	-0.0007	25.04	-0.0004	11.80	0.0003	44,957
	22.1875	13.25	-0.0007	33.92	-0.0004	20.67	0.0003	78,780

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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 is again sensitive to several key parameters in a general surgical population undergoing mixed anaesthesia. In this group increase in the probability of LPS resulted in the largest variation, to £84, 329 per QALY gained. The ICER was again driven by probability of

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

Scenario analysis

a) Inclusion of anaesthesia-related complication (PONV)

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[106] to investigate the potential impact of including this outcome on the cost effectiveness results.

For this scenario analysis we used the baseline (control group) risk of PONV as the estimated risk for standard clinical monitoring and applied the odds ratio 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 intravenous injection (unit cost = £5.39, BNF no 62, September 2011[33]).

Table 64 and Table 65 report the results of the scenario analysis for patients at high risk of intraoperative awareness and a general surgical population, respectively, undergoing general anaesthesia with TIVA. The incremental costs for BIS monitoring are reduced, from the value reported for the base case analyses (Table 52 and Table 54), by including an estimate of PONV. However the change in costs is slight and leave the ICERs largely unchanged.

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Odds ratio = 0.77: baseline risk = 0.3, risk with BIS monitoring = 0.248					
Standard clinical monitoring	24.19		-0.0011		
BIS	42.48	18.29	-0.0005	0.0007	26,933
Odds ratio = 0.56[a]: baseline risk = 0.3, risk with BIS monitoring = 0.194					
Standard clinical monitoring	24.19		-0.0011		
BIS	42.19	17.99	-0.0005	0.0007	26,500
[a] lower limit of 95% CI estimated by Liu[106]					

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Odds ratio = 0.77: baseline risk = 0.375, risk with BIS monitoring = 0.248					
Standard clinical monitoring	23.13		-0.0007		
BIS	37.15	14.02	-0.0004	0.0003	44,153
Odds ratio = 0.56[a]: baseline risk = 0.375, risk with BIS monitoring = 0.194					
Standard clinical monitoring	23.13		-0.0007		
BIS	38.86	13.73	-0.0004	0.0003	43,227
[a] lower limit of 95% CI estimated by Liu[106]					

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

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Odds ratio = 0.77: baseline risk = 0.3, risk with BIS monitoring = 0.248					
Standard clinical monitoring	14.31		-0.0011		
BIS	32.95	18.64	-0.0006	0.0005	35,592
Odds ratio = 0.56[a]: baseline risk = 0.3, risk with BIS monitoring = 0.194					
Standard clinical monitoring	14.31		-0.0011		
BIS	32.66	18.34	-0.0006	0.0005	35,031
[a] lower limit of 95% CI estimated by Liu[106]					

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Odds ratio = 0.77: baseline risk = 0.3, risk with BIS monitoring = 0.248					
Standard clinical monitoring	13.25		-0.0007		
BIS	29.20	15.95	-0.0004	0.0003	60,803
Odds ratio = 0.56[a]: baseline risk = 0.3, risk with BIS monitoring = 0.194					
Standard clinical monitoring	13.25		-0.0007		
BIS	28.91	15.66	-0.0004	0.0003	59,682
[a] lower limit of 95% CI estimated by Liu[106]					

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.

b) 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, while 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 a value of 1.0% reported for certain types of surgery (cardiac surgery, Caesarean section and trauma surgery).[79;112;113] 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 (see Table 68).

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
TIVA					
Standard clinical monitoring	26.22		-0.0021		
BIS	43.25	17.03	-0.0007	0.0014	12,497
Mixed anaesthesia					
Standard clinical monitoring	16.34		-0.0021		
BIS	34.15	17.81	-0.0010	0.0010	17,510

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 (see Table 69 and Table 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-four-fold 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 ICERs.

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Baseline probability of awareness = 0.99%					
Standard clinical monitoring	26.18		-0.0020		
BIS	38.17	11.99	-0.0007	0.0014	8,874
Baseline probability of awareness = 0.007%					
Standard clinical monitoring	22.56		-0.0004		
BIS	37.29	14.73	-0.0003	0.0001	116,252

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Baseline probability of awareness = 0.99%					
Standard clinical monitoring	16.30		-0.0020		
BIS	30.86	14.56	-0.0010	0.0010	14,443
Baseline probability of awareness = 0.007%					
Standard clinical monitoring	12.68		-0.0004		
BIS	29.23	16.54	-0.0003	0.0001	133,094

c) 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 which the monitor module was used in each year (patients per device year), throughout its assumed five year effective life. The estimate used for the general surgical population was 1000 patients per year (equivalent to 4 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. Table 71 and Table 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 (6 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 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	26.87	39,576	23.65	74,453
500	18.57	27,345	15.34	48,302
1000	17.53	25,816	14.30	45,033
1500	17.18	25,306	13.96	43,944

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	27.22	51,988	25.58	97,477
500	18.92	36,126	17.27	65,825
1000	17.88	34,144	16.23	61,869
1500	17.53	33,483	15.89	60,550

d) Impact of utility decrement for PTSD

The quality of life decrement applied in the base case analysis was based upon Freed and colleagues'[121] paper in 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 (see Table 73 below).

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = -0.5					
Standard clinical monitoring	24.19		-0.0034		
BIS	42.76	18.57	-0.0010	0.0024	7,872
Utility decrement for PTSD = -0.75					
Standard clinical monitoring	24.19		-0.0048		
BIS	42.76	18.57	-0.0014	0.0035	5,361
Mixed anaesthesia					
Utility decrement for PTSD = -0.5					
Standard clinical monitoring	14.31		-0.0034		
BIS	33.23	18.92	-0.0016	0.0017	10,884
Utility decrement for PTSD = -0.75					
Standard clinical monitoring	14.31		-0.0048		
BIS	33.23	18.92	-0.0023	0.0025	7,456

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 £7,872 per QALY gained and £10,884 per QALY gained respectively. Where the PTSD decrement was increased further, the ICER decreased again to £5,361 and £7,456 per QALY gained in the TIVA and mixed anaesthesia groups respectively.

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = -0.5					
Standard clinical monitoring	23.13		-0.0015		
BIS	37.43	14.30	-0.0005	0.0009	15,627
Utility decrement for PTSD = -0.75					
Standard clinical monitoring	23.13		-0.0020		
BIS	37.43	14.30	-0.0007	0.0013	10,931
Mixed anaesthesia					
Utility decrement for PTSD = -0.5					
Standard clinical monitoring	13.25		-0.0015		
BIS	29.48	16.23	-0.0008	0.0007	23,366
Utility decrement for PTSD = -0.75					
Standard clinical monitoring	13.25		-0.0020		
BIS	29.48	16.23	-0.0010	0.0010	16,579

The scenario analyses using alternative PTSD decrements in the general surgical population reflects the results in the high risk population: there is a substantial reduction in the ICERs where these are increased.

Entropy compared with Standard clinical monitoring

Base case

Total intravenous anaesthesia (TIVA)

The costs, QALYs and incremental cost effectiveness ratio (ICER) modelled for patients considered at high risk of intra-operative awareness undergoing general anaesthesia with TIVA, comparing standard clinical monitoring with monitoring by Entropy are presented in Table 75 below.

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

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

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 6 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 Entropy monitoring, with the incremental cost being £9.79. The breakdown of the total cost for standard clinical monitoring and Entropy are reported in Table 53 and the costs of anaesthetic drug use outlined in this table apply to the Ellerkman 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 Entropy monitoring is partially offset by the reduction in costs of patients with PTSD.

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

Cost	Standard clinical monitoring (£)	Entropy (£)
Depth of anaesthesia monitoring	0.00	11.05
Anaesthetic drugs	23.11	23.11
Post-operative nausea and vomiting	1.62	1.62
Post-operative cognitive dysfunction	0.00	0.00
PTSD	1.66	0.40

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 Entropy monitored patients, with a difference of 0.0007 QALYs, resulting in an ICER of £14,421 per QALY gained.

In a general surgical population (not just those at high risk of intra-operative awareness) undergoing general anaesthesia with TIVA, 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 4 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, QALYs and ICER modelled for this population, comparing standard clinical monitoring with monitoring by Entropy are presented in Table 77 (based on anaesthetic drug consumption from the RCT by Ellerkman and colleagues[62]) and in Table 78 (based on anaesthetic drug consumption from the RCT by Gruenewald and colleagues[55]).

Table 77 - Cost effectiveness of Entropy compared with standard clinical monitoring in a general surgical population undergoing TIVA (drug use based on Ellerkman et al)

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

Table 78 - Cost effectiveness of Entropy compared with standard clinical monitoring in a general surgical population undergoing TIVA (drug use based on Gruenewald et al)

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

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

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

Cost	Standard clinical monitoring (£)	Entropy (£)
Depth of anaesthesia monitoring	0.00	9.87
Anaesthetic drugs Ellerkman and colleagues Gruenewald and colleagues	 23.11 29.29	 23.58 29.85
Post-operative nausea and vomiting	1.62	1.62
Post-operative cognitive dysfunction	0.00	0.00
PTSD	0.59	0.14

The QALY loss incurred by patients undergoing 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 Ellerkman and colleagues RCT[62] were applied, and £31, 430 where anaesthetic consumption from Gruenewald and colleagues[55] were applied.

Mixed anaesthesia (induction with IV anaesthetic (propofol and sufentanil) and maintenance with IV and inhaled anaesthetic (sufentanil and sevoflurane))

The costs, QALYs and ICER modelled for patients considered at high risk of intra-operative awareness undergoing mixed anaesthesia, comparing standard clinical monitoring with monitoring by Entropy are presented in Table 80.

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

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

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, below. Sufentanil costs are not included as this 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 PTSD, costs for PTSD were lower in the group undergoing Entropy monitoring. The incremental cost of Entropy monitoring was £10.14.

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

Cost	Standard clinical monitoring (£)	Entropy (£)
Depth of anaesthesia monitoring	0.00	11.05
Anaesthetic drugs	15.93	15.93
Post-operative nausea and vomiting	1.62	1.62
Post-operative cognitive dysfunction	0.00	0.00
PTSD	1.66	0.75

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 Entropy patients of 0.0005. This yielded an ICER of £19,367 per QALY gained.

In a general surgical population 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 3 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 general anaesthesia with both IV and inhaled anaesthetic, comparing standard clinical monitoring with monitoring by Entropy are presented in Table 82.

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

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

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

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

Cost	Standard clinical monitoring (£)	Entropy (£)
Depth of anaesthesia monitoring	0.00	9.87
Anaesthetic drugs	15.93	11.37
Post- operative nausea and vomiting	1.62	1.62
Post- operative cognitive dysfunction	0.00	0.00
PTSD	0.59	0.27

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 (TIVA)

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 Ellerkman 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 cost are the same for both standard clinical monitoring and entropy and therefore cancel out in the calculation of incremental cost and in the ICER.

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

Parameter	Input value	Standard clinical monitoring		Entropy		Incremental		ICER (£/ QALY gained)
		Cost (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
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	4,834
Odds ratio awareness with DoA 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 (yrs)	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

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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 ICERs resulting from the one-way sensitivity analysis in a high risk population receiving TIVA ranged from £4, 834 - £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 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 £4, 834 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 Ellerkman and colleagues[62]) and in Table 86 (anaesthetic drug costs based on usage reported by Gruenewald and colleagues[55])

Table 85 - One way sensitivity analysis: Entropy compared with standard clinical monitoring in a general surgical population, undergoing TIVA (drug use based on Ellerkman et al)

Parameter	Input value	Standard clinical monitoring		Entropy		Incremental		ICER (£/ QALY gained)
		Cost (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
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
Odds ratio awareness with DoA 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 (yrs)	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

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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 Ellerkman and colleagues[62] (Table 85) reflect 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, 236 per QALY gained, and £41, 419 per

QALY gained,) a reduction in effectiveness of the Entropy module (£45, 292 per QALY gained,) and the probability of LPS and a reduction in the PTSD QoL decrement.

Table 86 - One way sensitivity analysis: Entropy compared with standard clinical monitoring in a general surgical population, undergoing TIVA (drug use based on Gruenewald et al)

Parameter	Input value	Standard clinical monitoring		Entropy		Incremental		ICER (£/ QALY gained)
		Cost (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
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
Odds ratio awareness with DoA 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 (yrs)	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

Notes

[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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

Again, the one-way sensitivity analysis results in the general surgical population receiving TIVA and applying costs from Gruenewald and colleagues[55] to inform costs, (Table 86) reflect those in the high risk group. Whilst the ICER 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 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 IV anaesthetic (remifentanil) and maintenance with IV and inhaled anaesthetic (remifentanil and sevoflurane)). The results are shown in Table 87 and Table 88.

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

Parameter	Input value	Standard clinical monitoring		Entropy		Incremental		ICER (£/ QALY gained)
		Cost (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
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	7,290
Odds ratio awareness with DoA 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 (yrs)	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

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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 ICERs of £7,290 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 Entropy module, giving an £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, whilst being insensitive to the LPS decrement and duration.

Table 88 - One way sensitivity analysis: 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	6,494
	-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
Odds ratio awareness with DoA 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 (yrs)	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

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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

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 ICERs of £6,494 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 Entropy module, which results in an ICER of £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

a) Inclusion of anaesthesia-related complication (PONV)

The systematic review of patient outcomes did not identify any robust data which reported an estimate of the effect of Entropy monitoring on risk of post-operative nausea and vomiting (PONV). We developed a scenario analysis using data from a meta analysis by Liu,[106] 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%,[103-105] for standard clinical monitoring and applied the odds ratio derived in the meta analysis (0.77, 95% CI 0.56 to 0.99) to estimate risk for 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 intravenous injection (unit cost = £5.39, BNF no 62, September 2011[33]).

Table 89 and Table 90 report the results of this scenario analysis for high risk patients and general surgical patients, respectively, undergoing general anaesthesia with TIVA.

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Odds ratio = 0.77: baseline risk = 0.3, risk with Entropy monitoring = 0.248					
Standard clinical monitoring	26.38		-0.0011		
Entropy	35.60	9.51	-0.0005	0.0007	14,010
Odds ratio = 0.56[a]: baseline risk = 0.3, risk with Entropy monitoring = 0.194					
Standard clinical monitoring	26.38		-0.0011		
Entropy	35.60	9.22	-0.0005	0.0007	13,576
[a] lower limit of 95% CI estimated by Liu[106]					

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

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Anaesthetic drug consumption based on Ellerkman and colleagues[62]					
Odds ratio = 0.77: baseline risk = 0.3, risk with Entropy monitoring = 0.248					
Standard clinical monitoring	25.32		-0.0007		
Entropy	34.92	9.61	-0.0004	0.0003	30,250
Odds ratio = 0.56: baseline risk = 0.3, risk with Entropy monitoring = 0.194					
Standard clinical monitoring	25.32		-0.0007		
Entropy	34.63	9.31	-0.0004	0.0003	29,324
Anaesthetic drug consumption based on Gruenewald and colleagues[62]					
Odds ratio = 0.77: baseline risk = 0.3, risk with Entropy monitoring = 0.248					
Standard clinical monitoring	31.50		-0.0007		
Entropy	41.20	9.70	-0.0004	0.0003	30,550
Odds ratio = 0.56: baseline risk = 0.3, risk with Entropy monitoring = 0.194					
Standard clinical monitoring	31.50		-0.0007		
Entropy	40.90	9.41	-0.0004	0.0003	29,624
[a] lower limit of 95% CI estimated by Liu[106]					

Again, changes in the OR of PONV due to Entropy monitoring make little difference to the ICER in a general surgical population undergoing TIVA. The base case ICER of £31, 131 applying Ellerkman 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).

Table 91 and Table 92 report the results of this scenario analysis for patients at high risk and for patients at average risk of intra-operative awareness, respectively, undergoing general

anaesthesia with mixed anaesthesia (induction with IV anaesthetic and maintenance with IV and inhaled anaesthetic).

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Odds ratio = 0.77: baseline risk = 0.3, risk with Entropy monitoring = 0.248					
Standard clinical monitoring	19.20		-0.0011		
Entropy	29.07	9.86	-0.0006	0.0005	18,833
Odds ratio = 0.56: baseline risk = 0.3, risk with Entropy monitoring = 0.194					
Standard clinical monitoring	19.20		-0.0011		
Entropy	28.77	9.57	-0.0006	0.0005	18,271
[a] lower limit of 95% CI estimated by Liu[106]					

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 ICERs of £18, 833 and £18, 271 per QALY gained respectively.

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Odds ratio = 0.77: baseline risk = 0.3, risk with Entropy monitoring = 0.248					
Standard clinical monitoring	18.14		-0.0007		
Entropy	22.84	4.71	-0.0004	0.0003	17,934
Odds ratio = 0.56: baseline risk = 0.3, risk with Entropy monitoring = 0.194					
Standard clinical monitoring	18.14		-0.0007		
Entropy	22.55	4.41	-0.0004	0.0003	16,813
[a] lower limit of 95% CI estimated by Liu[106]					

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.

b) Scenario analyses for probability of intra-operative awareness for patients at high risk of intra-operative awareness and for the general surgical population

Our review of published studies of the incidence of intra-operative 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[140], Ranta and colleagues[113] and Myles and colleagues[79]), while 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% (see Table 93). The effect of this is to reduce the ICER to £6,059 per QALY gained for TIVA and to £8,882 for mixed anaesthesia.

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
TIVA					
Standard clinical monitoring	28.41		-0.0021		
Entropy	36.67	8.26	-0.0007	0.0014	6,059
Mixed anaesthesia					
Standard clinical monitoring	21.23		-0.0021		
Entropy	30.26	9.03	-0.0010	0.0010	8,882

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%, see Table 94 and Table 95).

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Anaesthetic drug consumption based on Ellerkman and colleagues[62]					
Baseline probability of awareness = 0.99%					
Standard clinical monitoring	28.37		-0.0020		
Entropy	35.94	7.57	-0.0007	0.0014	5,605
Baseline probability of awareness = 0.007%					
Standard clinical monitoring	24.75		-0.0004		
Entropy	35.07	10.32	-0.0003	0.0001	81,406
Anaesthetic drug consumption based on Gruenewald and colleagues[55]					
Baseline probability of awareness = 0.99%					
Standard clinical monitoring	34.55		-0.0020		
Entropy	42.22	7.67	-0.0007	0.0014	5,676
Baseline probability of awareness = 0.007%					
Standard clinical monitoring	30.93		-0.0004		
Entropy	41.34	10.41	-0.0003	0.0001	82,157

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 Ellerkman and colleagues[62] or Gruenewald and colleagues[55]) these ranges from approximately £5, 600 per QALY gained to approximately £80, 000 per QALY gained respectively.

In threshold analyses we found that depth of anaesthesia monitoring with Entropy for patients undergoing general anaesthesia with TIVA was cost effective if the probability of awareness was greater than 0.192% to 0.194%, at a willingness to pay threshold of £30,000 per QALY gained. Depth of anaesthesia monitoring with Entropy was cost effective if the probability of awareness was greater than 0.315% to 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, since the exact values depend on which study the anaesthetic drug consumption is based on (Ellerkman and colleagues[62] or Gruenewald and colleagues[55]).

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Baseline probability of awareness = 0.99%					
Standard clinical monitoring	21.19		-0.0020		
Entropy	24.51	3.31	-0.0010	0.0010	3,286
Baseline probability of awareness = 0.007%					
Standard clinical monitoring	17.57		-0.0004		
Entropy	22.87	5.30	-0.0003	0.0001	42,599

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 £3, 286.

In threshold analyses we found that depth of anaesthesia monitoring with Entropy for patients undergoing mixed general anaesthesia was cost effective if the probability of awareness was greater than 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%.

c) 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 five year effective life. The estimate used for the general surgical population was 1000 patients per year (equivalent to 4 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 Entropy monitoring, compared with standard clinical monitoring, and the resulting effect on the ICER. Table 96 reports the incremental cost and ICER for Entropy compared, compared with standard clinical monitoring, at four selected values for the number of patients per device year: the base case value of 500 and also for a low value of 10 and high values of 1,000 (4 patients per day over 250 working days per year) and 1,500 (6 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.

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

Patients per device year	Standard clinical monitoring (£)	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
1,000	25.32	35.20	9.89	31,131
1,500	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
1,000	18.14	23.12	4.99	19,000
1,500	18.14	22.73	4.59	17,494

d) Impact of alternative assumptions on the utility decrement for PTSD

The quality of life decrement applied in the base case was based upon Freed and colleagues'[121] paper in 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 (see Table 97 and Table 98).

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = 0.50					
Standard clinical monitoring	26.38		-0.0034		4,152
Entropy	36.18	9.79	-0.0010	0.0024	
Utility decrement for PTSD = 0.75					
Standard clinical monitoring	26.38		-0.0048		
Entropy	36.18	9.79	-0.0014	0.0035	2,827
Mixed anaesthesia					
Utility decrement for PTSD = 0.50					
Standard clinical monitoring	19.20		-0.0034		
Entropy	29.35	10.14	-0.0016	0.0017	5,835
Utility decrement for PTSD = 0.75					
Standard clinical monitoring	19.20		-0.0048		
Entropy	29.35	10.14	-0.0023	0.0025	3,997

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 the ICER reduced to £10,907 per QALY gained and £5,835 per QALY gained respectively. Where the PTSD decrement was increased further, the ICER decreased again to £7,629 and £3,997 per QALY gained in the TIVA and mixed groups respectively.

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

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = 0.50					
Standard clinical monitoring	25.32		-0.0015		
Entropy	35.20	9.98	-0.0005	0.0009	10,803
Utility decrement for PTSD = 0.75					
Standard clinical monitoring	25.32		-0.0020		
Entropy	35.20	9.89	-0.0007	0.00013	7,556
Mixed anaesthesia					
Utility decrement for PTSD = 0.50					
Standard clinical monitoring	18.14		-0.0015		
Entropy	23.12	4.99	-0.0008	0.0007	7,176
Utility decrement for PTSD = 0.75					
Standard clinical monitoring	18.14		-0.0020		
Entropy	23.12	4.99	-0.0010	0.0010	5,091

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

Narcotrend compared with Standard Clinical Monitoring

Base case

Total intravenous anaesthesia (TIVA)

The costs, QALYs and ICER modelled for patients considered at high risk of intra-operative awareness undergoing general anaesthesia with TIVA, comparing standard clinical monitoring with depth of anaesthesia monitoring by Narcotrend are presented in Table 99 below.

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Standard clinical monitoring	33.45		-0.0011		
Narcotrend	37.31	3.86	-0.0005	0.0007	5,681

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 6 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 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 censors 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 (see Table 100).

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
Post-op nausea and vomiting	1.62	1.62
Post-op cognitive dysfunction	0.00	0.00
PTSD	1.66	0.40

Patients in both groups incurred a slight QALY loss, resulting from psychological sequelae of awareness (LPS and PTSD) and from post-operative cognitive dysfunction in older patients. This was lower in the Narcotrend monitored patients, with a difference of 0.0007 QALYs, resulting in an ICER of £5,681 per QALY gained.

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

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

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

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 4 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, due to the

lower incidence of awareness) which offset the additional costs associated with depth of anaesthesia monitoring, see Table 102.

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
Post-op nausea and vomiting	1.62	1.62
Post-op cognitive dysfunction	0.00	0.00
PTSD	0.59	0.14

Given the lower probability of intra-operative 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 post-operative cognitive dysfunction are lower for patients monitored with Narcotrend compared with those receiving standard clinical monitoring. Since better outcomes are modelled as being achieved at lower costs, Narcotrend dominates standard clinical monitoring for this population.

Mixed anaesthesia (induction with IV anaesthetic (remifentanil) and maintenance with IV and inhaled anaesthetic (remifentanil and desflurane))

The costs, QALYs and ICER modelled for patients considered at high risk of intra-operative awareness undergoing mixed anaesthesia, comparing standard clinical monitoring with depth of anaesthesia monitoring by Narcotrend are presented in Table 103.

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Standard clinical monitoring	38.99		-0.0011		
Narcotrend	43.20	4.21	-0.0006	0.0005	8,033

Naroctrend 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).

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
Post-op nausea and vomiting	1.62	1.62
Post-op cognitive dysfunction	0.00	0.00
PTSD	1.66	0.75

The reduced QALY loss in high risk patients undergoing monitoring with Narcotrend compared with patients undergoing standard monitoring was due to the lower probability of awareness in this group, with a difference of 0.0005 QALYs. This resulted in an ICER of 8,033 per QALY gained.

The costs, QALYs and ICER modelled for a general surgical population (not just those at high risk of intra-operative awareness) undergoing general anaesthesia with TIVA, comparing standard clinical monitoring with depth of anaesthesia monitoring by Narcotrend are presented in Table 105.

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

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

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 3 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, see Table 106. This arises from the relatively small additional cost of depth of anaesthesia monitoring with Narcotrend (the sensors are available a low cost, while the capital cost of the monitor is spread across a relatively large patient throughput) and from savings due to a reduction in the use of anaesthetic drugs (and to a lesser extent with lower PTSD-related costs, due to the lower incidence of awareness).

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
Post-op nausea and vomiting	1.62	1.62
Post-op cognitive dysfunction	0.00	0.00
PTSD	0.59	0.27

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

Deterministic sensitivity analysis

Total intravenous anaesthesia (TIVA)

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 Table 107 and Table 108.

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 (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
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	1,123
Odds ratio awareness with DoA monitor	0.1	33.45	-0.0011	37.08	-0.0004	3.63	0.0008	4,631
	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	5,812
	1	33.45	-0.0012	37.31	-0.0005	3.86	0.0007	5,436
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	3,861
Duration of PTSD (yrs)	5.6	33.45	-0.0010	37.31	-0.0004	3.86	0.0006	6,959
	9.6	33.45	-0.0014	37.31	-0.0005	3.86	0.0008	4,570
Proportion PTSD[b]	0.345	32.85	-0.0009	37.17	-0.0004	4.32	0.0005	8,640
	0.733	34.04	-0.0014	37.45	-0.0005	3.41	0.0009	4,002
LPS QoL decrement	-0.075	33.45	-0.0011	37.31	-0.0005	3.86	0.0007	5,779
	-0.05	33.45	-0.0011	37.31	-0.0005	3.86	0.0007	5,835
PTSD QoL decrement	-0.134	33.45	-0.0012	37.31	-0.0005	3.86	0.0007	5,207
	-0.068	33.45	-0.0008	37.31	-0.0004	3.86	0.0004	8,589
Probability people with PTSD seek treatment	0	31.80	-0.0011	36.91	-0.0005	5.12	0.0007	7,534
	1	36.67	-0.0011	38.09	-0.0005	1.42	0.0007	2,085
Unit cost of sensors (£)	0.42	33.45	-0.0011	37.17	-0.0005	3.72	0.0007	5,475
	0.70	33.45	-0.0011	37.45	-0.0005	4.00	0.0007	5,887

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as a proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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 one way sensitivity analysis of key parameters in the high risk surgical group receiving TIVA resulted in ICERs ranging from £1,123 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.

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		BIS		Incremental		ICER (£/ QALY gained)
		Cost (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
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	
Odds ratio awareness with DoA 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 (yrs)	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	

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as a proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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 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 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 IV anaesthetic (remifentanil) and maintenance with IV and inhaled anaesthetic (remifentanil and desflurane)). The results are shown in Table 109 and Table 110.

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

Parameter	Input value	Standard clinical monitoring		BIS		Incremental		ICER (£/ QALY gained)
		Cost (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
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	2,290
Odds ratio awareness with DoA monitor	0.25	38.99	-0.0011	42.87	-0.0005	3.87	0.0007	5,769
	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	8,206
	1	38.99	-0.0012	43.20	-0.0006	4.21	0.0005	7,707
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	5,865
Duration of PTSD (yrs)	5.6	38.99	-0.0010	43.20	-0.0006	4.21	0.0004	9,704
	9.6	38.99	-0.0014	43.20	-0.0007	4.21	0.0006	6,542
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	5,982
LPS QoL decrement	-0.075	38.99	-0.0011	43.20	-0.0006	4.21	0.0005	8,162
	-0.05	38.99	-0.0011	43.20	-0.0006	4.21	0.0005	8,236
PTSD QoL decrement	-0.134	38.99	-0.0012	43.20	-0.0007	4.21	0.0006	7,401
	-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	9,770
	1	42.21	-0.0011	44.65	-0.0006	2.44	0.0005	4,661
Unit cost of sensors (£)	0.42	38.99	-0.0011	43.06	-0.0006	4.07	0.0005	7,766
	0.70	38.99	-0.0011	43.34	-0.0006	4.35	0.0005	8,300

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as a proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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 £2,290 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 £2,290 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.

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		BIS		Incremental		ICER (£/ QALY gained)
		Cost (£)	QALYs	Cost (£)	QALYs	Cost	QALYs	
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	2,534
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	
Odds ratio awareness with DoA 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 (yrs)	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	

Notes
[a] changing the probability of LPS also changes the probability of PTSD, since PTSD is calculated as a proportion of the probability of LPS (i.e. people with PTSD are a sub-group 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 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% confidence interval is close to zero, indicating only limited savings in cost of anaesthetic gas to offset against the cost of BIS monitoring, resulting in a positive incremental cost.

Scenario analysis

a) Inclusion of anaesthesia-related complication (post-operative nausea and vomiting)

The systematic review of patient outcomes did not identify any robust data which reported an estimate of the effect of Narcotrend monitoring on risk of post-operative nausea and vomiting (PONV). We developed a scenario analysis using data from a meta analysis by Liu,[106] 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%,[103-105] for standard clinical monitoring and applied the odds ratio 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 intravenous injection (unit cost = £5.39, BNF no 62, September 2011[33]).

Table 111 and Table 112 report the results of this scenario analysis for high risk patients and general surgical patients, respectively, undergoing general anaesthesia with TIVA.

Table 111 - Scenario analysis: including an estimated effect of Narcotrend monitoring on the incidence of PONV in patients at high risk of awareness undergoing h TIVA

Intervention	Cost per patient (£)	Incremental Cost (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Odds ratio = 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	5,270
Odds ratio = 0.56: baseline risk = 036.74.3, risk with Narcotrend monitoring = 0.194					
Standard clinical care	33.45		-0.0011		
Narcotrend	36.74	3.28	-0.0005	0.0007	4,836

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.

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Odds ratio = 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
Odds ratio = 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 and Table 114 report the results of this scenario analysis for patients at high risk and for patients at average risk of intra-operative awareness, respectively, undergoing general anaesthesia with mixed anaesthesia (induction with IV anaesthetic and maintenance with IV and inhaled anaesthetic).

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Odds ratio = 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	7,499
Odds ratio = 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	6,937

Where the variations in the OR of PONV are applied to the high risk patients undergoing mixed anaeasthesia there is a slight reduction in the ICER. An OR of 0.77 results in an ICER of £7, 499 per QALY gained and an OR of 0.56 yields an ICER of £6,937 per QALY gained in this group.

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Odds ratio = 0.77: baseline risk = 0.3, risk with Narcotrend monitoring = 0.248					
Standard clinical care	37.93		-0.0007		Narcotrend dominates
Narcotrend	35.90	-2.02	-0.0004	0.0003	
Odds ratio = 0.56: baseline risk = 0.3, risk with Narcotrend monitoring = 0.194					
Standard clinical care	37.93		-0.0007		Narcotrend dominates
Narcotrend	35.61	-2.32	-0.0004	0.0003	

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.

b) Scenario analyses for probability of intra-operative awareness for patients at high risk of intra-operative awareness and for the general surgical population

Our review of published studies of the incidence of intra-operative 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[140], Ranta and colleagues[113] and Myles and colleagues[79]), while 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% (see Table 115).

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
TIVA					
Standard clinical care	35.48		-0.0021		
Narcotrend	37.80	2.32	-0.0007	0.0014	1,705
Mixed anaesthesia					
Standard clinical care	41.02		-0.0021		
Narcotrend	44.12	3.10	-0.0010	0.0010	3,047

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 £8, 033 to £3, 047 per QALY gained in the group receiving mixed, and from £5, 681 to £1,705 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%, see Table 116 and Table 117).

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Baseline probability of awareness = 0.99%					
Standard clinical care	35.44		-0.0020		Nacrotrend dominates
Nacrotrend	29.27	-6.17	-0.0007	-0.0014	
Baseline probability of awareness = 0.007%					
Standard clinical care	31.82		-0.0004		Nacrotrend dominates
Nacrotrend	28.40	-3.43	-0.0003	0.0001	

Where the outlying probabilities are applied the ICER is robust and Narcotrend continues to dominate in TIVA and mixed anaesthesia patients.

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
Baseline probability of awareness = 0.99%					
Standard clinical care	40.98		-0.0020		
Nacrotrend	37.57	-3.42	-0.0010	0.0010	Nacrotrend dominates
Baseline probability of awareness = 0.007%					
Standard clinical care	37.36		-0.0004		
Nacrotrend	35.93	-1.43	-0.0003	0.0001	Nacrotrend dominates

c) 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 five year effective life. The estimate used for the general surgical population was 1000 patients per year (equivalent to 4 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 1,500 (6 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.

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 (£)	Nacrotrend (£)	Incremental cost (£)	ICER (£/ QALY gained
TIVA				
100	32.39	49.03	16.65	52,414
500	32.39	30.81	-1.58	Nacrotrend dominates
1,000	32.39	28.53	-3.85	Nacrotrend dominates
1,500	32.39	27.7	-4.61	Nacrotrend dominates
Mixed anaesthesia				
100	37.93	26.68	18.76	71,484
500	37.93	38.46	0.53	2,035
1,000	37.93	36.18	-1.74	Nacrotrend dominates
1,500	37.93	35.42	-2.50	Nacrotrend dominates

d) Impact of alternative assumptions on the utility decrement for PTSD

The quality of life decrement applied in the base case was based upon Freed and colleagues'[121] paper in 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 (see Table 119) or mixed anaesthesia (Table 120).

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = 0.50					
Standard clinical care	33.45		-0.0034		
Nacrotrend	37.31	3.86	-0.0010	0.0024	1,636
Utility decrement for PTSD = 0.75					
Standard clinical care	33.45		-0.0048		
Nacrotrend	37.31	3.86	-0.0014	0.0035	1,114
Mixed anaesthesia					
Utility decrement for PTSD = 0.50					
Standard clinical care	38.99		-0.0034		
Nacrotrend	43.20	4.21	-0.0016	0.0017	2,420
Utility decrement for PTSD = 0.75					
Standard clinical care	38.99		-0.0048		
Nacrotrend	43.20	4.21	-0.0023	0.0025	1,658

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 £1, 636 and £1,114 per QALY gained for a 0.5 and 0.75 decrement respectively in the group undergoing TIVA. The ICER is reduced to £2, 420 and £1, 658 for a 0.5 and 0.75 decrement in the group undergoing mixed anaesthesia.

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 (£)	QALYs	Incremental QALYs	ICER (£/QALY gained)
TIVA					
Utility decrement for PTSD = 0.50					
Standard clinical care	32.39		-0.0015		Nacrotrend dominates
Nacrotrend	28.53	-3.85	-0.0005	0.0009	
Utility decrement for PTSD = 0.75					
Standard clinical care	32.39		-0.0020		Nacrotrend dominates
Nacrotrend	28.53	-3.85	-0.0007	0.0013	
Mixed anaesthesia					
Utility decrement for PTSD = 0.50					
Standard clinical care	37.93		-0.0015		Nacrotrend dominates
Nacrotrend	36.18	-1.74	-0.0008	0.0007	
Utility decrement for PTSD = 0.75					
Standard clinical care	37.93		-0.0020		Nacrotrend dominates
Nacrotrend	36.18	-1.74	-0.0010	0.0010	

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, 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 Entropy or Nacrotrend 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 maybe 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 may, depending on the utility losses associated with these conditions, be acceptable in cost effectiveness terms. The economic analysis suggests that, other than at comparatively low patient volumes, the acquisition cost of the depth of anaesthesia 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.

6. 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 on-line 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 patient's families and friends, for example, to provide social, emotional and practical support. Strain may be placed upon 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).

7. DISCUSSION

7.1 Statement of principal findings

7.1.1 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 (7 and 4 RCTs respectively). A notable feature of the primary studies within each of the BIS, 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 US and Canada, randomised at least 3020 patients per study group, and patients received only inhaled general anaesthesia. 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 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 Odds Ratio remained statistically significant 0.45 (95% CI 0.25, 0.81), favouring BIS, though with significant heterogeneity. We classified the trials into sub-groups based on the type of general anaesthesia used (inhaled only; mixed inhaled and intravenous; total intravenous). The pooled Peto Odds Ratios for the sub-groups of mixed inhaled and intravenous general anaesthesia, and total intravenous anaesthesia were both consistent with the overall pooled Odds Ratio (i.e.

statistically significant in favour of BIS). In contrast, the pooled estimate for the trials of inhaled general anaesthesia, including the BAG-RECALL RCT and another large RCT (the B-Unaware trial[27]), favoured standard clinical monitoring though the confidence intervals overlapped with 1 indicating potential advantage to both BIS and to standard clinical monitoring. Importantly the BAG-RECALL RCT was designed to overcome some of the methodological limitations of the B-Unaware trial, 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.

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 they were 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 section 3.1.3). 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. Greunewald 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 to those estimated from much larger studies (Table 1), although in the Gruenewald study awareness was experienced by only one patient.[55]

The case of awareness reported by Gruenewald and colleagues might have happened outside of the period of general anaesthesia, since 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 did not specify that their patients were at risk of awareness, the patients did appear to be relatively old (early 60s age), possibly overweight or obese, and half of them had notable illness (ASA physical score grade III). The awareness assessment conducted by Kerssens and colleagues involved asking patients five questions that were very similar to

those of the Brice interview. Both 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 three days after surgery (30 days after extubation); all other trials that assessed intraoperative awareness conducted follow-up assessments only one day or less post-surgery,[54;57;59;60] three 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 (section 3.1.3), 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 general anaesthesia 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. Whilst 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 to 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 intravenous), 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 (entropy) and Table 24 (Narcotrend) in Section 5.1 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 RCTs reported significant reductions of sevoflurane consumption under entropy-guided anaesthesia relative to standard clinical monitoring (i.e. favouring the entropy group) in both adults[61] and children,[54] but no difference in propofol consumption between BIS, entropy and standard clinical monitoring groups in adults.[62] However, the latter trial[62] has high risk of bias due to an imbalance in the patient attrition between the study groups (section 5.1.1). 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 (Table 9). Overall, the findings from the statistically powered RCTs indicate that 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

consumption was significantly lower in the BIS-guided group[45;49] or 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 entropy group.[54] In contrast to 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 entropy-guided anaesthesia[55;141] and Narcotrend-guided anaesthesia, [59;60;63], while 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 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 an RCT by Kerssens and colleagues,[49] and for total intravenous anaesthesia (propofol) using data from an 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 post-anaesthesia care unit (PACU) (5 RCTs); duration of stay in the PACU (2 RCTs); time to discharge from the PACU (5 RCTs) ; time to response to commands (3 RCTs) time to recovery of orientation (3 RCTs); time to first movement response (2 RCTs) time to recovery based on recovery scores (2 RCTs); time to spontaneous breathing (1 RCT); time to laryngeal mask airway removal (1 RCT); and time to phonation (1 RCT). “PACU stay” was an outcome in the Cochrane review[34] but does not appear to distinguish between PACU admission, 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 that were not specifically powered for this outcome, two detected a significant difference between the study groups in time to eye opening and five did not. In the four RCTs 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, entropy, Narcotrend) or type of anaesthesia (volatile, total intravenous, or mixed) could be explanatory variables. It is unclear whether 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 due to the authors' conflict of interests[57] (section 5.1.1); 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 six minutes, with the largest reductions being for Narcotrend-guided total intravenous 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 six minutes (the best achieved) in the time to extubation would have importance for patients or for service provision, given that it represents less than 10% of the total time patients were undergoing surgical procedures.

Outcomes related to PACU stay

None of the RCTs that assessed outcomes related to PACU stay were 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 total intravenous anaesthesia in adults (5.8 minutes),[63] BIS-guided mixed anaesthesia in adults (4.7 minutes),[48] and 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 to 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] whilst 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 these trials supplemented their BIS-guided anaesthesia with monitoring of clinical signs which makes it difficult to determine whether 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 general anaesthesia. These reductions in discharge times are relatively large compared to 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 entropy-guided anaesthesia[57] and Narcotrend-guided anaesthesia[59] compared with standard clinical practice. Both these trials were on adults receiving total intravenous anaesthesia. The third RCT, on children receiving total intravenous anaesthesia, 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 entropy-guided[54;57] or Narcotrend-guided[59] anaesthesia compared to standard clinical practice. The reported time savings were 4.8 minutes (median) in entropy-

guided total intravenous anaesthesia in adults,[57] 5.1 minutes (mean) in entropy-guided volatile anaesthesia in children,[54] and 5.6 minutes (mean) in Narcotrend-guided total intravenous anaesthesia 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 the RCTs measuring this outcome evaluated 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 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 total intravenous anaesthesia in adults and found a significantly shorter time to spontaneous breathing in the 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 post-operative nausea and vomiting (PONV) (4 RCTs), post-operative pain (2 RCTs), post-operative cognitive dysfunction in elderly patients (1 RCT) and emergence delirium in children (1 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.

Post-operative nausea and vomiting. The four RCTs reporting this outcome evaluated BIS-guided volatile anaesthesia in children,[51] entropy-guided total intravenous anaesthesia in adults,[55] and Narcotrend-guided total intravenous anaesthesia 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 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 visual analogue scale (no details provided) rather than frequency of occurrence, and found significantly higher (better) scores (indicating less frequent PONV) in the Narcotrend group compared to 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.

Post-operative pain. The two RCTs that assessed post-operative pain evaluated entropy-guided anaesthesia, either in adults under total intravenous anaesthesia,[55] or in children under mixed anaesthesia.[56] 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 entropy group than standard clinical monitoring for the adult population.[55] In the paediatric population, the CHEOPS scores were significantly lower in the entropy group at 60, 90 and 120 minutes after arrival in the PACU but not at 30 minutes after arrival.[56]

Post-operative cognitive dysfunction. The RCT that assessed this outcome evaluated BIS-guided intravenous anaesthesia in elderly patients.[47] At one week post-surgery the incidence of POCD was 32.5% in the entropy group and 39.1% in the standard clinical monitoring group. At three months post-surgery the incidences were 8.1% and 12.0% respectively. Only the three 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 Pediatric Anesthetic Emergence Delirium (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.

7.1.2 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[98] using data drawn from a prospective study by Ekman and colleagues[99] and from the RCTs 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 , while 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 anesthesia monitoring was estimated as $4,410 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 due to 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 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. Cost effectiveness was assessed

using incremental cost effectiveness ratios 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.

BIS compared with standard clinical monitoring

We presented a base case analysis for two modes of anaesthetic administration (total intravenous anaesthesia and mixed anaesthesia (induction with IV anaesthesia and maintenance with inhaled anaesthesia or a combination of inhaled and IV 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 general anaesthesia with TIVA we used the odds ratio of awareness with BIS monitoring (0.24), compared with standard clinical monitoring, reported in the meta-analysis in our systematic review of patient outcomes (section 5.1) 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 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 odds ratio, 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 (section 5.1). 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 under-dose 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 awareness, undergoing general anaesthesia 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 6 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 (90% of additional cost, per patient). By reducing the incidence of awareness and longer-term effects of post-operative cognitive dysfunction BIS monitoring was associated with improved outcomes. The ICER, for BIS compared with standard clinical monitoring in this population was £27,345. Deterministic sensitivity analyses indicated 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 general anaesthesia 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 4 cases of LPS (from 5.2 to 1.3), which included a reduction of 2 cases of PTSD (from 2.8 to 0.7). While 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 (£44,702) despite the lower incremental cost estimated for this population, arising from reduced anaesthetic consumption. Deterministic sensitivity analyses indicated the ICER was sensitive to the same input parameters as for the population at high risk of awareness. In all 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 general anaesthesia (induction with IV and maintenance including inhaled anaesthetic) we used the pooled odds ratio 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 the odds ratio 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 £36,126. Deterministic sensitivity analyses indicated 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 odds ratio of awareness with BIS vs standard clinical care), highest probability of LPS (0.48), longest duration of PTSD (21.6 years), highest probability of PTSD (0.239) and lowest cost of sensors (£13.3, 75% of based 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 general anaesthesia, were the same as for TIVA (16, 5.2 and 2.8 per 10,000 patients, respectively), while 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 (£16.23 compared with £14.20). 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 (£61,869). Deterministic sensitivity analyses indicated 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.

Entropy compared with standard clinical monitoring

A base case analysis was presented for two modes of anaesthetic administration (total intravenous anesthesia and mixed anaesthesia (induction with IV anaesthesia and maintenance with inhaled anaesthesia or a combination of inhaled and IV 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 Entropy on the incidence of intraoperative awareness or on post-operative cognitive dysfunction. In the absence of evidence specific to Entropy we have applied the effectiveness estimates derived for BIS, described above. This meant that the modelled clinical effectiveness of Entropy was identical to that reported for BIS – this is an untested assumption and must be considered a weakness in the evidence base for 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 general anaesthesia with TIVA, the modelled cost per patient with 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 Entropy monitoring was approximately half that of BIS monitoring, with the majority being attributable to the sensors attached to the patient (80% of additional cost per patient). Entropy monitoring was associated with improved outcomes, based on applying clinical effectiveness evidence reported for BIS. The ICER for 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 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 Entropy on PONV and to assumptions regarding patient throughput (except at comparatively low volumes, below 500 cases per year per module).

In the population of general surgical patients, undergoing general anaesthesia with TIVA, 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 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 (£3,131 - £31,430). Deterministic sensitivity analyses indicated 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 ICERS below £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 Entropy we have applied the effectiveness estimates derived for BIS in this analysis. For patients undergoing mixed general anaesthesia (induction with IV and maintenance including inhaled anaesthetic) the pooled odds ratio 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 general anaesthesia, 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 Entropy compared with standard clinical monitoring in this population was £19,367. Deterministic sensitivity analyses indicated 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 odds ratio of awareness with monitoring vs standard clinical care) and greater probability of LPS (0.48). The majority of the ICERs remained below £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 general anaesthesia, Entropy monitoring had higher costs than standard clinical monitoring. In contrast to 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 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 Entropy monitoring than for high risk patients, the lower incremental cost resulted in an equivalent ICER (£19,000). Deterministic sensitivity analyses indicated 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 odds ratio 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 (total intravenous anesthesia and mixed anaesthesia (induction with IV anaesthesia and maintenance with inhaled anaesthesia or a combination of inhaled and IV 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 post-operative cognitive dysfunction. 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 general anaesthesia 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 Entropy monitoring, and approximately a quarter that of BIS – primarily due to differences in the cost of the sensors attached to the patient. In contrast to BIS and 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 £5,681. 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 Entropy on PONV.

In the general surgical population, undergoing general anaesthesia 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 general anaesthesia (induction with IV and maintenance including inhaled anaesthetic) the pooled odds ratio 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 general anaesthesia, 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 £8,033. Deterministic sensitivity analyses indicated 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 odds ratio 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 general anaesthesia, Narcotrend monitoring had higher costs than standard clinical monitoring. While the proportionate reduction in consumption of inhaled anaesthetic (desflurane) was lower than the reduction in IV 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, since 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 £2,534. Conclusions from the scenario analyses were similar to those undertaken for high risk patients.

7.2 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 only know of 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)[106] 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, post-operative nausea and vomiting, 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 post-operative recovery, not limited to ambulatory anaesthesia, was conducted by Punjasawadwong and colleagues (2007).[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 intravenous anaesthetic), although for anaesthetic consumption the precision of the existing effect estimates was not necessarily improved, due to 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 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-, 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 (section 7.3). 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 incremental cost effectiveness ratios 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 QALYs 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 QALYs 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 quality of life impact of those conditions).

In the model general anaesthesia 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 post-operative cognitive dysfunction, which have consequences for quality of life. 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 ICERs 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 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 Entropy and Narcotrend. These are untested assumptions and must be considered a weakness in the cost effectiveness evidence base. While 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 Entropy and Narcotrend – again this is an untested assumption. Evidence on the incidence of post-operative cognitive dysfunction 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 post-operative complication). The best evidence we could identify that reported post-operative cognitive dysfunction in patients who had been assessed pre-operatively compared with a matched group of non-surgical controls is over ten years old and it is not clear whether this will reflect incidence of POCD in current practice.

While we were able to identify some evidence on the incidence of PTSD in patients who experienced awareness during general anaesthesia, we did not identify any studies reporting overall quality of life 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 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 over-dose or under-dose 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 under-dose, whereas POCD and PONV maybe more closely associated with over-dose. The potential for savings in terms of anaesthetic drug use may also primarily arise in this latter group. While 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.

7.3 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 post-operative cognitive dysfunction (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 only available 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 2011[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 due to the technology being considered (BIS, 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. Also the Narcotrend-Compact M monitor which used NarcoWin software seems to have been upgraded to the Narcotrend NI Module.

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. For example a range of BIS versions have been used in the studies (including 3.1, 3.12U, 3.2, 3.21 3.22, 3.3, 3.4) which may have produced different results. It is not clear what alterations have been made to the algorithms and how these impact the trial results as the algorithms are proprietary and not completely published. In the Narcotrend industry submission to NICE two trials are discussed which show 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-state entropy difference was also used as a target value in some trials. Narcotrend values ranged from D_0 to C_1, D_2– E_0 adjusted to $D_0 – D_1$, and $D_2 – E_0$, 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. Ellerkman 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 trial 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 intravenous 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 generalizability 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.

8. 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 to 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 not associated with a statistically significant reduction in intraoperative awareness in patients classified as at higher risk, though there is uncertainty in effect estimates due to significant heterogeneity. Caution is advised due to 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, though 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

Our systematic review of patient outcomes found considerably more RCTs had investigated the effects of titrating depth of anaesthesia according to BIS than according to E-entropy or Narcotrend. Our literature searches identified three on-going RCTs that would meet the inclusion criteria of our systematic review (Appendix 12), all of which are investigating anaesthesia depth titrated according to BIS values. A further trial which is similar to our inclusion criteria is worth mentioning. The Michigan Awareness Control Study is an RCT comparing BIS-guided and MAC-guided electronic alerts for the prevention of awareness under general anaesthesia.[142] In one arm of the study, anaesthesia providers will receive an electronic page if the BIS value is >60. In the other arm of the study, anaesthesia providers will receive a page if the age-adjusted MAC is <0.5. The target sample size is 30,000 patients (aged >18 years) at both low and high risk for awareness. The primary outcome measure will be intraoperative awareness with explicit recall measured at 28-30 days post-anasethesia. A secondary outcome will be incidence of PTSD in patients with definite or possible awareness. A meta-analysis combining the results of the study with the results of the B-Unaware[27] trial and the BAG-RECALL trial[44] (included in our review) will also be conducted. The completion date was June 2010, and it is not known when the results of the trial will be published. This will therefore be the largest statistically powered RCT of EEG-guided general anaesthesia to detect awareness, and the results may reduce some of the uncertainty regarding the effects of BIS monitoring on awareness in the existing evidence base.

There is a need for further evaluation of the utility of E-entropy-guided and Narcotrend-guided depth of anaesthesia in RCTs, with adequate sample sizes. The data from current E-entropy and Narcotrend RCTs are inadequate to statistically pool quantitative effect estimates for relevant outcomes and population subgroups. In particular, there were no trials of the use of Narcotrend in children, and only two paediatric studies of E-entropy.

As demonstrated in our systematic review of patient outcomes, there are numerous prognostic variables that could influence the effectiveness of EEG-titrated depth of anaesthesia. These include (inter alia) the study population, type of surgery, and type of anaesthesia used. Currently-available RCTs provide little insight into the importance of these variables. Future RCTs should, where possible, be stratified to enable the effects of these variables to be explored.

Any further RCTs that are conducted to investigate the effects of EEG-titrated monitoring of depth of anaesthesia on the incidence of intraoperative awareness should incorporate adequate length of follow up to detect delayed cases of awareness. Cases of awareness may emerge after the first post-operative week, but in nearly all of the currently available RCTs of BIS, 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 RCTs should be strongly encouraged. Future studies should also evaluate the effects of anaesthesia over-dosing, including short-term effects such as nausea and vomiting, as well as longer-term impact on cognitive function.

9. REFERENCES

(1) Jones JG, Konieczko K. Hearing and memory in anaesthetised patients. *BMJ* 1986; 292:1291.

(2) Kerssens C, Klein J, Bonke B. Awareness: Monitoring versus remembering what happened. *Anesthesiology* 2003; 99(3):570-575.

(3) Enlund M, Hassan HG. Intraoperative awareness: detected by the structured Brice interview? *Acta Anaesthesiologica Scandinavica* 2002; 46:345-349.

(4) Russell IF. Comparison of wakefulness with two anaesthetic regimens. *British Journal of Anaesthesia* 1986; 58:965-968.

(5) Moerman N, Bonke B, Oosting J. Awareness and recall during general anaesthesia. *Anesthesiology* 1993; 79:454-464.

(6) Ranta S, Laurila R, Saario J, Ali-Melkkila T, Hynynen M. Awareness with recall during general anaesthetia: incidence and risk factors. *Anesthesia and Analgesia* 1998; 86:1084-1089.

(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. *Anesthesia and Analgesia* 2009; 108(2):527-535.

(8) Leslie K, Chan MTV, Myles PS, Forbes A, McCulloch TJ. Posttraumatic stress disorder in aware patients from the B-Aware trial. *Anesthesia and Analgesia* 2010; 110(3):823-828.

(9) Sandin RH, Enlund G, Samuelsson P, Lennmarken C. Awareness during anaesthesia: a propsective case study. *Lancet* 2000; 355:707-711.

(10) Osterman JE, Hopper J, Heran WJ, Keane TM, van der Volk BA. Awareness under anaesthesia and the development of posttraumatic stress disorder. *General Hospital Psychiatry* 2001; 23:198-204.

(11) Lennmarken C, Bildfors K, Enlund G, Samuelsson P, Sandin R. Victims of awareness. *Acta Anaesthesiologica Scandinavica* 2002; 46:229-231.

(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. *British Journal of Anaesthesia* 2000; 84(1):6-10.

(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. *Anesthesia and Analgesia* 2004; 99:833-839.

(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(2):269-274.

(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 Anaesthesiologica Scandinavica* 2009; 53:873-882.

(16) Leslie K, Davidson AJ. Awareness during anesthesia: a problem without solutions?. [Review]. *Minerva Anestesiologica* 2010; 76(8):624-628.

(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. *Interactive CardioVascular and Thoracic Surgery* 2006; 5(3):207-216.

(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. *British Journal of Anaesthesia* 2008; 101(2):178-185.

(19) Domino KB, Posner KL, Caplan RA, Cheney FW. Awareness during anaesthesia. A closed claims analysis. *Anesthesiology* 1999; 90:1053-1061.

(20) Hardman JG, Aitkenhead AR. Awareness during anesthesia. *Continuing Education in Anaesthesia, Critical Care and Pain* 2005; 5(6):183-186.

(21) Thomas G, Morgan M. Monitoring neuromuscular blockade and depth of anaesthesia. *Anaesthesia and Intensive Care Medicine* 2011; 12:271-274.

(22) Orser BA, Mazer D, Baker AJ. Awareness during anesthesia. *Canadian Medical Association Journal* 2008; 178(2):185-188.

(23) Bruhn J, Myles PS, Sneyd R, Struys MMRF. Depth of anaesthesia monitoring: What's available, what's validated and what's next? *British Journal of Anaesthesia* 2006; 97(1):85-94.

(24) Brice DD, Hetherington RR, Utting JE. A simple study of awareness and dreaming during anaesthesia. *British Journal of Anaesthesia* 1970; 42:535-541.

(25) Wennervita J, Ranta SO, Hynynen M. Awareness and recall in outpatient anesthesia. *Anesthesia and Analgesia* 2002; 95(1):72-77.

(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 Anesthesiology* 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(11):1097-1108.

(28) Eger EI, Sonner JM. Response to Enlund M: TIVA, awareness and the Brice interview. *Anesthesia and Analgesia* 2006; 102:965.

(29) Lubke GH, Kerssens C, Gershon RY, Sebel PS. Memory formation during general anesthesia for emergency cesarean sections. *Anesthesiology* 2000; 92(4):1029-1034.

(30) Hudson AE, Hemmings HC. Are anaesthetics toxic to the brain? *British Journal of Anaesthesia* 2011.

(31) The Association of Anaesthetists of Great Britain and Ireland. Recommendations for Standards of Monitoring During Anaesthesia and Recovery. 4th Edition. 2007. London, The Association of Anaesthetists of Great Britain and Ireland.

(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(9):850-854.

(33) Joint Formulary Committee. British National Formulary 62. London: British Medical Association and Royal Pharmaceutical Society of Great Britain; 2011.

(34) Punjasawadwong Y, Phongchiewboon A, Bunchungmongkol N. Bispectral index for improving anaesthetic delivery and postoperative recovery. *Cochrane Database of Systematic Reviews* 2007; Issue 4.

(35) Higgins JPT, Altman DG, G, Juni P, Moher D, Oxman AD et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011; 343.

(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, et al. Good practice guidelines for decision-analytic modelling in health technology assessment: A review and consolidation of quality assessment. *Pharmacoeconomics* 24[4], 355-371. 2006.

(38) National Institue for Health and Clinical Excellence (NICE). Guide to the methods of technology appraisal. 2008. 4-7-2011.

(39) Treasury HM. The Green Book. Appraisal and Evaluation in Central Government. http://www.hm-treasury.gov.uk/data_greenbook_index.htm . 2010. 1-4-2009.

(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. *Chinese Medical Journal* 2011; 124(22):3664-3669.

(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. *Journal of Anaesthesiology Clinical Pharmacology* 2011; 27(4):500-505.

(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. *Journal of Clinical Anesthesia* 2011; 23(8):636-642.

(43) Hampshire P, Guha A, Welters I, Murphy J, Poole L. Bispectral index monitoring reduces sedative and vasopressor requirements during percutaneous tracheostomy. *Critical Care* 2011; 15:S123-S124.

(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 England Journal of Medicine* 2011; 365(7):591-600.

(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. *Anesthesia & Analgesia* 2001; 92(4):877-881.

(46) Bhardwaj N, Yaddanapudi S. A randomized trial of propofol consumption and recovery profile with BIS-guided anesthesia compared to standard practice in children. *Paediatric Anaesthesia* 2010; 20(2):160-167.

(47) Chan MTV, Cheng B, Gin T, Lee T, Koo E. Cognitive dysfunction after anesthesia: a randomized controlled trial [abstract]. *Journal of Neurosurgical Anesthesiology* 2010; 22(4):408-409.

(48) Kamal NM, Omar SH, Radwan KG, Youssef A. Bispectral index monitoring tailors clinical anesthetic delivery and reduces anesthetic drug consumption. *Journal of Medical Sciences* 2009; 9(1):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(3):518-524.

(50) Leslie K, Myles PS, Forbes A, Chan MT. The effect of bispectral index monitoring on long-term survival in the B-aware trial. *Anesthesia & Analgesia* 2010; 110(3):816-822.

(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. *Journal of the Chinese Medical Association* 2011; 74(1):28-36.

(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. *Pediatric Dentistry* 2004; 26(3):256-260.

(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 Dentistry* 2005; 27(6):500-504.

(54) Choi SR, Lim YH, Lee SC, Lee JH, Chung CJ. Spectral entropy monitoring allowed lower sevoflurane concentration and faster recovery in children. *Acta Anaesthesiologica Scandinavica* 2010; 54(7):859-862.

(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(12):1224-1229.

(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. *Paediatric Anaesthesia* 2010; 20(12):1105-1110.

(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(2):274-279.

(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 Anaesthesiologica Taiwanica*: Official Journal of the Taiwan Society of Anesthesiologists 2008; 46(3):106-111.

(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. *Chinese Journal of Cancer* 2010; 29(1):110-113.

(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? *British Journal of Anaesthesia* 2007; 99(5):686-693.

(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? *Anesthesia & Analgesia* 2006; 103(6):1469-1477.

(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. *Anaesthesia & Intensive Care* 2010; 38(1):159-166.

(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(1):34-41.

(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. *Anesthesia & Analgesia* 2005; 101(2):427-434.

(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(4):849-852.

(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(6):264-269.

(67) Basar H, Ozcan S, Buyukkocak U, Akpinar S, Apan A. Effect of bispectral index monitoring on sevoflurane consumption. *Eur J Anaesthesiol* 2003; 20(5):396-400.

(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):1-4.

(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(1):63-69.

(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(3):342-347.

(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(4):808-815.

(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(8):364-369.

(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(4):819-830.

(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(4):443-451.

(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(4):394-399.

(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. *Technology, Computing, and Simulation* 2007; 104(5):1145-1149.

(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(8):862-868.

(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(2):105-110.

(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(9423):1757-1763.

(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(5):1165-1169.

(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(6):435-439.

(82) Puri GD, Murthy SS. Bispectral index monitoring in patients undergoing cardiac surgery under cardiopulmonary bypass. *Eur J Anaesthesiol* 2003; 20(6):451-456.

(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(6):1667-1674.

(84) Song D, Joshi GP, White PF. Titration of volatile anesthetics using bispectral index facilitates recovery after ambulatory anesthesia. *Anesthesiology* 1997; 87(4):842-848.

(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(1):6-17.

(86) Tufano R, Palomba R, Lambiase G, Giurleo LG. [The utility of bispectral index monitoring in general anesthesia]. *Minerva Anestesiol* 2000; 66(5):389-393.

(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(4):811-817.

(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(1):13-18.

(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(1):20-25.

(90) Assare H, Anderson RE, Jakobsson J. Sevoflurane requirements during ambulatory surgery: a clinical study of bispectral index and auditory evoked potential guided anaesthesia. *Ambulatory Surgery* 2002; 9:207-211.

(91) Luginbuhl M, Wuthrich S, Petersen-Felix S, Zbinden AM, Schnider TW. Different benefit of bispectral index (BIS) in desflurane and propofol anesthesia. *Acta Anaesthesiologica Scandinavica* 2003; 47:165-173.

(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. *Paediatric Anaesthesia* 2010; 20(1):56-62.

(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 Anaesthesiologica Scandinavica* 2008; 52(8):1038-1045.

(94) Ministry of Health and Long-Term Care.Medical Advisory Secretariat. Bispectral index monitor. 2004.

(95) Han DW, Linares-Perdomo O-J, Lee JS, Kim JH, Kern SE. Comparison between cerebral state index and bispectral index as measures of electroencephalographic effects of sevoflurane using combined sigmoidal e max model. *Acta Pharmacologica Sinica* 2011; 32(10):1208-1214.

(96) 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. *British Journal of Clinical Pharmacology* 2011; 71(6):871-885.

(97) Nordström O, Engström AM, Persson S, Sandin R. Incidence of awareness in total i.v. anaesthesia based on propofol, alfentanil and neuromuscular blockade. *Acta Anaesthesiologica Scandinavica* 1997; 41:978-984.

(98) Abenstein JP. Is BIS monitoring cost-effective? Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine & Biology Society 2009; 2009:7041-7044.

(99) Ekman A, Lindholm ML, Lennmarken C, Sandin R. Reduction in the incidence of awareness using BIS monitoring. *Acta Anaesthesiol Scand* 48, 20-26. 2004.

(100) 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* 6[30]. 2002.

(101) Chernin EL. Pharmacoeconomics of Inhaled Anesthetic Agents: Considerations for the Pharmacist. *American Journal of Health System Pharmacy* 2004; 61(Suppl 4):S18-S22.

(102) Dion P. The cost of anaesthetic vapours. *Canadian Journal of Anaesthesia* 1992; 39:633.

(103) Chatterjee S, Rudra A, Sengupta S. Current Concepts in the Management of Postoperative Nausea and Vomiting. *Anesthesiology Research and Practice* 2011.

(104) Ku CM, Ong BC. Postoperative Nausea and Vomiting: a Review of Current Literature. *Singapore Medical Journal* 2003; 44(7):366-374.

(105) Gan TJ. Postoperative nausea and vomiting - Can it be eliminated? *Journal of the American Medical Association* 2002; 287(10):1233-1236.

(106) Liu SS. Effects of Bispectral Index Monitoring on Ambulatory Anesthesia. *Anesthesiology* 2004; 101:311-315.

(107) Deiiner S, Silverstein JH. Postoperative delirium and cognitive function. *British Journal of Anaesthetics* 2009; 103(Suppl 1):i41-i46.

(108) Moller JT, Cluitmans P, Rasmussen LS, Canet J, Rabbitt P, Jolles J et al. Long term postoperative cognitive dysfunction in the elderly: ISPOCD1 study. *The Lancet* 2012; 351:857-861.

(109) 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; 44:1246-1251.

(110) The Information Centre for Health and Social Care. HES online - main procedures and intervention tables. http://www.hesonline.nhs.uk/ . 2012.

(111) 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(1):49-55.

(112) Lyons G, Macdonald R. Awareness during Caesarean section. *Anaesthesia* 1991; 46(1):62-64.

(113) Ranta S, Jussila J, Hynynen M. Recall of awareness during cardiac anaesthesia: influence of feedback information to the anaesthesiologist. *Acta Anaesthesiologica Scandinavica* 1996; 40(5):554-560.

(114) Liu WHD, Thorp TAS, Graham SG, Aitkenhead AR. Incidence of awareness with recall during general anaesthesia. *Anaesthesia* 1991; 46(6):435-437.

(115) Schwender D, Kunze-Kronawitter H, Dietrich P, Klasing S, Forst H, Madler C. Conscious awareness during general anaesthesia: patients' perceptions, emotions, cognition and reactions. *British Journal of Anaesthesia* 1998; 80:133-139.

(116) Samuelsson P, Brudin L, Sandin RH. Late psychological symptoms after awareness among consecutively included surgical patients. *Anesthesiology* 2007; 106:26-32.

(117) 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. 184-192.

(118) American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th Edition ed. Washington DC: American Psychiatric Association; 1994.

(119) Kessler RC, Sonnega A, Bromet E, Hughes M, Neslon CB. Posttraumatic Stress Disorder in the National Comorbidity Survey. *Archives of General Psychiatry* 1995; 52(12):1048-1060.

(120) Doctor JN, Zoellner LA, Feeny NC. Predictors of health-related quality-of-life utilities among persons with posttraumatic stress disorder. *Psychiatric Services* 2011; 62(3):272-277.

(121) 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. *Psychiatric Services* 2009; 60(9):1230-1238.

(122) Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics* 21, 271-292. 2002.

(123) Laffaye C, Kennedy C, Stein MB. Post-traumatic stress disorder and health-related quality of life in female victims of intimate partner violence. *Violence & Victims* 2003; 18(2):227-238.

(124) 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. *Journal of Psychiatric Research* 2007; 41(1-2):90-96.

(125) Meeske KA, Ruccione K, Globe DR, Stuber ML. Posttraumatic stress, quality of life, and psychological distress in young adult survivors of childhood cancer. *Oncology Nursing Forum* 2001; 28(3):481-489.

(126) 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. *Journal of Clinical Monitoring & Computing* 2007; 21(5):295-302.

(127) Weil G, Passot S, Servin F, Billard V. Does spectral entropy reflect the response to intubation or incision during propofol-remifentanil anesthesia? *Anesthesia & Analgesia* 2008; 106(1):152-159.

(128) Magruder KM, Frueh B, Knapp R, et al. Prevalence of posttraumatic stress disorder in Veteran's Affairs primary care clinics. *General Hospital Psychiatry* 27, 169-179. 2005.

(129) Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics* 21, 271-292. 2002.

(130) 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 in Health* 12[2], 346-353. 2009.

(131) 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. *Journal of Traumatic Stress* 2007; 20(4):637-642.

(132) Shiner B, Watts BV, Pomerantz A, Young-Xu Y, Schnurr PP. Sensitivity of the SF-36 to PTSD symptom change in veterans. *Journal of Traumatic Stress* 2011; 24(1):111-115.

(133) 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 Research* 2011; 186(1):85-90.

(134) SCHARR. SF-6D. http://www.shef.ac.uk/scharr/sections/heds/mvh/sf-6d . 2012.

(135) National Institue for Health and Clinical Excellence (NICE). Post-traumatic stress disorder (PTSD). The management of PTSD in adults and children in primary and secondary care. National costing report. http://guidance.nice.org.uk/CG26/CostingReport/pdf/English . 2005. 1-2-2012.

(136) National Institue for Health and Clinical Excellence (NICE). Post-traumatic stress disorder (PTSD). The management of PTSD in adults and children in primary and secondary care. Clinical Guideline 26. http://www.nice.org.uk/CG26 . 2005. 1-2-0012.

(137) Singleton N, Bumpstead R, O'Brien M, Lee A, Meltzer H. Psychiatric morbidity among adults living in private households, 2000. http://www.ons.gov.uk/ons/index.html . 2001. London, HMSO. 1-2-2012.

(138) 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. *Medical Care* 2008; 46(4):388-393.

(139) Department of Health. NHS Reference Costs 2010/2011. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_131140 . 2012.

(140) Phillips AA, McLean RF, Devitt JH, Harrington EM. Recall of intraoperative events after general anaesthesia and cardiopulmonary bypass. *Canadian Journal of Anaesthesia* 1993; 40:922-926.

(141) Russell IF. The Narcotrend 'depth of anaesthesia' monitor cannot reliably detect consciousness during general anaesthesia: an investigation using the isolated forearm technique. *British Journal of Anaesthesia* 2006; 96(3):346-352.

(142) 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 Anesthesiology* 2009; 9(pp 7).

(143) Khosravi S, Hahn JO, Dosani M, Dumont GA, Ansermino JM. Monitor-decoupled pharmacodymamics of propofol in children using State Entropy as the clinical end point. Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine & Biology Society 2010; 2010:544-547.

(144) Neto AM, Santos JC, Santos EC. Molecular dynamics study of the flagella inside the carbon torus. *Journal of Nanoscience & Nanotechnology* 2011; 11(2):1230-1236.

(145) Baulig W, Seifert B, Schmid ER, Schwarz U. Comparison of spectral entropy and bispectral index electroencephalography in coronary artery bypass graft surgery. *Journal of Cardiothoracic & Vascular Anesthesia* 2010; 24(4):544-549.

(146) Bispectral index monitor: an evidence-based analysis. 2004.

(147) El Menesy TM, Hamden GA, Amin SM, Zakaria D. Continuous intraoperative high state and response entropies reduced desflurane emergence agitation in pediatric laparoscopic surgery. *Egyptian Journal of Anaesthesia* 2008; 24(2):85-93.

(148) Pelletier N, Passot S, Zufferey P, Molliex S. Benefit of EEG Guided Anaesthesia: A Meta-Analysis. *Anesthesiology* 2011; 115.

(149) 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(4):216-220.

(150) 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(5):1096-1104.

(151) Riad W, Schreiber M, Saeed AB. Monitoring with EEG entropy decreases propofol requirement and maintains cardiovascular stability during induction of anaesthesia in elderly patients. *European Journal of Anaesthesiology* 2007; 24(8):684-688.

(152) Singh M, Kulkarni D, Gopinath R. Bispectral index monitoring may not enhance performance of the "wake-up" test during scoliosis surgery. *Journal of Anaesthesiology Clinical Pharmacology* 2009; 25(2):179-182.

(153) 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. *European Journal of Anaesthesiology* 2005; 22(10):741-747.

(154) 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).

10. APPENDICES

Appendix 1: Report methods for synthesis of evidence of clinical 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 Centre for Reviews & Dissemination (CRD) guidance for undertaking reviews in health care.

Population

The relevant study population for this assessment is patients receiving general anaesthesia for surgery, including adults and children in whom the technology is licensed. Elderly and obese patients undergoing general anaesthesia 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
- Bispectral Index (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, anti-sickness drugs and muscle relaxants.
- Health related quality of life (HRQoL)

Data on these indirect outcomes are likely to be used to estimate Quality-Adjusted Life Years (QALYs) as final health outcomes.

Study design

We will prioritise RCTs for inclusion in the systematic review of clinical effectiveness (see Section 4.8). Where RCTs of technologies are not identified we will consider non-randomised controlled trials and controlled observational studies for inclusion, providing they include relevant outcomes as specified in Section 5.4.

Systematic reviews will only be retrieved in order 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.[143] 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 (Ovid); Embase (Ovid); the Cochrane Library (Cochrane Database of Systematic Reviews; Cochrane Central Register of Controlled Trials), the Database of Abstracts of Reviews of Effectiveness (DARE); Health Technology Assessment Database (HTA); NHS Economic Evaluation Database (NHS EED); 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[143] – see Section 5.5). 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; Clinical Trials.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 (see Section 6) 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 which 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.

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 in section 5 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 as described in section 5.2, 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 et al.[144] and Philips et al.[145] 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 health related quality of life (HRQoL) or health state utility for adults with episodes of intraoperative awareness. Where available, quality of life data will be used in our economic model (see section 6.2).

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[145] 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 (PSA) will be undertaken if both the data and modelling approach permit this. The outputs of any PSA 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.

Appendix 2: Literature search strategies

Medline search strategy for BIS, Narcotrend and 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 Entropy, hence the inclusion of BIS terms at the end of the strategy (from line 65 onwards)

Medline search strategy for BIS, Narcotrend and 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/ (13101)
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/ (46626)

12 exp Anesthetics, General/ (98559)
13 (an?esthetic* or an?esthesia or an?esthetist*).tw. (184909)
14 Intraoperative Period/ (11282)
15 Anesthesia, Intravenous/ (9798)
16 Anesthetics, Inhalation/ (9572)
17 Anesthesiology/ (15249)
18 exp Infusions, Intravenous/ (44602)
19 Surgical Procedures, Operative/ (48143)
20 General Surgery/ (31636)
21 (surgery or surgical).tw. (1018003)
22 Perioperative Period/ (254)
23 Signal Processing, Computer-Assisted/ (30306)
24 Intraoperative Complications/ (23721)
25 Perioperative Care/ (6700)
26 Monitoring, Physiologic/ (41597)
27 Adjuvants, Anesthesia/ (2653)
28 Electromyography/ (62736)
29 exp Electroencephalography/ (113311)
30 Mental Recall/ (25043)
31 Wakefulness/ (13087)
32 Consciousness/ (8829)
33 Perception/ (17362)
34 Intraoperative Awareness/ or Awareness/ (12290)
35 Arousal/ (26845)
36 Deep Sedation/ (309)
37 Conscious Sedation/ (5918)
38 Drug Therapy, Computer-Assisted/ (1263)
39 Pain Measurement/ (50337)
40 cerebral cortex/de (15104)
41 Evoked Potentials/ or Evoked Potentials Auditory/ (57136)
42 Signal Processing, Computer-Assisted/ (30306)
43 (surgery or surgical or operating or operation*1).tw. (1240833)
44 (intraoperative* or "intra-operative*" or "intra operative*").tw. (73745)
45 (perioperative* or "peri-operative*" or "peri operative*").tw. (45446)
46 "depth of anaesthesia monitor*".tw. (39)
47 "depth of anesthesia monitor*".tw. (31)
48 "Anesthesia and Analgesia"/ (3320)
49 Postoperative Period/ (30192)
50 (postoperative or post?operative).tw. (257047)
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. (767912)
52 (EEG or EMG or FEMG or encephalogra* or electroencephalogra* or electromyogra*).tw. (103627)
53 Brice.tw. (18)
54 or/6-53 (2633781)
55 5 and 54 (326)
56 limit 55 to (english language and yr="1995 -Current") (277)
57 animals/ (4924118)
58 56 not 57 (259)
59 (letter or comment or editorial).pt. (1097745)
60 58 not 59 (240)
61 crystal*.tw. (146923)
62 60 not 61 (229)
63 coma/ or coma.tw. (25605)
64 62 not 63 (228)

65 exp economics/ (449064)
66 exp economics hospital/ (17691)
67 exp economics pharmaceutical/ (2299)
68 exp economics nursing/ (3854)
69 exp economics medical/ (13581)
70 exp "Costs and Cost Analysis"/ (161041)
71 Cost Benefit Analysis/ (52655)
72 exp models economic/ (8329)
73 exp fees/ and charges/ (7794)
74 exp budgets/ (11145)
75 (economic* or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic*).tw. (350335)
76 (value adj1 money).tw. (20)
77 budget$.tw. (14911)
78 or/65-77 (681466)
79 ((energy or oxygen) adj cost).tw. (2386)
80 (metabolic adj cost).tw. (626)
81 ((energy or oxygen) adj expenditure).tw. (13708)
82 or/79-81 (16090)
83 78 not 82 (677823)
84 (letter or editorial or comment or historical article).pt. (1367063)
85 83 not 84 (624009)
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 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: Adults and children aged over 2 years receiving general anaesthesia for surgery. Not included: • 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 will not be included.	Yes ↓ next question	Unclear ↓ next question	No → EXCLUDE1
Technology: Any of the following: • E-Entropy[b] • Bispectral Index (BIS) • Narcotrend **Comparators:** Standard clinical observation[c], including one or more of the following markers: • end-tidal anaesthetic gas concentrations / minimum alveolar concentration (for inhaled anaesthesia) • heart rhythm • blood pressure • oxygen levels (pulse oximeter) • lacrimation • sweating	Yes ↓ next question	Unclear ↓ next question	No → EXCLUDE2
Outcomes: One or more of the following: • Probability of intraoperative awareness • 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 anti-sickness drugs and muscle relaxants. • Health related quality of life (HRQoL)	Yes ↓ next question	Unclear ↓ next question	No → EXCLUDE3
Study design: RCT; quasi-randomised or non-randomised controlled trial; controlled before and after study [d]	Yes ↓ next question	Unclear ↓ next question	No → EXCLUDE4

Systematic reviews to be retrieved for reference checking only 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 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", "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-RCTs for a technology if no RCTs have already been identified.

Appendix 4: Reasons for the exclusion of full-text publications from systematic review of patient outcomes

Of 31 full-text publications that were screened against the systematic review eligibility criteria, 10 were excluded for the following reasons:

Exclusion criterion = study design (5 publications):
Not primary research (2 studies):

- Punjasawadwong et al.[34] – a Cochrane review comparing BIS against standard practice.
- Anon[146] – a systematic review comparing BIS against standard practice, but pre-dating the Cochrane review by Punjasawadwong et al.[34]

Primary research other than RCT (3 studies):

- ElMenesy *et al*.[147]
- Pelletier *et al*.[148]
- Smajic *et al*.[149]

Exclusion criterion = Comparator (standard practice unclear or not defined) (4 publications):

- Bauer *et al*.[150]
- Riad *et al*.[151]
- Singh *et al*.[152]
- Weber *et al*.[153]

Publication retracted by journal (1 publication):

- Mayer *et al*.[76]

Appendix 5: Data extraction and critical appraisal forms used in the systematic review of patient outcomes

Reviewer 1: JS		Reviewer 2: GF	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 140 **Author:** Aime et al **Year:** 2006 **Study design**: RCT **Number of centres:** 1 **Country**: France **Sponsor**: GE Healthcare Monitoring Solutions loaned the authors a S5 monitor and provided the probes. No other funding source reported. **Trial name:** Not reported	**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: RE and SE 40-60. Intermittent bolus doses of sufentanil given if RE-SE 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**	**Total numbers involved**: n=140. Group 1 n=40, group 2 n=40, group 3 n=60 Pre-medication used: 100 mg hydroxyzine orally 1 hour before surgery. General anaesthetic used: IV propofol 2-3mg/kg (induction). Sevoflurane in 60% nitrous oxide with oxygen. Regional anaesthesia used: None Analgesia used: IV sufentanil 0.2-0.3µg /kg injected over 15-30s (induction), 0.15-0.20 µg · kg^{-1} · h^{-1} with 5 µg bolus given 5 mins before surgical incision. IV morphine for post-operative analgesia started approx. 20 mins prior to scheduled end of surgery (0.1-0.15mg/kg), plus paracetamol, nefopam, non-steroidal anti-inflammatory drugs. Muscle relaxants used: IV atracurium 0.5mg/kg. Anti-nausea drugs used: Not stated Other drugs used: Esmolol (for tachycardia), nicardipine 1-2mg (hypertension), ephedrine 3-6mg IV / phenylephrine 20-100 µg IV (for hypotension), atropine 0.5mg IV (bradycardia) Type of surgery: abdominal; gynaecological, urological, orthopaedic Duration of surgery: Precise duration not stated. Minimum 1 hour Duration of general anaesthesia: ranged from 170.8 (± 90.6) mins (standard practice group) to 190.8 (± 84.9 mins) (spectral entropy guided group). **Inclusion criteria**: Aged 18-80 years, ASA physical status I, II, III, scheduled for elective abdominal, gynaecologic, urologic or orthopaedic surgery expected to last at least 1 hour. **Exclusion criteria:** History of any	**Primary outcomes**: Reduction in sevoflurane consumption **Secondary outcomes**: • Sufentanil consumption • BIS and 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 1st and 3rd postoperative days. **Methods of assessing outcomes:** Sevoflurane consumption measured by sevoflurane vaporizer weight: mean for one patient; mean for one patient normalized to the duration of anaesthetic; mean for one patient normalized to the duration of anaesthetic and also to the weight of the patient. Intraoperative recall measured by standardized interview (Brice et al, 1970)

	of anaesthetist: Described as "more than three months of routine use"	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**: Gender male, n (%): Group 1 = 14 (41) Group 2 = 23 (62%) Group 3 = 23 (43%) Age yrs, mean (SD): Group 1 = 57 (±19) Group 2 = 58 (±18) Group 3 = 54 (± 15) Ethnic groups, n (%): not reported 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 Co-morbidities: None reported **Losses to follow up**: None reported **Place of anaesthetic administration:** Operating room	

Definitions: ASA = American Society of Anesthesiologists; RE = response entropy; SD = Standard Deviation; SE = state entropy

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 (min)	7.6 (± 4.1)	7.2 (± 4.7)	8.0 (± 3.9)	NR
Time to extubation (min)	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 sevoflurane consumption normalised $g \cdot h^{-1}$ sevoflurane consumption normalised $g \cdot kg^{-1} \cdot h^{-1}$	 21.3 (± 11.1) 7.2 (± 3.0) 0.10 (± 0.04)	 22.8 (± 14.4) 7.8 (± 3.4) 0.10 (± 0.05)	 25.6 (± 17.2) 9.4 (± 5.6) 0.14 (± 0.09)	 0.49 0.07 0.003

Health related quality of life	NR	NR	NR	NR
Nausea / vomiting / anti-sickness drugs	NR	NR	NR	NR
Pain / pain relieving drugs (for one patient) sufentanil induction dose $\mu g \cdot kg^{-1}$ sufentanil maintenance consumption $\mu g \cdot h^{-1}$ sufentanil maintenance consumption $\mu g \cdot kg^{-1} \cdot h^{-1}$	 0.22 (± 0.05) 14.0 (± 6.7) 0.20 (± 0.09)	 0.21 (± 0.05) 13.6 (± 6.1) 0.18 (± 0.09)	 0.23 (± 0.06) 14.9 (± 8.3) 0.22 (± 12)	 0.18 0.66 0.26
Other morbidity ephedrine use (n) nicardipine use (n) esmolol (n) atropine (n)	 3 1 0 1	 2 2 0 0	 4 2 1 0	 NR NR NR NR
Mortality	NR	NR	NR	NR

Definitions: NR= Not reported; SD = Standard deviation

Additional Results/comments (e.g., early response factors, quality of life):

- Percentage of time passed (induction, maintenance, recovery and total) with bradycardia (<75% of baseline values), normal range of heart rate, tachycardia (more than 125% of baseline values), hypotension (<75% of baseline values), normal range of mean arterial blood pressure, and hypertension (more than 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 only statistically significantly different between study arms 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: Not reported.
Blinding: Not reported.
Analysis by intention to treat: 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 entropy-guided group (Group 2) were statistically significantly heavier (p=0.04). More males were included in the entropy-guided group.
Method of data analysis: chi-square test for nominal data. One-way analysis of variance with Bonferroni 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 $\cdot kg^{-1} \cdot h^{-1}$. 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 entropy-guided groups.
Attrition/drop-out: 6 patients excluded from Group 3 (1 not extubated at the end of surgery due to hypothermia, 3 required intraoperative propofol administration, and missing data in 2 cases), 6 patients excluded from Group 1 (3 not extubated at the end of surgery because of hypothermia, 2 required intraoperative propofol administration, and monitor data were lost in 1 case), and 3 from Group 2 (all were not extubated at the end of surgery dye to hypothermia, 2 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.
Inter-centre variability: Not applicable.
Conflict of interests: None declared. Some of the monitoring equipment used was provided by GE Healthcare

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

Appendix 5 (continued)

Reviewer 1: JS		Reviewer 2: GF	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 672 **Author:** Avidan et al **Year:** 2011 **Study design**: RCT **Number of centres:** 3 **Country**: USA/Canada **Sponsor**: Foundation for Anesthesia Education & Research; American Society of Anesthesiologists; Winipeg Regional Health Authority & University of Manitoba Department of Anesthesia; Department of Anesthesiology at Washington in St. Louis; University; Department of Anesthesiology at University of Chicago. **Trial name:** BIS or Anesthetic Gas to Reduce Explicit Recall trial (BAG-	**Group 1:** BIS (Covidien) Target device/index value: 40-60 (audible alarms used outside of this range) **Group 2:** End-tidal anesthetic agent concentration (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	**Total numbers involved**: 6041 randomised 3021 (Group 1) 3020 (Group 2) Pre-medication 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 Anti-nausea 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 general anaesthesia: Not stated **Inclusion criteria**: 18 years or older, undergoing general anaesthesia 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 neurologic deficits. "Minor risk factors" for awareness as used in the B-Aware study were not used as enrolment criteria.	**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. 3 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 4th expert reviewer who reviews cases for the Anesthesia Awareness Registry of the American Society of Anesthesiologists, made the final determination.

RECALL)	practitioners to check BIS/ETAC and consider patient awareness.	**Baseline measurements**: Gender male, n (%): Group 1 =1621 (56.7) Group 2 = 1679 (58.9) Age yrs, mean (SD): Group 1 = 60 (± 14.2) Group 2 = 61 (± 14.4) 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) Co-morbidities: 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: Not reported	

Definitions: ASA = American Society of Anesthesiologists; MAC = Minimum alveolar concentration; BMI = Body Mass Index; SD = Standard Deviation;

Outcome	**Group 1**	**Group 2**	**Difference, BIS-ETAC percentage points (95% CI)**	**p-value**
Intraoperative Awareness, n/N (%) Definite Definite or possible	 7/2861 (0.24) 19/2861 (0.66)	 2/2852 (0.07) 8/2852 (0.28)	 0.17 (-0.03 to 0.38) 0.38 (0.03 to 0.74)	 0.98 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
Health related quality of life	NR	NR	NR	NR
Nausea / vomiting / Anti-sickness 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 30 day mortality	 33/2907 (1.14%) 57/2907 (1.96%)	 38/2902 (1.31%) 64/2902 (2.21%)	 NR 0.24 (-0.50 to 0.99)	 NR NR

Definitions: NR= Not reported, CI = Confidence Interval

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 9 patients had definite intraoperative awareness (incidence 0.16%, 95% CI 0.08 to 0.30), and 27 patents had define 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, s 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 5 of the 9 patients who experienced possible awareness did not have either BIS values of greater than 60 or ETAC values of less than 0.7 age-adjusted MAC.
- Overall, during the maintenance of anaesthesia the BIS was less than 60 a median of 94.0% of the time (interquartile range, 93.6 to 100), and the ETAC was greater than 0.7 age-adjusted MAC a median of 84.8% of the time (interquartile range, 67.2 to 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 intensive care unit 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 intention to treat: A modified intention-to-treat 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-square test, Fisher's exact test, unpaired Mann-Whitney U test or unpaired Student's t-test used for other comparisons.
Sample size/power analysis: 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, as 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/drop-out: Of 3021 randomised to Group 1, 114 (3.8%) were excluded post-randomisation. Of the remaining 2907, 46 (1.6%) were lost to follow-up and 2861 were assessed for intraoperative awareness. Of 3020 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 and 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 intravenous 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 aneasthetic 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.
Inter-centre 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
Definitions: ASA = American Society of Anaesthetists

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 post-randomisation exclusions and loss to follow-up and reasons were similar between study arms.
Reporting bias		
Selective reporting.	Low	No evidence to suggest selective reporting

Appendix 5 (continued)

Reviewer 1: GF		Reviewer 2: JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 1313 **Author:** Bannister et al **Year:** 2001 **Study design**: RCT **Number of centres:** Not reported; appears to be 1 **Country**: USA **Sponsor**: Supported in part by a grant from Aspect medical systems (device manu-facturer)	**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:** Not reported	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 3-18 years age group meets the systematic review age inclusion criterion and is reported here (mean age in the younger group ≤ 2.2 years). Pre-medication used: Midazolam 0.3 – 0.75 mg/kg (Group 1: 77.5%, Group 2: 88.6%). General anaesthesia (induction and maintenance): Sevoflurane 8% in 60% N_2O in oxygen. Regional anaesthesia: None. Analgesia: Fentanyl 1–2 µg/kg or morphine 0.05–0.1 mg/kg. Muscle relaxants: Nonpolarising IV neuromuscular block (no other details). Anti-nausea 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 general anaesthesia: Not reported **Inclusion criteria:** Not reported other than age 6 – 18 years and undergoing tonsillectomy and/or adenoidectomy **Exclusion criteria:** Not reported **Baseline measurements:** Gender 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	**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%

		Ethnic groups, n (%): Not reported Weight, kg, mean ± SD: Group 1: 26.9 ± 10.6 Group 2: 27.7 ± 14.7 ASA grade: Not reported Risk factors for awareness: None reported Co-morbidities: None reported Losses to follow up: None reported Place of anaesthetic administration: Not reported	

Definitions: ASA = American Society of Anesthesiologists; PACU=post-anaesthesia care unit

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 Last 15 minutes of GA End of procedure	 1.8 ± 0.4 1.6 ± 0.6 1.1 ± 0.6	 2.4 ± 0.6 2.1 ± 0.7 1.5 ± 0.7	 <0.05 <0.05 NS
Health related quality of life	NR	NR	NR
Nausea / vomiting / Anti-sickness 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

Definitions: GA = general anaesthesia; NR= not reported; NS=not statistically significant (p≥0.05)

Additional Results/comments (e.g., early response factors, quality of life):

- 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 inter-group 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: Not reported
Blinding: Single observer blinded to the patient groups was responsible for all PACU discharge assessments.
Analysis by intention to treat: 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 gender, 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 gender distribution.
Sample size/power analysis: Not reported
Attrition/drop-out: None reported

General comments
Generalisability: North American paediatric population aged 6–18 years undergoing tonsillectomy and/or adenoidectomy under sevoflurane for general anaesthesia; socio-economic details not reported. Not specifically identified as at risk for intraoperative awareness.
Inter-centre variability: Not applicable (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.

Domain	Author's judgement	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 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

Appendix 5 (continued)

Reviewer 1: JS		Reviewer 2: JB	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 771 **Author:** Bhardwaj & Yaddanapudi **Year:** 2010 **Study design**: RCT **Number of centres:** 1 **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 20µg · kg^{-1} min^{-1} to achieve a BIS value between 45 and 60) **Group 2:** Standard clinical practice (propofol infusion rate manually altered by 20µg · kg^{-1} min^{-1} 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**	**Total numbers involved**: 50 Group 1 = 25 Group 2 = 25 Pre-medication used: midazolam 0.5 mg · kg^{-1} General anaesthetic used: propofol 3 mg · kg^{-1} (induction). Propofol 150 µg · kg^{-1} · min^{-1} with nitrous oxide in oxygen (FiO_2 0.33) (maintenance). Regional anaesthesia used: None Analgesia used: morphine 0.1 mg · kg^{-1} (induction). Additional dose of opioid (fentanyl or morphine) was administered if signs of inadequate anaesthesia detected. Muscle relaxants used: Atracurium (0.5 mg · $kg^{-1)}$ used to facilitate tracheal intubation. Anti-nausea drugs used: Not reported Other drugs used: Atropine used to treat bradycardia (HR <80 of baseline). Neostigmine (0.05 mg · kg^{-1} and atropine (0.025 mg · kg^{-1}) used for reversal of neuromuscular blockade. Type of surgery: elective urogenital surgery Duration of surgery, mins. Mean (SD). Group 1 = 65.6 (29.2) Group 2 = 71.8 (27.3) Duration of general anaesthesia, mins. 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 general anaesthesia. **Exclusion criteria:** Patients with epilepsy and those taking drug known to affect EEG. **Baseline measurements**: Gender, male. n (%).	**Primary outcomes**: Reduction in consumption of propofol **Secondary outcomes**: Recovery from anaesthesia. **Length of follow up**: Not applicable (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.

	anaesthetist: Not stated	Group 1 = 21/25 (84%) Group 2 = 24/25 (96%) Age, yrs. Mean (SD). Group 1 = 6.3 (3.2) Group 2 = 6 (3) Ethnic groups, n (%): Not reported 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: Not reported Co-morbidities: Not reported **Losses to follow up**: Not applicable **Place of anaesthetic administration:** Premedication took place prior to transfer to the operation theatre. General anaesthesia was initiated in the operation theatre.	

Definitions: ASA = American Society of Anesthesiologists; HR = heart rate; SD = Standard devation

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^{-1} · min^{-1} (37.8)	106.6 µg · kg^{-1} · min^{-1} (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 mins (29.2)	86 mins (28.5)	NR Mean difference -4 (95% CI -20 to 13.5)
Health related quality of life	NR	NR	NR
Nausea / vomiting / Anti-sickness 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 Hypotension Bradycardia	 5/25 (20%) 6/25 (24%) 8/25 (32%)	 5/24 (21%) 7/24 (29%) 6/24 (25%)	 NR NR NR
Mortality	NR	NR	NR
Definitions: NR= Not reported; CI = confidence interval			

Additional Results/comments (e.g., early response factors, quality of life):

- 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 n both groups (2 patients in Group 1 compared to 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: Not reported.
Analysis by intention to treat: 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, gender).
Method of data analysis: Age, weight, HR, SBP, 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 $\mu g \cdot kg^{-1} \cdot min^{-1}$ (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/drop-out: 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.
Inter-centre variability: Not applicable
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, though is widely accepted as a tool in paediatric anaesthesia research.
Definitions:

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 though it does not say whether they were opaque.
Performance bias		
Blinding of participants and personnel	Unclear	Not reported
Detection bias		
Blinding of outcome assessment	Unclear	Not reported
Attrition bias		
Incomplete outcome data	Low	Only one exclusion from the study
Reporting bias		
Selective reporting	Low	No evidence to suggest selective reporting

Appendix 5 (continued)

Reviewer 1: GF		Reviewer 2: JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 1171 **Author:** Chan et al **Year:** 2010 **Study design**: RCT **Number of centres:** 2 **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: Not reported **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:** Not reported	**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. Pre-medication used: Not reported General anaesthetic used: Not explicitly reported but implied that both an inhalational agent and IV propofol were involved. Regional anaesthesia used: Not reported Analgesia used: Not reported Muscle relaxants used: Not reported Anti-nausea drugs used: Not reported Other drugs used: Not reported Type of surgery: Stated as major non-cardiac surgery (no other details) Duration of surgery: Not reported Duration of general anaesthesia: Not reported **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**: Not reported **Place of anaesthetic administration:** Not reported	**Outcomes** (not stated whether primary or secondary): • Postoperative cognitive dysfunction (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 8 neuropsychology tests before and at 1 and 3 weeks after surgery (no information on the tests reported). POCD was confirmed when 2 or more test parameters or the combined Z score > 1.96 (no further information given).

Definitions: ASA = American Society of Anesthestists; GA=general anaesthesia; IV=intravenous; POCD= Postoperative cognitive dysfunction

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 Target plasma propofol concentration	 25.3% reduction vs Group 2 [a] 20.7% reduction vs Group 2 [a]	 NR NR	 NR NR
Health related quality of life	NR	NR	NR
Nausea / vomiting / Anti-sickness drugs	NR	NR	NR
Pain / pain relieving drugs	NR	NR	NR
Other morbidity (e.g. cognitive dysfunction), n (%) [b] POCD, 1 week post-surgery POCD, 3 months post-surgery	 146 (32.5) 36 (8.1)	 177 (39.1) 54 (12.0)	 0.07 0.03 [OR (95% CI) = 1.6 (1.0–2.4)]
Mortality	NR	NR	NR
Definitions: CI=confidence interval; ETAC: End-tidal anaesthetic concentration; NR=not reported; OR=odds ratio; POCD= Postoperative cognitive dysfunction			

Additional Results/comments (e.g., early response factors, quality of life):

- 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 intravenous 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: Not reported.
Blinding: Not reported.
Analysis by intention to treat: 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: Not reported.
Sample size/power analysis: Not reported.
Attrition/drop-out: Not reported. The starting number of patients (921) is 20 more than the total sample size indicated for outcomes data (449 + 452 = 901); unclear whether this difference reflects attrition pre- or post-randomisation.

General comments
Generalisability: Elderly Chinese patients (>60 years) undergoing major non-cardiac surgery under general anaesthesia, but limited information on the types of anaesthesia (appears to include both inhaled and intravenous); unclear population characteristics (gender, weight, comorbidities not reported); unclear surgical procedures (no information reported); and unclear which groups some outcomes were reported for. Not reported whether population was at high risk of intraoperative awareness.
Inter-centre variability: Not reported.
Conflict of interests: None reported.

Domain	Author's judgement	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

[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.

Appendix 5 (continued)

Reviewer 1:JS		Reviewer 2: GF	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 21 **Author:** Choi et al **Year:** 2010 **Study design**: RCT **Number of centres:** Not stated (presume single centre) **Country**: South Korea **Sponsor**: Dong-A University	**Group 1:** 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. Pre-medication used: intravenous midazolam (0.15 mg/kg) General anaesthetic used: 5% vol% sevoflurane in oxygen at fresh gas flow of 51/min (induction). Sevoflurane administration was started at 2.5 vol% in air and oxygen 1.51/min. Regional anaesthesia used: not stated Analgesia used: intra-operative 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. Anti-nausea drugs used: Not reported Other drugs used: Not reported 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 general anaesthesia, 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 anti-seizure medication. **Baseline measurements**: Gender male, n (%): Group 1 = 25/39 (64) Group 2 = 27/39 (69) Age, yrs. Median (range): Group 1 = 4.0 (3.0-12.0) Group 2 = 6.0 (3.0 – 11.0)	**Primary outcomes**: 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 post-operative 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 intra-operative recall in the post-anaesthetic care unit and on the first post-operative day by an independent nurse. Time to the various recovery parameters were measured following discontinuation of sevoflurane. Complete recovery was defined as a score of 9 or more on a modified Aldrete score.

		Ethnic groups, n (%): Not reported 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 Co-morbidities: None reported **Losses to follow up**: Not reported **Place of anaesthetic administration:** Not stated	

Definitions: ASA = American Society of Anesthesiologists; i.v. = intravenous

Outcome	**Group 1**	**Group 2**	**p-value**
Intraoperative awareness / recall	Anaesthesia and surgery related memories were not reported by any patients in the post-operative interview.		
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time to emergence from anaesthesia, mins. Mean (SD): Eye–opening Orientation Complete recovery	 14.3 (3.6) 18.2 (4.0) 24.3 (7.3)	 18.0 (3.3) 23.3 (5.0) 28.8 (5.7)	 NS <0.05 <0.05
Time to extubation, minutes. Mean (SD)	8.3 (1.4)	11.9 (2.5)	<0.05
Time 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
Health related quality of life	NR	NR	NR
Nausea / vomiting / Anti-sickness drugs	NR	NR	NR
Pain / pain relieving drugs	NR	NR	NR
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

Definitions: NR= Not reported; NS= Not statistically significant

Additional Results/comments (e.g., early response factors, quality of life):
Systolic and diastolic blood pressure were significantly higher in Group 1 compared to 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 intention to treat: Not reported. Analysis excludes two patients out of the 80 enrolled due to "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-square 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 α of 0.05 and a statistical power of 0.9. Forty patients were enrolled in each group for redundancy.
Attrition/drop-out: Two patients out of the 80 enrolled were excluded from the analysis due to "technical problems"

General comments
Generalisability: Results applicable to Korean children without any apparent co-morbidities undergoing tonsillectomy / adenoidectomy. Not stated to be at increased risk for intra-operative awareness.
Inter-centre variability: Not applicable (presumed single centre)
Conflict of interests: None 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, though it is not reported whether 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, though 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

Appendix 5 (continued)

Reviewer 1: JB		Reviewer 2: JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 30 **Author:** Ellerkmann et al **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 SE to target value of 50 during maintenance. Target SE value of 60 to facilitate rapid emergence from anaesthesia (15 mins 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 mins before expected end of surgery) In the Entropy and BIS group a propofol bolus of 0.25 mg/kg could be give in the presence of a sudden increase in SE 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 1mg/kg/hour as necessary for clinical parameters.	**Total numbers involved**: 90 Group 1: 30 Group 2: 30 Group 3: 30 Pre-medication 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 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 remifentamil reduced to 0.08 µg/kg/minute in order to tolerate tube. Anti-nausea drugs used: not reported Other drugs used: 0.3 ml of iv vasopressor (akrinor, 1ml 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: not reported Duration of general anaesthesia, mins. 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 one 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.	**Primary outcomes**: Reduction in propofol consumption **Secondary outcomes**: Remifentanil consumption, recovery time, duration of anaesthesia, intraoperative awareness, BIS and 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 post-anaesthesia care unit Intraoperative awareness by 'standardised interview' (first and third day post-operative days) (Nordstrom 1997)

	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 Entropy monitors were covered behind a curtain; in the BIS and Entropy group, either only the BIS monitor or only the Entropy module was uncovered. **Length of experience / training of anaesthetist:** 'Experienced anaesthesiologist'	**Baseline measurements**: Gender male, n (%): Group 1 = 15/25 (60%) Group 2 = 18/27 (67%) Group 3 = 15/27 (56%) Age (yrs), mean (SD): Group 1 = 58.1 (14.2) Group 2 = 50.6 (15.7) Group 3 = 53.6 (18.4) Ethnic groups, n (%): not reported 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: not reported Co-morbidities: not reported **Losses to follow up:** none **Place of anaesthetic administration:** Premedication prior to operating theatre; general anaesthesia initiated in operating theatre.	

Definitions: ASA = American Society of Anesthesiologists, SD = standard deviation; SE = state entropy

Outcome	**Group 1 – 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 to emergence from anaesthesia mins, 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 Grp 1 vs Grp 2 Ns (no p value given) for Grp1/2 vs Grp 3
Time to extubation	NR	NR	NR	
Time to discharge to / from the recovery room	NR	NR	NR	

Anaesthetic consumption Propofol µg/kg/min Mean (SD) Remifentanil µg/kg/min Mean (SD) Bolus of propofol following rise in BIS or Entropy (SE) above 65 or sudden unexpected somatic response, n	106 (24) 0.08 (0.02) 12	104 (20) 0.08 (0.02) 8	101 (22) 0.09 (0.02) 10	P=0.27 Grp 1/2 vs Grp 3 P=0.56
Health related quality of life	NR	NR	NR	
Nausea / vomiting / Anti-sickness drugs	NR	NR	NR	
Pain / pain relieving drugs	NR	NR	NR	
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR	
Mortality	NR	NR	NR	

Definitions: NR= Not reported, SE = State Entropy, 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 one minute after extubation.
- Various 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: Not reported
Blinding: Not reported
Analysis by intention to treat: No
Comparability of treatment groups at baseline: No differences between groups in age, weight and height by ANOVA; not reported for gender and ASA status.
Method of data analysis: Normally distributed data compared with between-group analysis of variance and Tukey HSD 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 suign 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/drop-out: patients excluded from analysis due to 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 (ie pain perception completely blocked) which could limit generalisability. Also more than one type of surgery was included and more than one regional anaesthesia technique which might contribute to different levels of analgesia. Authors state that similar results may not have been obtained with less experienced anaesthestists. Results applicable to adult patients receiving intravenous general anesthesia (and regional anaesthesia) assumed not to have significant morbidities.
Inter-centre variability: not applicable
Conflict of interests: 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	Grp 1 17% patients excluded from analysis; Grp 2 and Grp 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.		

Appendix 5 (continued)

Reviewer 1: GF		Reviewer 2: JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 98 **Author:** Gruenewald et al **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:** 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 min 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 (BP > 120% of baseline), hypo-tension (BP < 80% of baseline), tachycardia (> 90 beats/min), bradycardia (HR < 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	Total numbers involved: 72 Group 1: 37; Group 2: 35 Pre-medication 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/min. 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 min before end of surgery. Muscle relaxants used: Rocuronium 0.6 mg/kg. Anti-nausea 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 general anaesthesia, min, mean ± SD: Group 1: 110 ± 39; Group 2: 111 ± 46 Inclusion criteria: Not reported (implied adult female population). Exclusion criteria: Pregnancy, neurological or neuromuscular disease, use of CNS-active medication, abuse of alcohol or illicit drugs Baseline measurements: Gender male, n (%): 0 (0) Age, years, mean ± SD: Group 1: 38 ± 9; Group 2: 33 ± 9 Ethnic groups, n (%): Not reported Weight, kg, mean ± SD: Group 1: 68 ± 15; Group 2: 68 ± 13	**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 (OAAS, 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. Postoperative nausea and vomiting:

	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	ASA grade 1/2, n: Group 1: 14/23; Group 2: 11/24 Risk factors for awareness: Not reported Co-morbidities: Not reported Losses to follow up: Not reported Place of anaesthetic administration: Not reported	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
Definitions: ASA = American Society of Anesthesiologists; BP=blood pressure; HR=heart rate; OAAS=Observer Assessment of Alertness and Sedation scale			

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 (%) Stated no difference between groups in awareness or explicit memory assessed 24 hours post-surgery (no further quantitative data provided)	0 (0)	1 (2.8) [a]	Not reported
Patient distress and sequelae resulting from perioperative awareness	Not reported	Not reported	Not reported
Time to emergence from anaesthesia, minutes Median [IQR] (range) time to eye opening	3 [1–5] (0–9)	4 [3–6] (0–14)	NS
Time to extubation	Not reported	Not reported	Not reported
Time to discharge to / from the recovery room	Not reported	Not reported	Not reported
Mean ± SD anaesthetic consumption (induction + maintenance), µg/kg/min: Propofol Remifentanil	 81 ± 22 0.46 ± 0.08	 95 ± 14 0.39 ± 0.08	 <0.01 <0.001
Health related quality of life	Not reported	Not reported	Not reported
Nausea / vomiting: Nausea and vomiting, n (%) (on arrival in recovery room) Anti-sickness drugs: None reported	15 (41)	13 (37)	NS
Pain: Median [IQR] (range) pain intensity score (on arrival in recovery room) Pain relieving drugs: Stated analgesia (piritramide) did not differ between groups (no quantitative data reported)	6 [4–7] (2–10)	4 [3–5] (1–10)*	0.03
Other morbidity (e.g. cognitive dysfunction)	Not reported	Not reported	Not reported
Mortality	Not reported	Not reported	Not reported
Definitions: IQR=interquartile range; NR= not reported; NS not statistically significant (p≥0.05)			

Additional Results/comments (e.g., early response factors, quality of life):

- Patients in Group 2 had significantly more hypertension, hypotension, tachycardia, bradycardia and somatic responses (purposeful movements) compared to 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 [IQR] (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 opaque.
Blinding: OAAS, postoperative nausea and vomiting, 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 intention to treat: 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 nearly identical 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 ANOVA '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 min, an α 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 intra-operative awareness, given the low incidence rate.
Attrition/drop-out: Not reported

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.
Inter-centre variability: Not applicable: 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.

[a] implied this was a female patient who did not report feeling any pain
* asterisk as reported with the original data - meaning not stated

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 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 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

Appendix 5 (continued)

Reviewer 1: JB		Reviewer 2: JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 1116 **Author:** Kamal **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 sevofluarane was increased; If BIS in target range fentanyl 25-50 µg IV given; if BIS <50 sevoflurane decreased and patient checked for lack of analgesia; if lack of analgesia fentanyl 25-50 µg IV given; if no lack of analgesia labetalol 5-10 mg IV 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/min) anaesthesia was deepened by increasing inspired sevoflurane or adjusting fentanyl 25-	**Total numbers involved**: 60 Group 1 = 30 Group 2 = 30 Pre-medication used: none used General anaesthetic used: Propofol 1-2 mg/kg IV and fentanyl 2-3 µg/kg IV (induction). Sevoflurane and 50% nitrous oxide with oxygen 2 l/min (continued). Nitrous oxide discontinued, sevoflurane adjusted for BIS index in Group 1 and as usual practice in Group 2 (10 mins 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 IV. Intermittent boluses of atracurium 0.2-0.3 mg/kg IV. Anti-nausea drugs used: not reported Other drugs used: Ephedrine 3-6mg IV or phenylephrine 20-100 µg IV (for hypotension). Atropine 0.02 mg/kg IV (for bradycardia). Glycopyrate 0.01 mg/kg and neostigmine 0.05 mg/kg IV 5 min before discontinuation of anaesthesia (to reverse residual neuromuscular blockade. Type of surgery: elective moderate abdominal surgery Duration of surgery, mins. Mean (SD) : Group 1 = 91.7 (11.3) Group 2 = 85.8 (17.4) Duration of general anaesthesia, mins. 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	**Primary outcomes**: Not specified **Secondary outcomes**: Not specified **Outcomes**: Recovery times (awakening, tracheal extubation, orientation, arrival at PACU, discharge from PACU) BIS index values Anaesthestic 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 min 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.

	50 µg IV or labetalol 5-10 mg IV according to anaesthesiologists 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	psychoactive medication and a body mass index >40. **Baseline measurements**: Gender male, n (%): Group 1 = 18 (62%) Group 2 = 20 (71%) Age (yrs), mean (SD): Group 1 = 51.6 (7.4) Group 2 = 52.1 (5.2) Ethnic groups, n (%): not reported 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*:* not reported Co-morbidities: not reported **Losses to follow up**: none **Place of anaesthetic administration:** not reported	

Definitions: ASA = American Society of Anesthesiologists; PACU = post anaesthesia care unit; IV = intravenous

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 to emergence from anaesthesia after termination of anaesthesia (awakening eye opening, mins)	4.1 (1.6)	4.4 (1.9)	ns
Time to extubation	4.3 (2.1)	4.8 (2.3)	ns
Time to discharge to / from the recovery room Arrival at PACU (min) PACU discharge (min)	 9.4 (1.9) 53.9 (14.7)	 14.1 (2.8) 78.6 (21.5)	 P<0.01 P<0.01
Anaesthetic consumption Sevoflurane mL, mean (SD) End tidal sevoflurane vol %, mean (SD) Propofol mg, mean (SD) Fentanyl µg, mean (SD)	 5.7 (1.9) 0.43 (0.3) 161.7 (27.5) 383.7 (62.6)	 8.4 (2.3) 0.59 (0.1) 157.9 (35.8) 389.4 (41.5)	 P<0.01 p≤0.01 ns ns
Health related quality of life	NR	NR	
Nausea / vomiting / Anti-sickness drugs	NR	NR	
Pain / pain relieving drugs	NR	NR	
Other morbidity (e.g. cognitive dysfunction)	NR	NR	
Mortality	NR	NR	

Definitions: NR= Not reported

Additional Results/comments (e.g., early response factors, quality of life):

- Orientation (min) 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 post-operative 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 intention to treat: No as 3 patients not included in analysis.
Comparability of treatment groups at baseline: authors state groups comparable but no p values 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-square test.
Sample size/power analysis: Not reported
Attrition/drop-out: As above. One patient in Group 1 was desaturated intra-operatively 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 (ie starting point of recovery process variable). Results applicable to adults receiving inhaled anaesthesia for moderate abdominal surgery.
Inter-centre variability: Not applicable
Conflict of interests: 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 3 patients not included in analysis (see above)
Reporting bias		
Selective reporting.	Low	No evidence of selective reporting
Other bias		
Other sources of bias.		

Appendix 5 (continued)

Reviewer 1: GF		Reviewer 2: JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 802 **Author:** Kerssens et al **Year:** 2009 **Study design**: RCT **Number of centres:** Not reported **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: Not reported **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:** Not reported	**Total numbers involved:** 128 Number randomised: Group 1: 67; Group 2: 61. Pre-medication used: Stated benzodiazepines were not given to any patients pre- or intra-operatively. General anaesthetic used: Induction: Propofol 2mg/kg. Maintenance: Sevoflurane in oxygen using standard ventilation parameters (not specified). Regional anaesthesia used: Used only for post-operative pain management. Analgesia used: Fentanyl 3 µg/kg (induction); 50–100 µg (maintenance). Muscle relaxants used: Vercuronium bromide 0.1 mg/kg with additional doses as necessary (tracheal intubation). Anti-nausea 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: Not reported. Duration of general anaesthesia, 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 general anaesthesia. **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 Mini-Mental State Examination (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	**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:** Physiologic parameters, BIS, end-tidal gas concentrations (every 5s) and vital signs (every 3 s) were automatically recorded to a computer using Rugloop (Demed, Belgium) Recall assessment: 6 hours after surgery, consisting of 5 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 pre-determined neutral words were 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 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 (three-alternative forced choice).

		sample were similar): Gender 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 (%): Not reported Weight, kg, mean ± SD: Group 1: 87.9 ± 18.9; Group 2: 84.4 ± 14.8 BMI, 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 Co-morbidities: None reported (patients with co-morbidities were excluded) Losses to follow up: Attrition reported, with reasons, both pre- and post-randomisation Place of anaesthetic administration: Not reported	

Definitions: ASA = American Society of Anesthesiologists; EEG=electroencephalogram; BMI=body mass index; MMSE=Mini-Mental State Examination; 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 post-operatively selecting a word presented during anaesthesia (target) or not presented during anaesthesia (distractor), mean ± SD: Target Distractor	0.371 ± 0.132 0.315 ± 0.117	0.323 ± 0.132 0.338 ± 0.119	Not reported [c] Not reported [c]
Patient distress and sequelae resulting from perioperative awareness	Not reported	Not reported	Not reported
Time to emergence from anaesthesia	Not reported	Not reported	Not reported
Time to extubation	Not reported	Not reported	Not reported
Time to discharge to / from the recovery room	Not reported	Not reported	Not reported

Anaesthetic consumption End-tidal gas concentration, %, mean ± SD: Maintenance phase During word presentation	 1.31 ± 0.29 [a] 1.30 ± 0.31 [a]	 1.56 ± 0.29 [b] 1.60 ± 0.37 [b]	 <0.001 NS [d]
Health related quality of life	Not reported	Not reported	Not reported
Nausea / vomiting / Anti-sickness drugs	Not reported	Not reported	Not reported
Pain / pain relieving drugs Fentanyl analgesia, mean ± SD: Preoperative, µg/kg Intraoperative, including induction dose, µg/kg/h Postoperative, µg/kg	 0.27 ± 0.43 [a] 2.83 ± 1.04 [a] 0.47 ± 0.66 [a]	 0.40 ± 0.47 [b] 2.70 ± 1.18 [b] 0.55 ± 1.10 [b]	 NS [d] NS [d] NS [d]
Other morbidity (e.g. cognitive dysfunction)	Not reported	Not reported	Not reported
Mortality	Not reported	Not reported	Not reported
Definitions: NS=not significant			

[a] reported for post-attrition subgroup (n=62)
[b] reported for post-attrition subgroup (n=47)
[c] see additional comments for interpretation of within-group differences
[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. $p\geq0.05$)

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 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 was reported only for baseline; stated that post-operative STAI results can be found elsewhere, together with results of a depression questionnaire, but no references were provided

Appendix 5 (continued)

Reviewer 1: JB		Reviewer 2: JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 177 **Author:** Kreuer et al **Year:** 2005 **Study design**: RCT **Number of centres:** 1 **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 mins 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 mins before expected end of surgery desflurane adjusted to target value of C_1. In Group1 and 2 if anaesthesia judged inadequate although target value achieved, infusion rate of remifentanil increased by 0.05 µg/kg/min. **Group 3:** Standard anaesthetic practice protocol If anaesthesia inadequate desflurane concentration increased in steps of 0.5% vol as necessary. If insufficient remifentanil increased by 0.05µg/kg/min. Hypotension treated with desflurane concentration reduced in steps of	**Total numbers involved**: 120 Group 1 = 40 Group 2 = 40 Group 3 = 40 Pre-medication used: midazolam 7.5 mg orally in the evening and on the morning before surgery. General anaesthetic used: Induction – remifentanil infusion 0.4µg/kg/min, 5 min later 2mg/kg propofol for hypnosis. After intubation remifentanil reduced to constant rate of 0.2µg/kg/min, Desflurane adjusted according to EEG target values or clinical variable. 15 mins 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: not reported Analgesia used: 100mL infusion of 0.9% NaCl + metamizol 25 mg/kg for postoperative pain relief. Muscle relaxants used: 0.5mg/kg atrcurium Anti-nausea drugs used: not reported Other drugs used: Hypotension treated with an IV vasopressor (Akrinor, 1ml contains 100mg of cafedrine and 5mg of theodrenaline) given at dose chosen by investigator. Atropine 0.5mg for bradycardia. Type of surgery: minor orthopaedic surgery Duration of surgery: not reported Duration of general anaesthesia (mins). 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 hr.	**Primary outcomes**: Time taken to spontaneous opening of eyes. **Secondary outcomes**: Not explicitly stated (times to tracheal extubation and arrival at PACU, consumption of desfluane) **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 day.

	0.5 vol%. Desflurane reduced 15 mins 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	**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**: Gender male, n (%): Group 1 = 20/40 (50) Group 2 = 20/40 (50) Group 3 = 20/40 (50) Age (yrs), mean (range): Group 1 = 46.5 (14.1) Group 2 = 44.7 (15.6) Group 3 = 43.6 (16.0) Ethnic groups, n (%): not reported 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: not reported Co-morbidities: not reported **Losses to follow up**: not reported **Place of anaesthetic administration:** in the operating room	

Definitions: ASA = American Society of Anaesthetists, PACU = postanaesthestic care unit

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 to eye opening (mins), mean (SD) Reduction compared with std practice (%)	4.2 (2.1) -10.6	3.7 (2.0) -21.3	4.7 (2.2) n/a	Ns

Time to extubation (min), mean (SD) Reduction compared with std practice (%)	4.4 (2.2) -12.0	3.6 (2.0)* -28.0	5.0 (2.4) n/a	*p<0.05 Group 2 vs Group 3
Time to discharge to PACU (min), mean (SD) Reduction compared with std practice (%)	8.4 (2.4)* -10.6	8.0 (1.9)* -15.0	9.4 (2.4) n/a	*p<0.05 Group 2 vs Group 3
Anaesthetic consumption per patient Desflurane mg, mean (SD) Reduction compared with std practice (%) Desflurane mg/min, mean (SD) Reduction compared with std practice (%) Normalised remifentanil infusion rates, µg/kg/min, mean (SD)	 4861.7 (2948.3) -12.4 416.2 (99.1)* -6.2 0.22 (0.05)	 4655.9 (2891.7) -16.1 374.6 (124.2)* -15.7 0.22 (0.06)	 5547.3 (2396.4) n/a 443.6 (71.2) n/a 0.23 (0.07)	 ns *p<0.05 ns
Health related quality of life	NR	NR	NR	
Nausea / vomiting / Anti-sickness drugs	NR	NR	NR	
Pain / pain relieving drugs	NR	NR	NR	
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR	
Mortality	NR	NR	NR	
Definitions: NR= Not reported; ns = not statistically significant				

Additional Results/comments (e.g., early response factors, quality of life):

- 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;
- 5 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 intention to treat: 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 min at $\alpha = 0.05$.
Attrition/drop-out: none.

General comments
Generalisability: Observed differences were minimal and not clinically significant. Results applicable to patients receiving general anaesthesia with desflurane-remifentanil for minor orthopaedic surgery.
Inter-centre variability: n/a
Conflict of interests: funding source stated but no other details 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
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.		

Appendix 5 (continued)

Reviewer 1: JB		Reviewer 2: JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 207 **Author:** Kreuer et al **Year:** 2003 **Study design**: RCT **Number of centres:** 1 **Country**: Germany **Sponsor**: Support soley from departmental sources	**Group 1:** BIS A-2000 monitor (software version 3.0) Propofol TCI continuously adjusted to target value of 50 BIS. 15 mins 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 mins 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 0.05µg/kg/min.	**Total numbers involved**: 120 Group 1 = 40 Group 2 = 40 Group 3 = 40 Pre-medication used: 0.15 mg/kg diazepam orally in the evening and on the morning before surgery. General anaesthetic used: Induction – remifentanil infusion 0.4µg/kg/min, 5 min later propofol TCI, initially started at 3.5µg/ml. After intubation remifentanil reduced to constant rate of 0.2µg/kg/min, Propofol TCI adjusted according to EEG target values or clinical variables. 15 mins 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: not reported Analgesia used: 100mL infusion of 0.9% NaCl + metamizol 25 mg/kg for postoperative pain relief. Muscle relaxants used: 0.1mg/kg cisatrcurium. Anti-nausea drugs used: not reported Other drugs used: Hypotension treated with an IV vasopressor (Akrinor, 1ml contains 100mg of cafedrine and 5mg of theodrenaline) given at dose chosen by investigator. Atropine 0.5mg for bradycardia. Type of surgery: minor orthopaedic surgery Duration of surgery: not reported Duration of general anaesthesia (mins). 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 one hour	**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.

	Hypotension treated with propofol concentration reduced in steps of 0.5µg/ml. Propofol reduced 15 mins 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	**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**: Gender male, n (%): Group 1 = 20/40 (50) Group 2 = 20/40 (50) Group 3 = 20/40 (50) Age (yrs), mean (SD): Group 1 = 43.8 (4.2) Group 2 = 44.8 (15.9) Group 3 = 46.1 (14.5) Ethnic groups, n (%): not reported Weight (kg). Mean (SD): Group 1 = 78.3 (13.8) Group 2 = 76.6 (11.7) Group 3 = 82.7 (17.8) 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: not reported Co-morbidities: not reported **Losses to follow up**: not reported **Place of anaesthetic administration:** in the operating room	

Definitions: ASA = American Society of Anesthesiologists; TCI = target-controlled infusion; PACU postanaesthesia care unit

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 to emergence from anaesthesia (mins), mean (SD) Reduction compared With std practice (%)	3.5 (2.9)* -63.4	3.4 (2.2)* -62.4	9.3 (5.2) n/a	*p<0.001 Group1/2 vs Group 3

Time to extubation (min), mean (SD) Reduction compared with std practice (%)	4.1 (2.9)* -57.7	3.7 (2.2)* -61.9	9.7 (5.3) n/a	*p<0.001 Group 1/2 vs Group 3
Time to discharge to PACU (min), mean (SD) Reduction compared with std practice (%)	7.0 (3.2)* -43.5	6.6 (2.8)* -46.7	12.4 (5.7) n/a	*p<0.001 Group 1/2 vs Group 3
Anaesthetic consumption per patient Propofol mg, mean (SD) Reduction compared with std practice (%) Propofol mg/kg/hr, mean (SD) Reduction compared with std practice (%) Normalised remifentanil infusion rates, µg/kg/min, mean (SD)	 720.6 (245.3)* -25.7 4.8 (1.0)* -29.4 0.22 (0.07)	 721.3 (401.2)** -25.7 4.5 (1.1)* -33.8 0.21 (0.07)	 970.5 (384.4) n/a 6.8 (1.2) n/a 0.20 (0.07)	 *p<0.001 **p<0.05 *p<0.001 ns
Health related quality of life	NR	NR	NR	
Nausea / vomiting / Anti-sickness drugs	NR	NR	NR	
Pain / pain relieving drugs	NR	NR	NR	
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR	
Mortality	NR	NR	NR	
Definitions: NR= Not reported				

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;
- 5 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 intention to treat: 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 Students *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 min at $\alpha = 0.05$.
Attrition/drop-out: none reported.

General comments
Generalisability: Sex differences observed within groups (see above). Results applicable to patients receiving intravenous general anaesthesia with propofol-remifentanil for minor orthopaedic surgery.
Inter-centre variability: n/a
Conflict of interests: not reported
Other:
Definitions:

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.		

Appendix 5 (continued)

Reviewer 1: GF		Reviewer 2: JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 246 **Author:** Lai et al **Year:** 2010 **Study design**: RCT **Number of centres:** 1 **Country**: China **Sponsor**: Not reported	**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 D2 and E0, then the fentanyl infusion rate was adjusted 10 minutes before end of surgery to target NT values between D0 and D1. 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 BPM), mean arterial pressure (normal = baseline value ± 20%), and body movement **Length of experience / training of anaesthetist:** Not reported	**Total numbers involved**: 40 Group 1: 20; Group 2: 20 Pre-medication used: None reported. General anaesthetic used (total intravenous anaesthesia): Induction: Propofol 3 mg/kg/h. Maintenance: Propofol 4–8 mg/kg/h. 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 2μg/kg. Maintenance: Fentanyl 1μg/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). Anti-nausea drugs used: None reported. Other drugs used: Tachycardia (>100 BPM): Fentanyl 1μg/kg, with metoprolol 1mg added as necessary. Hypertension (>20% above baseline value): urapidil 10–15 mg. Body movement: Fentanyl 1μg/kg. Bradycardia (<50 BPM): 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: Not reported Duration of general anaesthesia, minutes, mean ± SD: [a] Group 1: 91 ± 30; Group 2: 88 ± 31; difference NS. **Inclusion criteria**: Patients with liver cancer scheduled to undergo	**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 visual analogue scores (VAS) **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.

		microwave coagulation under the guidance of computed tomography (CT) **Exclusion criteria:** Neurologic or psychiatric problems; hearing defects; alcohol or drug dependence **Baseline measurements**: Gender male, n (%): Not reported Age, years, mean (range): Group 1: 44 (25–69); Group 2: 41 (20–70); difference NS Ethnic groups, n (%): Probably Chinese (not reported) 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 Co-morbidities: 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)	

Definitions: ASA = American Society of Anesthesiologists; BPM: beats per minute; CT: computed tomography; 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)	Not applicable
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time to emergence from anaesthesia, minutes, mean ± SD: Arousal time Duration of orientation recovery	 4.9 ± 2.2 6.6 ± 3.2	 9.5 ± 2.9 12.2 ± 3.5	 <0.01 <0.01
Time to extubation	Not applicable	Not applicable	Not applicable
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
Health related quality of life	NR	NR	NR

Nausea / vomiting / Anti-sickness drugs: Nausea or vomiting reported after surgery, n (%)	0 (0)	0 (0)	Not applicable
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
Definitions: NR= Not reported			

Additional Results/comments (e.g., early response factors, quality of life):

- Stated there were no differences in heart rate or blood pressure between the two groups pre-operation, 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: Not reported.
Blinding: Not reported.
Analysis by intention to treat: 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: Gender 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: Not reported
Attrition/drop-out: Not explicitly reported but there do not appear to have been any drop outs.

General comments
Generalisability: Liver cancer patients eligible for microwave coagulation. Gender and ethnicity not reported, but appears to be a Chinese population. Early 40s in age, with ASA grade <III, most without concurrent hypertension, receiving total intravenous anaesthesia with propofol and fentanyl. No specific risk factors for intraoperative awareness identified.
Inter-centre variability: Not applicable (one centre).
Conflict of interests: Not reported.

[a] variance parameter not specified; assumed by reviewer to be SD
[b] postoperative visual analogue scores reported as an outcome - data not extracted by reviewer as no explanation or interpretation of the scores was provided
[c] not stated whether 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 any checks were made to

		ensure fidelity of the published version to the original work.

Appendix 5 (continued)

Reviewer 1: GF		Reviewer 2: JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 692 **Author:** Liao et al **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 Pediatric 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	**Total numbers involved**: 160 Group 1: 52. Group 2: 54 (Group 3: 54 – data not extracted) Pre-medication used: Stated none. General anaesthetic 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: Intravenous fentanyl 1μg/kg five minutes before incision. Muscle relaxants used: Stated none (patients breathed spontaneously). Anti-nausea drugs used: None reported. Other drugs used: In the post-anaesthesia care unit (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 general anaesthesia (GA), minutes, mean ± SD: Group 1: 39.5 ± 11.7; Group 2: 41.8 ± 14.0 (p=0.44 for 3-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 3-group comparison). **Inclusion criteria**: Pre-puberty children, aged 3–12 years, with ASA physical status I or II, scheduled for elective urologic outpatient surgery	**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 of 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) Pediatric Anesthetic Emergence Delirium (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.

	BIS monitors were positioned out of the anaesthesiologist's line of sight. **Length of experience / training of anaesthetist:** Not reported; 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)	**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 3-group comparisons; data for Group 3 not extracted):** Gender 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 (not reported) 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 Co-morbidities: 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	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.

Definitions: ASA = American Society of Anesthesiologists; GA: General anaesthesia; PACU: Post-anaesthesia care unit; PAED: Pediatric Anesthetic Emergence Delirium score

Outcome	**Group 1** (n=52)	**Group 2** (n=54)	**p-value (a) for 3-group comparison; (b) post-hoc comparison Group 1 v Group 2**
Intraoperative awareness with explicit recall, n (%)	0 (0)	0 (0)	Not applicable
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR

Time to emergence from anaesthesia, minutes, mean ± SD: Spontaneous movement Phonation Eyes opening	 3.6 ± 2.7 8.4 ± 5.2 15.0 ± 16.4	 6.1 ± 5.7 12.9 ± 9.0 16.1 ± 11.3	 (a) 0.02; (b) <0.05 (a) 0.11 (a) 0.17
Time to extubation – not applicable Time to laryngeal mask airway removal, minutes, mean ± SD	 1.8 ± 1.6	 2.1 ± 2.4	 (a) 0.93
Time to discharge from the recovery room, minutes, mean ± SD	64.5 ± 10.1	66.8 ± 9.0	(a) 0.03; (b) <0.05
Anaesthetic consumption Sevoflurane, g/minute, mean ± SD Mean end-tidal sevoflurane concentration, %, during Maintenance (See also additional comments below concerning anaesthetic consumption at different time points)	 0.6 ± 0.2 2.5 ± 0.4	 0.9 ± 0.3 2.9 ± 0.5	 (a) <0.001; (b) <0.01 (a) 0.001 (b) <0.01
Health related quality of life	NR	NR	NR
Nausea / vomiting / Anti-sickness drugs Postoperative nausea, n (%) Postoperative vomiting, n (%)	 5 (10) 2 (4)	 6 (11) [a] 3 (6) [a]	 (a) 0.95 (a) 0.88
Pain / pain relieving drugs, n (%) Did not receive analgesic or sedative agents Rescue requiring more analgesic or sedative agents Fentanyl use, µg, mean ± SD	 4 (8) [a] 9 (17) 24.8 ± 11.1	 5 (9) 6 (11) [a] 23.4 ± 9.1	 (a) 0.83 (b) 0.6 (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)	Not applicable
Definitions: NR= Not reported			

Additional Results/comments (e.g., early response factors, quality of life):

- 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 3-group comparisons of these variables all >0.7; data not extracted). Baseline data were also reported for the BMI-for-age percentile (3-group comparison, p=0.52) and Induction Compliance Checklist score (3-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 3-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: Not reported.
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 the nurse who assessed intraoperative recall was blinded to the patient group.
Analysis by intention to treat: 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-square or Fisher 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 (α=0.05).
Attrition/drop-out: None reported, but sample sizes for post-operative outcomes suggest there were no drop outs.

General comments
Generalisability: Pre-pubertal predominantly male, probably Chinese, paediatric outpatient population with ASA health status <3, who received general anaesthesia with sevoflurane. Not identified as being at high risk of intraoperative awareness.
Inter-centre variability: Not applicable (appears to be one centre).
Conflict of interests: 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	Unclear	Stated that both anaesthesiologists were blinded to the

personnel		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 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 post-operative outcomes suggest there were no drop outs.
Reporting bias		
Selective reporting.	Low	No evidence to suggest reporting bias

Appendix 5 (continued)

Reviewer 1: JS		Reviewer 2: GF	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 1315 **Author:** Messieha **Year:** 2004 **Study design**: RCT **Number of centres:** 1 (presumed) **Country**: USA **Sponsor**: Not stated	**Group 1:** "BIS known" - BIS (Aspect Medical Systems), no further detail given. Target device/index value: 60 to 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. Pre-medication used: ketamine 3mg/kg; midazolam 0.05 mg/kg; glycopyrrolate 0.2 mg, intra-muscular 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 1mg/kg Anti-nausea 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 general anaesthesia: 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**: Gender, male. N (%) Group 1 = 4 (40)	**Primary outcomes**: 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 assessment of other outcomes was blinded.

		Group 2 = 7 (70) (p=0.3) Age (yrs), mean (SD) Group 1 = 7.4 (± 3), range 3 to 13 years Group 2 = 5.5 (± 3), range 2 to 12 years. (p=0.2) Ethnic groups, n (%): Not reported. 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. Co-morbidities – cerebral palsy, n (%) Group 1 = 2 (20%) Group 2 = 2 (20%) (p=1.0) **Losses to follow up**: Not reported **Place of anaesthetic administration:** Presedation was given prior to transfer to the operating room. Upon transfer general anaesthesia was started.	

Definitions: ASA = American Society of Anesthesiologists; PACU = Post-Anaesthetic Care Unit

Outcome	**Group 1**	**Group 2**	**Group 3**
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, minutes. Mean (SD)	9 (± 5)	13 (± 5)	0.07

Time to PACU discharge, minutes. Mean (SD) Duration of PACU stay, minutes. Mean (SD)	60 (± 13) 45 (± 8)	90 (± 11) 71 (± 9)	<0.001 <0.001
Anaesthetic consumption	NR	NR	NR
Health related quality of life	NR	NR	NR
Nausea / vomiting / Anti-sickness drugs	NR	NR	NR
Pain / pain relieving drugs	NR	NR	NR
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

Definitions: 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: Not reported.
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 intention to treat: Not reported.
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: Not reported
Attrition/drop-out: Not reported

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.
Inter-centre variability: Not applicable (presumed to be one centre)
Conflict of interests: 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	Not reported
Performance bias		
Blinding of participants and personnel	Unclear	Not reported
Detection bias		
Blinding of outcome assessment	Unclear	BIS values recorded by blinded observer. Not clear whether assessment of other outcomes was blinded.

Attrition bias		
Incomplete outcome data	Unclear	Not reported
Reporting bias		
Selective reporting.	Low	No evidence to suggest selective reporting

Appendix 5 (continued)

Reviewer 1: JS		Reviewer 2: GF	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 1316 **Author:** Messieha et al **Year:** 2005 **Study design**: RCT **Number of centres:** 1 (presumed) **Country**: USA **Sponsor**: Not stated	**Group 1:** "BIS known" - BIS (Aspect Medical Systems), no further detail given. Target device/index value: 55 to 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 to 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 Pre-medication 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, IV administered at the start of the case. Muscle relaxants used: rocuronium bromide 1mg/kg, single dose administered at the beginning of the case. Reversal was administered at the end of the case (drug not stated). Anti-nausea drugs used: ondansetron 0.15 mg/kg, IV. 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 general anaesthesia: Not stated. **Inclusion criteria**: Aged 2 – 18 scheduled to undergo complete dental rehabilitation under general anaesthetic. Patients with mild cerebral palsy without significant neurological deficit also enrolled. **Exclusion criteria:** None stated.	**Primary outcomes**: 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 assessment of other outcomes was blinded.

		Baseline measurements: Gender male: female ratio Group 1 = 4:10 Group 2 = 2:3 (numbers not reported) Age (yrs), mean (SD) Group 1 = 4 (± 2) Group 2 = 4 (± 2) Ethnic groups, n (%): Not reported. 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. Co-morbidities – Children with mild cerebral palsy were eligible, but it is not stated how many were included. **Losses to follow up**: Not reported **Place of anaesthetic administration:** Presedation was given 15 to 20 minutes prior to transfer to the operating room. Upon transfer general anaesthesia was started.	

Definitions: ASA = American Society of Anesthesiologists; PACU = Post-Anaesthetic Care Unit; IV = intravenous

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 to extubation, minutes. Mean (SD)	5 (± 2)	10 (± 7)	0.04
Duration of PACU stay, minutes. Mean (SD)	47 (± 17)	63 (± 17)	0.02
Anaesthetic consumption	NR	NR	NR
Health related quality of life	NR	NR	NR
Nausea / vomiting / Anti-sickness drugs	NR	NR	NR
Pain / pain relieving drugs	NR	NR	NR
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

Definitions: 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: Not reported.
Blinding: Describes the study as observer-blind, but no other information provided. The observer recorded BIS values. Unclear whether the measurement of other outcomes was blinded.
Analysis by intention to treat: Not reported and not discernible (attrition not reported).
Comparability of treatment groups at baseline: Described by authors as comparable in terms of American Society of Anesthesiologists Physical Status, weight and gender.
Method of data analysis: Student's t test and Mann-Whitney rank sum test.
Sample size/power analysis: Not reported.
Attrition/drop-out: Not reported.

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.
Inter-centre variability: Not applicable (presumed to be one centre)
Conflict of interests: 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	Not reported
Performance bias		
Blinding of participants and personnel	Unclear	Not reported
Detection bias		
Blinding of outcome assessment	Unclear	BIS values recorded by blinded observer. Not clear whether assessment of other outcomes was blinded.

Attrition bias		
Incomplete outcome data	Unclear	Not reported
Reporting bias		
Selective reporting.	Low	No evidence to suggest selective reporting

Appendix 5 (continued)

Reviewer 1: GF		Reviewer 2:JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 101 **Author:** Rundshagen et al **Year:** 2007 **Study design**: RCT **Number of centres:** Not stated (appears to be single) **Country**: Not stated, appears to be Germany (multi-national 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/min then the remifentanil infusion ± 0.1 µg/kg/min. Commencement of monitoring: 5–10 min 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:**	**Total numbers involved**: 48 Group 1: 24; Group 2: 20 Pre-medication used: Midazolam 0.1 mg/kg orally, 45 min pre-surgery. General anaesthetic used (intravenous): Induction: Remifentanil 0.5 µg/kg/min continuous infusion followed 1 min later by target-controlled infusion of propofol, with an estimated plasma concentration 3 µg/ml. Maintenance: Remifentanil and propofol (doses not stated). FI_{O_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 2g for 20 min before and piritramide 7.5 mg for 5 min before the suggested end of surgery. Piritramide or morphine (doses not stated) as needed for early postoperative pain in the post-anaesthetic care unit (PACU). Muscle relaxants used: Rocuronium 0.6 mg/kg, before intubation. Anti-nausea 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: Not reported. Duration of general anaesthesia, min, 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.	**Primary (powered) outcome**: • Time to extubation **Secondary outcomes**: • Post-operative 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 post-operative 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). Post-operative nausea and fatigue was assessed after 10, 30 and 90 min in the PACU using a 100 mm visual analogue scale (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.

	Stated that all patients were treated by one experienced consultant anaesthetist; no details provided.	**Baseline measurements**: Gender, 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 (%): Not reported 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 Risk factors for awareness: None reported. Co-morbidities: None reported that would be likely to affect EEG (for other co-morbidities see additional comments). **Losses to follow up**: Not reported. Attrition reported but unclear whether pre- or post-randomisation. **Place of anaesthetic administration:** General anaesthesia was induced upon arrival in the operating room	

Definitions: ASA = American Society of Anesthesiologists; BIS: bispectral index NCT: Narcotrend index; PACU = post-anaesthetic care unit

Outcome	**Group 1**	**Group 2**	**p-value**
Intraoperative awareness / recall: Explicit memory during anaesthesia, n (%) Recalled dreaming during anaesthesia, n (%)	 0 (0) 2 (8)	 0 (0) 0 (0)	 NR NR
Patient distress and sequelae resulting from perioperative awareness	NR	NR	NR
Time to emergence from anaesthesia	NR	NR	NR
Time to extubation, min, 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/min, mean ± SD Remifentanil dose, µg/kg/min, mean ± SD Propofol plasma concentration, µg/ml, mean ± SD: [a] Intubation Skin incision Extubation 10 min after extubation 90 min after extubation	 0.093 ± 0.042 0.31 ± 0.10 3.7 ± 1.6 3.4 ± 1.5 1.5 ± 1.3 1.5 ± 1.6 0.9 ± 1.3	 0.114 ± 0.035 0.34 ± 0.11 2.9 ± 1.4 3.1 ± 1.2 1.5 ± 1.4 1.0 ± 0.9 0.7 ± 1.0	 0.089 0.449 >0.05 >0.05 >0.05 >0.05 >0.05
Health related quality of life	NR	NR	NR

Nausea / vomiting / Anti-sickness drugs: Nausea and fatigue visual analogue scale scores, mean ± SD: [b]			
Nausea, 10 min post-surgery	6.88 ± 15.2	24.06 ± 34.04	0.005
Nausea, 30 min post-surgery	15.44 ± 23.8	18.58 ± 24.9	0.146
Nausea, 90 min post-surgery	9.18 ± 19.0	12.00 ± 27.4	0.095
Fatigue, 10 min post-surgery	47.74 ± 20.7	45.31 ± 18.9	0.740
Fatigue, 30 min post-surgery	57.30 ± 22.4	46.32 ± 23.3	0.088
Fatigue, 90 min 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

Definitions: NR= Not reported

Additional Results/comments (e.g., early response factors, quality of life):

- Baseline data for patients' height, type of operation (peripheral/abdominal/thorax), and Apfel score (risk of post-operative 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 cafedrin (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/min 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 ($p \leq 0.034$ for both parameters combined), but the combined difference was not significant when age-corrected data were analysed.
- Co-morbidities 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: Not reported.

Blinding: Not reported.

Analysis by intention to treat: Unclear. Analysis does not include all the patients who started but it is unclear whether attrition happened pre-or post-randomisation.

Comparability of treatment groups at baseline: Groups were similar for the reported variables of gender, 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. Inter-group comparisons for propofol concentrations and visual analogue scores were tested by repeated-measures analysis of

variance or non-parametric statistics. Inter-group 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 α=5%, and d=1.0, the required sample size was estimated using a power table to be 13 subjects per group.
Attrition/drop-out: 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 whether the withdrawals occurred pre-or post-randomisation nor how they were distributed among the two study groups.

General comments
Generalisability: Appears to be a German adult population, predominantly of ASA grade II, but some grade I & III, with cardiovascular co-morbidities, undergoing various elective surgical procedures, and receiving propofol and remifentanil general anaesthesia. Ethnicity not reported. No explicit risk factors for intraoperative awareness identifiable.
Inter-centre variability: Not applicable (appears to be a single centre study).
Conflict of interests: None 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

Appendix 5 (continued)

Reviewer 1: GF		Reviewer 2: JB	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 8 **Author:** Talawar et al **Year:** 2010 **Study design**: RCT **Number of centres:** 1 **Country**: India **Sponsor**: Stated no external funding used	**Group 1:** 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 min 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:** Not reported	**Total numbers involved**: 50 Group 1: 25 Group 2: 25 Pre-medication used: None reported. General anaesthetic used: Induction: IV propofol 3–5 mg/kg for patients with an IV 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 litre 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: IV fentanyl 1 µg/kg (appears to be after insertion of the laryngeal mask airway). Maintenance: IV 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 IV boluses of fentanyl 0.5 µg/kg every 10 min until pain subsided. Muscle relaxants used: None used. Anti-nausea 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 general anaesthesia, minutes, median (range): Group 1: 68 (32–125); Group 2: 72 (47–180); difference p=0.23.	**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 • Post-operative 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 noninvasively. 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 the Children's Hospital of Eastern Ontario Pain Score (CHEOPS) (reference cited) every 30 min for the first 2 hours. Note non-independence of postoperative analgesia and postoperative pain scores (see left).

		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. **Baseline measurements**: Gender 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 (%): Not reported 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. Co-morbidities: None reported. **Losses to follow up**: None reported (all patients included in analysis) **Place of anaesthetic administration:** Operating room	

Definitions: ASA = American Society of Anesthesiologists

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 to emergence from anaesthesia, minutes: Recovery time (time to awakening), median (range) Recovery time (time to awakening), mean ± SD Time to reach Steward score of 6, median (range) Time to reach Steward score of 6, mean ± SD	 7 (3–18) 8.2 ± 4.49 6 (1–15) 7.08 ± 3.78	 10 (5–21) 10.96 ± 3.86 8 (2–24) 8.36 ± 4.8	 0.017 (2.72; 0.34–5.1) 0.464 (1.3; -1.2–3.7)
Time to extubation	NA	NA	NA
Time to discharge to / from the recovery room, minutes: Time to discharge for PACU, median (range) Time to discharge for PACU, mean ± SD	 15 (5–31) 15.32 ± 6.6	 19 (10–40) 19.32 ± 7.12	 0.045 (4.0; 0.07–7.9)

Anaesthetic (isoflurane) consumption, %, mean[a]:			
Immediately before laryngeal mask airway (LMA) 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 LMA removal	0.35	0.38	≥0.05
Health related quality of life	NR	NR	NR
Nausea / vomiting / Anti-sickness drugs	NR	NR	NR
Pain / pain relieving drugs			
Postoperative pain scores, mean (SE):			
30 minutes after admission to PACU	4.88 (0.319)	4.76 (0.09)	0.71 (0.12; -0.53–0.77)
60 minutes	4.48 (0.10)	4.76 (0.08)	0.01 (-0.28; 4.59–4.92) [b]
90 minutes	4.56 (0.10)	4.76 (0.08)	0.01 (-0.2; 4.59–4.92) [b]
120 minutes	4.88 (0.21)	5.44 (0.33)	0.01 (-0.56; 4.77–6.09) [b]
Required additional fentanyl intra-operatively, 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

Definitions: CHEOPS: Children's Hospital of Eastern Ontario Pain Score; ECG: electrocardiogram; IV: intravenous; LMA: laryngeal mask airway; MAC: minimum alveolar concentration; NA: not applicable; NR= Not reported; PACU: post-anaesthesia care unit

Additional Results/comments (e.g., early response factors, quality of life):

- 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 SE and RE 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 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 intention to treat: 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 2 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 χ^2 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 since 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 min. Anticipating a 5 min difference in recovery time between the study groups, with an α error of 0.05 and 90% power, a sample size of 15 in each group was calculated.
Attrition/drop-out: 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 IV propofol or inhaled sevoflurane, followed by maintenance under inhaled isoflurane. No specific risk factors for intraoperative awareness identified.
Inter-centre variability: Not applicable (one centre).
Conflict of interests: Stated none.

[a] mean estimated from graph by reviewer (95% CI was reported but has not been not extracted by the reviewer as it was stated to which group(s) or difference the CI applies)
[b] as reported: confidence interval 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 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 assessment of other outcomes was blinded.
Attrition bias		
Incomplete outcome data	Low	Analysis by intention to treat with no discernible attrition
Reporting bias		
Selective reporting.	Low	No evidence to suggest selective reporting

Appendix 5 (continued)

Reviewer 1: GF		Reviewer 2: JB	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 175 **Author:** Vakkuri et al **Year:** 2005 **Study design**: RCT **Number of centres:** 6 **Country**: Finland (3), Sweden (2), Norway (1) **Sponsor**: Technical assistance, financial support, and equipment for data collection and analysis for this study were provided by Datex-Ohmeda, Helsinki, Finland	**Group 1:** 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 (SE) between 45 and 65 until last 15 minutes of anaesthesia then ideally 65 (not exceeding 70) during last 15 minutes. Response-state entropy difference (RE-SE) <10. Heart rate and blood pressure to be kept within ±20% of baseline (pre-operative 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	**Total numbers involved**: 335 randomised (number randomised per group not reported). Numbers after attrition: Group 1: 160; Group 2: 160 Pre-medication 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/h and propofol ≤ 9 mg/kg/h in a mixture of O_2 (35-50%) and N_2O (50-65%). In Group 1, propofol was titrated to maintain the target SE; alfentanil or propofol boluses were permitted if SE suddenly increased; and alfentanil infusion was adjusted if the RE-SE 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: Not reported (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. Anti-nausea drugs used: None reported. Type of surgery: Different types of gynaecologic, abdominal, urologic, orthopaedic, breast, thyroid and inguinal hernia operations. Duration of surgery: Not reported. Duration of general anaesthesia,	**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 • Post-operative pain • Post-operative nausea and vomiting • Intraoperative awareness • Nurse estimation of post-operative 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 post-operative 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

	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.	minutes, mean ± SD: Group 1: 106 ± 48; Group 2: 107 ± 49; difference NS **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 to 150 minutes. **Exclusion criteria:** Known psychiatric or neurologic 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; body mass index >33 kg/m^2; any serious medical condition that would interfere with cardiovascular response assessment; cardiac, vascular or cranial neurosurgery; intra-operatively activated epidural analgesia; emergency or other non-elective surgery. **Baseline measurements** (reported only for analysed population after attrition; N=320); all differences stated NS: Gender 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 (%): Not reported 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. Co-morbidities: None reported (note extensive exclusion criteria for comorbid patients) **Losses to follow up**: Reported with reasons but not separable by study	Anaesthetic consumption: infusion rates of anaesthetics were noted manually in the anaesthetic record Drug consumption: noted manually in the anaesthetic record Pain scores: measured with a visual analogue scale (VAS) (no details given) Nausea and vomiting: measured with a visual analogue scale ‘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 post-operative day

		group **Place of anaesthetic administration:** Operating room	

Definitions: ASA = American Society of Anesthesiologists; NS: not statistically significant; PACU: post-anaesthesia care unit; RE: response entropy; SE: state entropy; VAS: visual analogue scale

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 to emergence from anaesthesia, minutes: Time to spontaneous breathing, median (range) Time to eyes open, median (range) Time to squeezes hand on command, median (range) Time to orientation to time and place, median (range)	 4.74 (0.00–18.0) 6.08 (0.15–37.5) 8.60 (1.17–47.4) 10.3 (1.17–48.7)	 7.07 (-1.00–28.5) 10.8 (2.23–43.2) 12.7 (2.43–48.1) 15.1 (4.08–113)	 <0.001 <0.001 <0.001 <0.001
Time to extubation, minutes, median (range)	5.80 (3.00–27.3)	9.16 (1.67–32.3)	<0.001
Time to discharge to / from the recovery room, minutes: Time to discharge from OR to PACU, median (range) Time to discharge from PACU, median (range)	 10.3 (3.83–42.4) 134 (50–1,293)	 13.0 (5.00–49.8) 150 (7–1,020)	 <0.001 0.21
Anaesthetic consumption: [a] Propofol, mg/kg/min, median (range) Alfentanil, µg/kg/min, median (range)	 0.10 (0.04–0.23) 0.60 (0.12–2.2)	 0.11 (0.03–0.21) 0.57 (0.16–1.6)	 <0.001 0.54
Health related quality of life	NR	NR	NR
Nausea / vomiting / Anti-sickness 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 Opioid analgesic requirements in the PACU	 NR NR	 NR NR	Both outcomes: stated no difference between groups
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

Definitions: NR= Not reported; OR = operating room

Additional Results/comments (e.g., early response factors, quality of life):

- Stated that the aim in all patients was to provide smooth, haemodynamically stable anaesthesia with the shortest possible emergence time and without intra-operative awareness.
- The initial 8–9 patients in 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 gynaecologic 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 versus 60 ± 10 beats/minute; p=0.029) and blood pressure (83 ± 10 versus 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 compared to 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 opaque). With sequential coding, the subjects were treated in blocks of 10 (5 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 intention to treat: No: 15 patients excluded after randomisation were omitted from the analysis.
Comparability of treatment groups at baseline: Ethnicity was not reported but age, gender, 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 α of 0.05.
Attrition/drop-out: 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 2 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 due to '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 general anaesthesia with alfentanil and propofol. Population noted not to be at particular risk of intraoperative awareness.
Inter-centre 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

Appendix 5 (continued)

Reviewer 1: JB		Reviewer 2: GF	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 71 **Author:** Wu et al **Year:** 2008 **Study design**: RCT **Number of centres:** 1 **Country**: Taiwan **Sponsor**: Supported in part by the National Science Council	**Group 1: Entropy** RE and SE values shown on GE Datex-Ohemda S/5^{TM} Anaesthesia Monitor. Target device/index value: RE and SE target values 35-45, corresponding to stable 2% EtSevo in the absence of major surgical stimulation. Gradient between RE and SE within 5-10. Anaesthesia monitored by entropy unless haemodynamic changes of 30% persisted for more than 5 mins. **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 less than 30%. Commencement of monitoring: In the operation room (appears to be before induction, though not explicitly stated so). **Length of experience / training of anaesthetist:** Not reported	**Total numbers involved**: 68 patients enrolled and randomised; data for 65 Group 1 = 34 Group 2 = 31 Pre-medication used: none reported General anaesthetic used: Sevoflurane as sole inhalational anaesthetic Induction – fentanyl 2µg/kg, propofol 2mg/kg and 2mL of 2% lidocaine. Maintenance: After intubation sevoflurane delivered in a mixed flow of 0.3L/min air and 0.7L/min 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 Anti-nausea drugs used: not reported Other drugs used: Hypertension treated with nicardipine 0.25 mg (heart rate<90/min) or labetolol 2.5mg (heart rate >90/min). Ephedrine 4 mg to treat hypotension (MAP <70% of baseline). Atropine 0.5mg IV bolus for bradycardia (heart rate <45/min). Type of surgery: Total knee replacement Duration of surgery: approx. 1.5 hrs Duration of general anaesthesia, mean ± SD: Group 1 = 133.74±30 min Group 2 = 144.84±30 min **Inclusion criteria**: ASA I or II scheduled to undergo total replacement **Exclusion criteria:** History of cerebrovascular disease, treatment with psychoactive medication, existing cardiac dysrhythmia or weight less than 70% or more than 130% of ideal body	**Primary outcomes**: Consumption of sevoflurane **Secondary outcomes**: Tourniquet-induced hyperdynamic responses; Pain status in the PACU; Post-operative nausea and vomiting; Level of awareness; Subjective complaints; Post-operative analgesic needs; Device values; Haemodynamic parameters **Length of follow up**: 72 hours postoperative for post-operative nausea and vomiting (follow up for level of awareness and other outcomes unclear) **Methods of assessing outcomes:** Consumption of sevoflurane determined by GE Datex Ohemda S/5^{TM} Anaesthetic Delivery Unit System Physiologic changes at 5 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.

		weight. **Baseline measurements**: Gender male, n (%): Group 1 = 28 (82%) Group 2 = 25 (81%) Age (yrs), mean (SD): Group 1 = 68.03 (6.1) Group 2 = 68.90 (6.5) Ethnic groups, n (%): Not reported 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: Not reported **Losses to follow up**: reported with reasons Group 1=0, Group 2=3 **Place of anaesthetic administration:** operation room	

Definitions: ASA = American Society of Anesthesiologists; EtSevo = end-expiratory concentration of sevoflurane; iv = intravenous; PACU = post-anaesthesia care unit; MAP = mean arterial pressure

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, sevoflurane mL, mean (SD)	27.79 (7.4)	31.42 (6.9)	P=0.023
Health related quality of life	NR	NR	
Nausea / vomiting / Anti-sickness drugs Post-operative nausea and vomitings	No statistically significant difference between groups		NR
Pain / pain relieving drugs Post-operative pain status and analgesic use	No statistically significant difference between groups		NR
Mortality	NR	NR	NR

Definitions: NR= Not reported; n/a = not applicable

Additional Results/comments:

- No cardiovascular or cerebrovascular complication in any patient of either group post-operative.
- 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 min 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 intention to treat: No (not all randomised patient analysed)
Comparability of treatment groups at baseline: stated no statistically significant differences in age, gender, ASA physical status, height, and weight.
Method of data analysis: For nominal data, statistical analysis performed using Chi squared test. Age, gender, 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: not reported
Attrition/drop-out: 3 patients from group 2 not included in results due to missing data (reasons no 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 intra-operative awareness identified.
Inter-centre variability: Not applicable – assumed single centre
Conflict of interests: 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	3 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)

Reviewer 1: GF		Reviewer 2: JS	
Reference and Design	**Technology**	**Participants**	**Outcome measures**
Ref ID: 1472 **Author:** Zhang et al. **Year:** 2011 (enrolment November 2008 – November 2010) **Study design**: RCT **Number of centres:** 13 **Country**: China **Sponsor**: Not reported (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: Not reported **Length of experience / training of anaesthetist:** Not reported	**Total numbers involved:** Number randomised not reported. Stated 5309 provided outcome data but only 5228 were analysed (Group 1 = 2919; Group 2 = 2309). *Pre-medication 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:* Not reported *Analgesia used:* Drugs and doses at the discretion of the anaesthetist *Muscle relaxants used:* Drugs and doses at the discretion of the anaesthetist *Anti-nausea drugs used:* Not reported *Other drugs used:* Not reported 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 (ortho-paedic) 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 general anaesthesia: Not reported **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**:	**Primary outcome**: Intraoperative awareness **Secondary outcomes**: 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 1st and 4th 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).

		Gender 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 (%): Not reported; 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. Co-morbidities: Not reported **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 5309 was the total number randomised. **Place of anaesthetic administration:** Not reported	
Definitions: ASA = American Society of Anesthesiologists; TIVA = total intravenous anaesthesia			

Outcome	**Group 1**	**Group 2**	**p-value; OR (95% CI)**
Intraoperative awareness / recall, n (%) Confirmed awareness Possible awareness Confirmed or possible awareness	 4/2919 (0.14) 4/2919 (0.14) 8/2919 (0.27)	 15/2309 (0.65) 6/2309 (0.26) 21/2309 (0.9)	 0.002; OR=0.21 (0.07 – 0.63) 0.485 p<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
Health related quality of life	NR	NR	NR
Nausea / vomiting / Anti-sickness drugs	NR	NR	NR

Pain / pain relieving drugs	NR	NR	NR
Other morbidity (e.g. cognitive dysfunction)	NR	NR	NR
Mortality	NR	NR	NR

Definitions: NR= Not reported

Additional Results/comments (e.g., early response factors, quality of life):

- 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: Not reported.
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 intention to treat: 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, gender, 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 inter-group comparisons and also χ^2 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% 2-sided type I error. To allow for missing data, 5000-6000 patients were recruited.
Attrition/drop-out: 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 post-operative interviews (n=2 in Group 1; n=2 in Group 2); post-operative 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.
Inter-centre variability: Not reported
Conflict of interests: Device manufacturer (Aspect Medical Systems) provided BIS electrodes.

[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 due to rounding errors).

Domain	Author's judgement	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 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

Health technology

Bispectral index monitoring (BIS)

Interventions and comparators

What interventions/ strategies were included?

General anaesthesia (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 $9,000
Cost of each BIS electrode sensor was $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:

NA

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)?

NA

List the utility values used in the evaluation

NA

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).

NA

What is the model time horizon?

NA

What, if any, discount rates have been applied in the model? Same rate for costs and outcomes?

NA

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, ie $1.07 per patient. Thus cost of BIS monitoring is $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$4,410 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, quality of life 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?

NA

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)

	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	

Yes / No / ? (unclear)

Appendix 6 (continued)

Study Characteristics

Reference

Satisha, 2010

Health technology

Bispectral index monitoring (BIS)

Interventions and comparators

What interventions/ strategies were included?

General anaesthesia (GA) with BIS

Was a no treatment/ supportive care strategy included?

GA without BIS

Research question

What are the stated objectives of the evaluation?

To determine the effects of introducing BIS at all potential sites of GA administration in practice on recovery from GA, the incidence of postoperative nausea and vomiting and expenditure on anaesthetic drugs.

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?

Adult only on operating lists in the main theatre suite.

Institutional setting Where is/are the intervention(s) being evaluated usually provided?

Medway Maritime Hospital

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?

UK, £. 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

Data derived from this study.

	Phase 1 – before introduction of BIS monitoring (n=427)	Phase 2 – after introduct BIS monitoring (n=299)
Nausea in PACU	63	22
Vomiting in PACU	26	6
Overall recovery time	34	35
Sevoflurane (GA agent)	231	158
Desflurane (GA agent)	120	83

Other GA agents shown in Table 1.

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.

Intervention costs derived from this study.
Reduction in drug costs per patient of £0.86.
Cost of each semi-reusable BIS sensor was £6.60.
Cost of BIS monitoring $16 per patient.

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:

NA

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)?

NA

List the utility values used in the evaluation

NA

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 model

Extract transition probabilities for [natural history/disease progression] model and show sources (or refer to table in text).

NA

What is the model time horizon?

NA

What, if any, discount rates have been applied in the model? Same rate for costs and outcomes?

NA

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

The clinical outcome was based upon the B-Aware study that showed that the use of BIS monitoring in patients at high risk of awareness reduced the incidence from 0.91 to 0.17%. The authors calculated that if BIS sensors were used in all patients at high risk of awareness the number needed to treat one episode of awareness would be 138.

Provide a summary of the costs estimated for each intervention/ strategy assessed in the evaluation

See above section on intervention costs.

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$2200.

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, quality of life 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?

NA

Conclusions/ Implications

Give a brief summary of the author's conclusions from their analysis

The cost of implementing BIS may be less than the cost of not preventing the episode due to potential lawsuits. Results from the prospective audit suggest that some of the advantages demonstrated for BIS monitoring in RCTs may not be seen to the same degree in everyday clinical practice.

What are the implications of the evaluation for practice?

Not stated

SHTAC Commentary

This study is a simple calculation based on the results of an audit of BIS monitoring. As such the economic evaluation is of poor quality.

Critical appraisal checklist of economic evaluation (Questions in this checklist based on Philips et al)

	Item	Satisha	Comments
1	Is there a clear statement of the decision problem?	?	Aim of study is prospective trial to determine whether effects of intraoperative BIS can be replicated in everyday clinical practice.
2	Is the comparator routinely used in UK NHS?	Y	UK study
3	Is the patient group in the study similar to those of interest in UK NHS?	Y	UK study
4	Is the health care system comparable to UK?	Y	UK study
5	Is the setting comparable to the UK?	Y	UK study
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 BIS monitoring cost is 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?	?	Partly. Unclear where the BIS monitoring cost is from
16	Have the costs and outcomes been discounted?	N	
17	Has uncertainty been assessed?	N	
18	Has the model been validated?	N	

Yes / 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. Ontario Health Technology Assessment Series 2004; Vol. 4, No. 9	Not full economic evaluation
Hayes Inc. Bispectral index monitoring for anesthesia awareness. 2005.	Unobtainable
Bard JW. The BIS monitor: a review and technology assessment. *AANA Journal* 2001; 69(6):477-483.	Review
Lehmann A, Karzau J, Boldt J, Thaler E, Lang J, Isgro F. Bispectral index-guided anesthesia in patients undergoing aortocoronary bypass grafting. *Anesthesia & Analgesia* 2003; 96(2):336-343.	Wrong comparator
Liu SS. Effects of Bispectral Index monitoring on ambulatory anesthesia: a meta-analysis of randomized controlled trials and a cost analysis. *Anesthesiology* 2004; 101(2):311-315.	Cost analysis
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. *Anesthesia and Analgesia* 2007; 104(5):1145-1149.	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, randomized, controlled trial of costs and outcomes. *Anesthesia and Analgesia* 2002; 95(4):805-12, table.	Not full economic evaluation
Nielsen JS, Thogersen B, Ording H. Monitoring depth of anaesthesia - a health technology assessment. 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 anesthesia Utilidad del indice biespectral (BIS) en la reduccion del costo de farmacos para la anestesia. *Cirugia y Cirujanos* 2003; 71(4):300-303.	Not English language (Spanish)
White PF, Tang J, Ma H, Wender RH, Sloninsky A, Kariger R. Is the patient state analyzer with the PSArray2 a cost-effective alternative to the bispectral index monitor during the perioperative period? *Anesthesia & Analgesia* 2004; 99(5):1429-1435.	Not full economic evaluation
Windisch PA, Worsham GM. The effect of the bispectral index on medication utilization in the operating room and time to discharge from the postanesthesia care unit. *Hospital Pharmacy* 2002; 37(4):386-390.	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 Anaesthesiologica Scandinavica* 1999; 43(5):545-549.	Not full economic evaluation

Appendix 8: Pooled intravenous anaesthetic consumption for Narcotrend RCTs

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/hr) 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 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
			Lower	Upper			
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

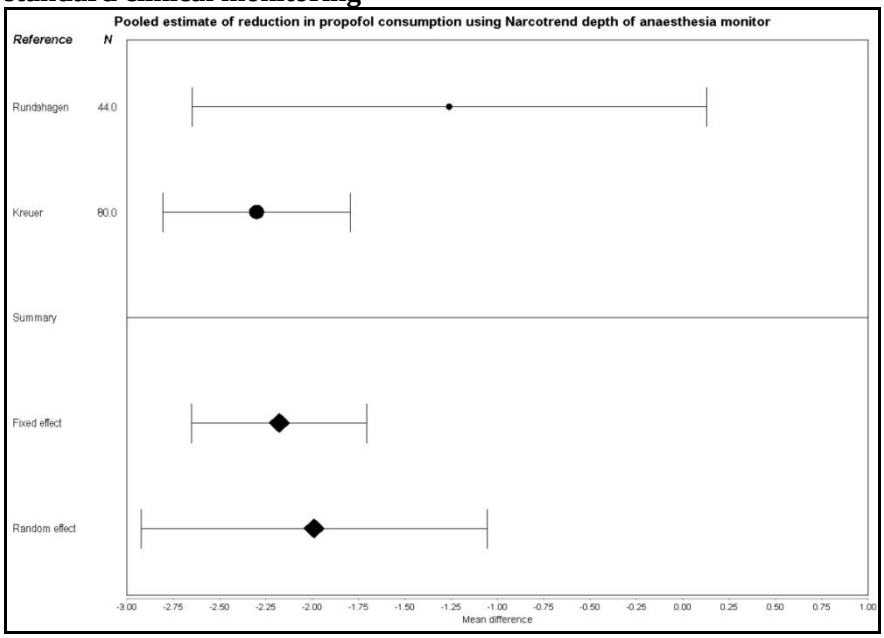


Table 123 reports the normalised remifentanil consumption (μg/kg/hr) and mean difference in each of the included trials. Pooled estimates for the mean difference are reported in Table 124 (Figure 11 presents a forest plot for the analysis).

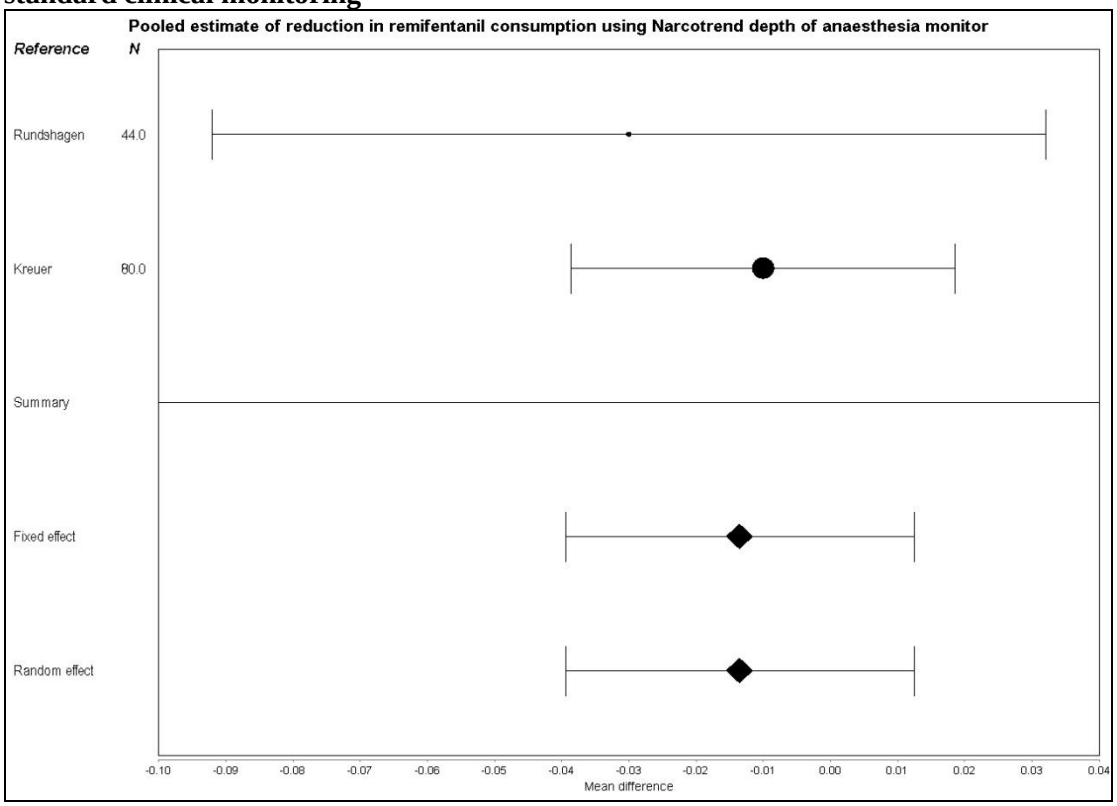
Table 123 Remifentanil consumption in RCTs using Narcotrend depth of anaesthesia monitoring

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
			Lower	Upper			
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

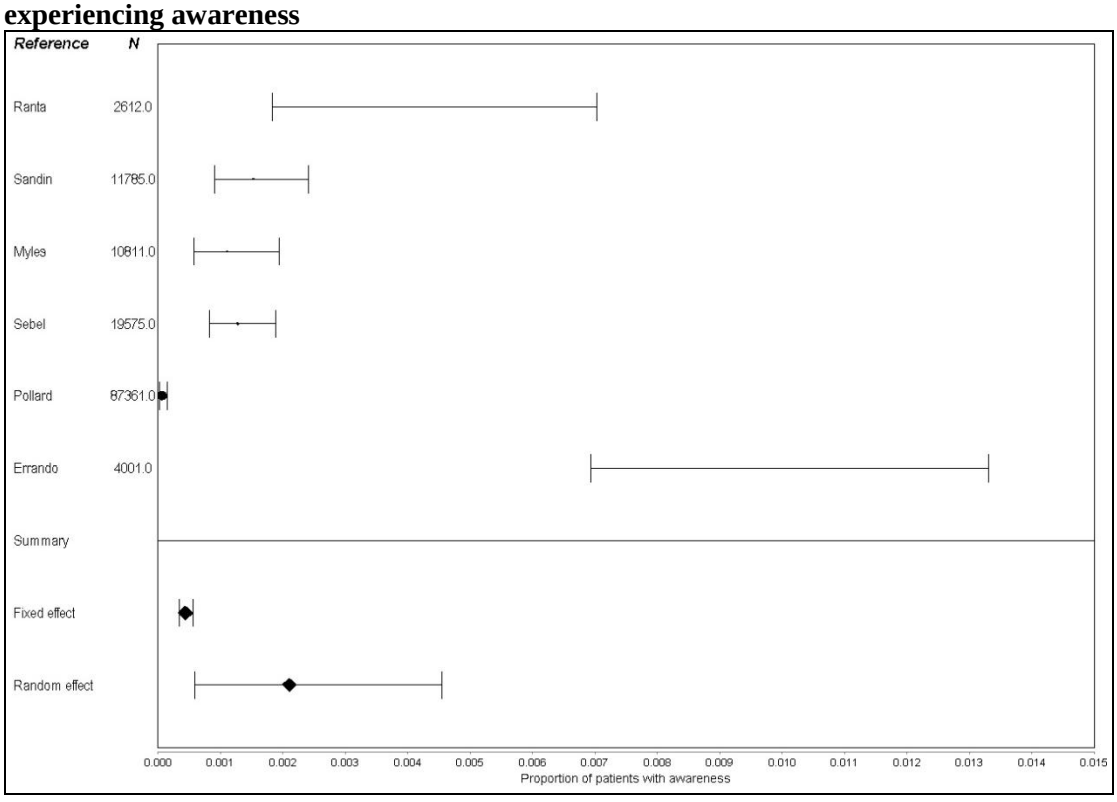


Appendix 9: Derivation of the pooled estimates of cumulative incidence of awareness used in the model

Table 39 in section 5.3.2 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 effects model and DerSimonian-Laird weights for the random effects model.

Figure 8.1 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 Q = 212.55 (df=5), p=0.0000, I^2 = 97.6% for fixed effect model).

Figure 8.1 Forest plot for pooled estimate of proportion of general surgical patients experiencing awareness



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 Q = 7.85 (df=3), p=0.0493, I^2 = 61.8% for fixed effect model]). (Figure 8.2)

Figure 8.2 Forest plot for pooled estimate of proportion of general surgical patients experiencing awareness (excluding outliers)

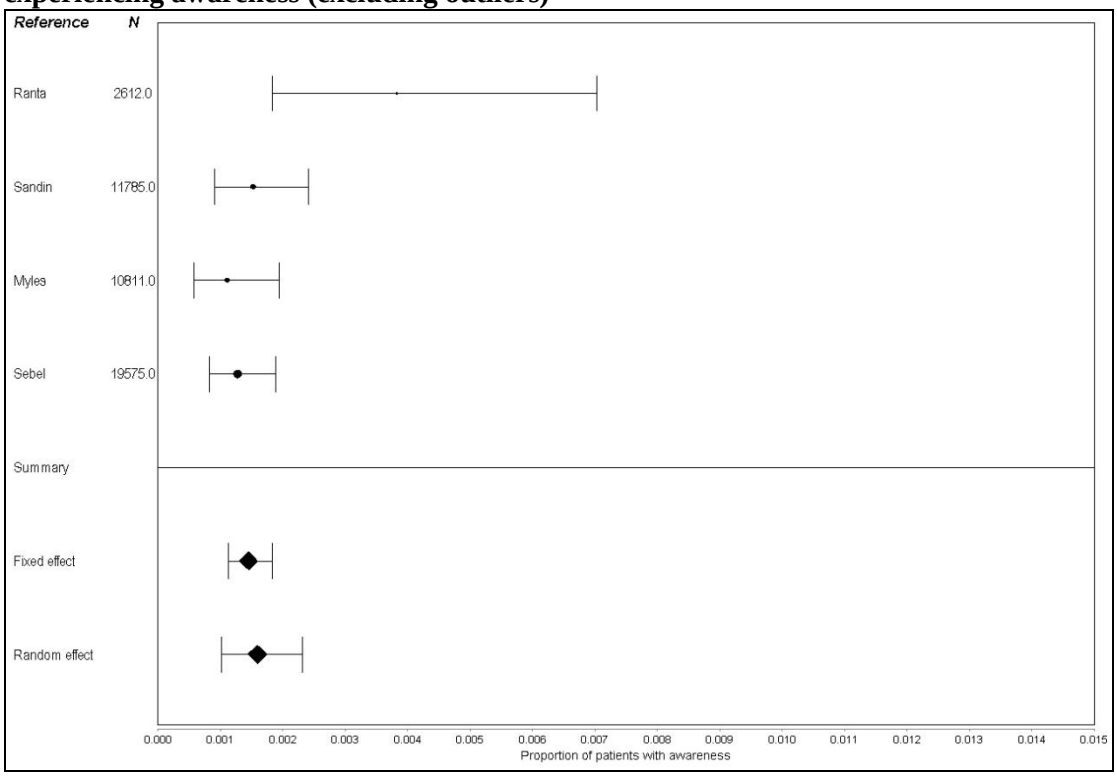
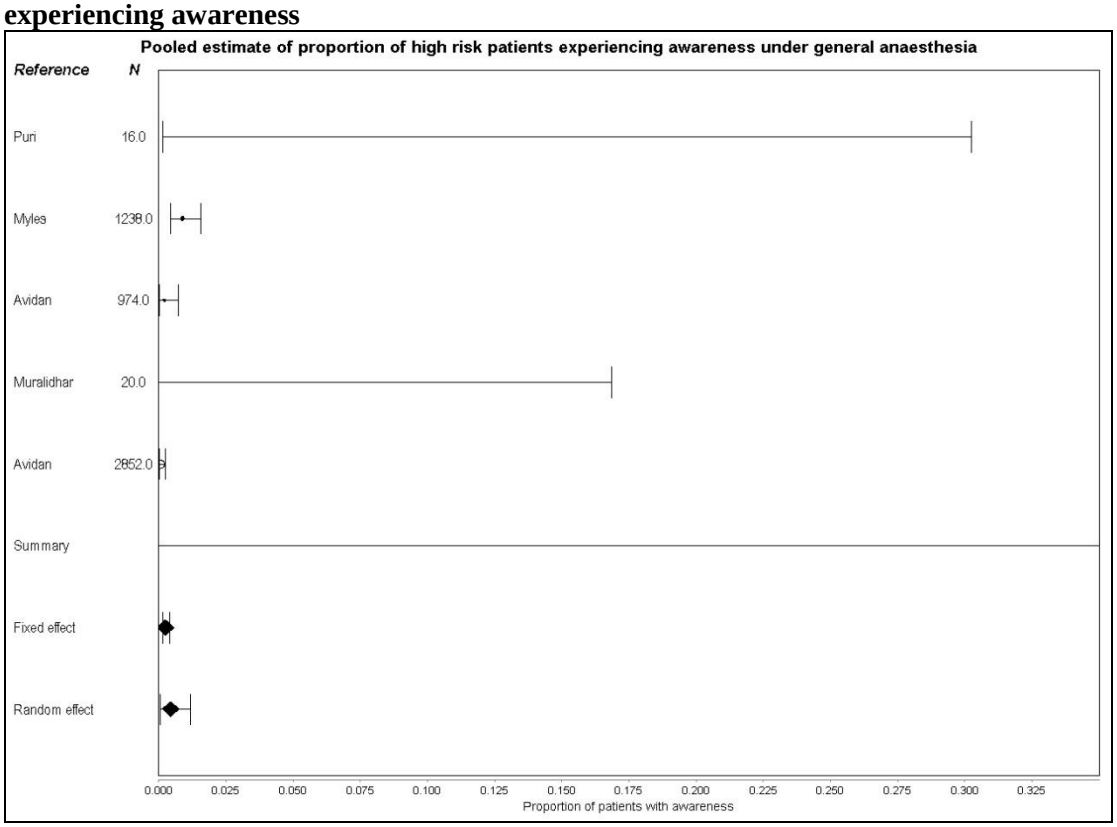


Figure8.3 shows the forest plot for studies in high risk surgical populations. A pooled estimate from all these studies gives a cumulative incidence of awareness of 0.45% (95% CI 0.06% to 1.19%) assuming random effects (Cochran Q = 19.97 (df=4), p=0.0005, I^2 = 80.0% for fixed effect model).

Figure 8.3 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

$$S(t) = \exp(-\lambda\, t^{\gamma})$$

with scale parameter λ and shape γ. Taking the log of both sides gives

$$\log(S(t)) = -\lambda\, t^{\gamma}$$

Taking the log of both sides again, gives

$$\log(-\log(S(t))) = \log(\lambda) + \gamma \log(t)$$

which is a linear function and can be fit using least squares methods to provide estimates of log(λ) 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. The process of extracting data from a chart usually begins with the user identifying key reference points on the chart (for example 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.

The following table reports the parameter estimates for linear regression for the Weibull survival function.

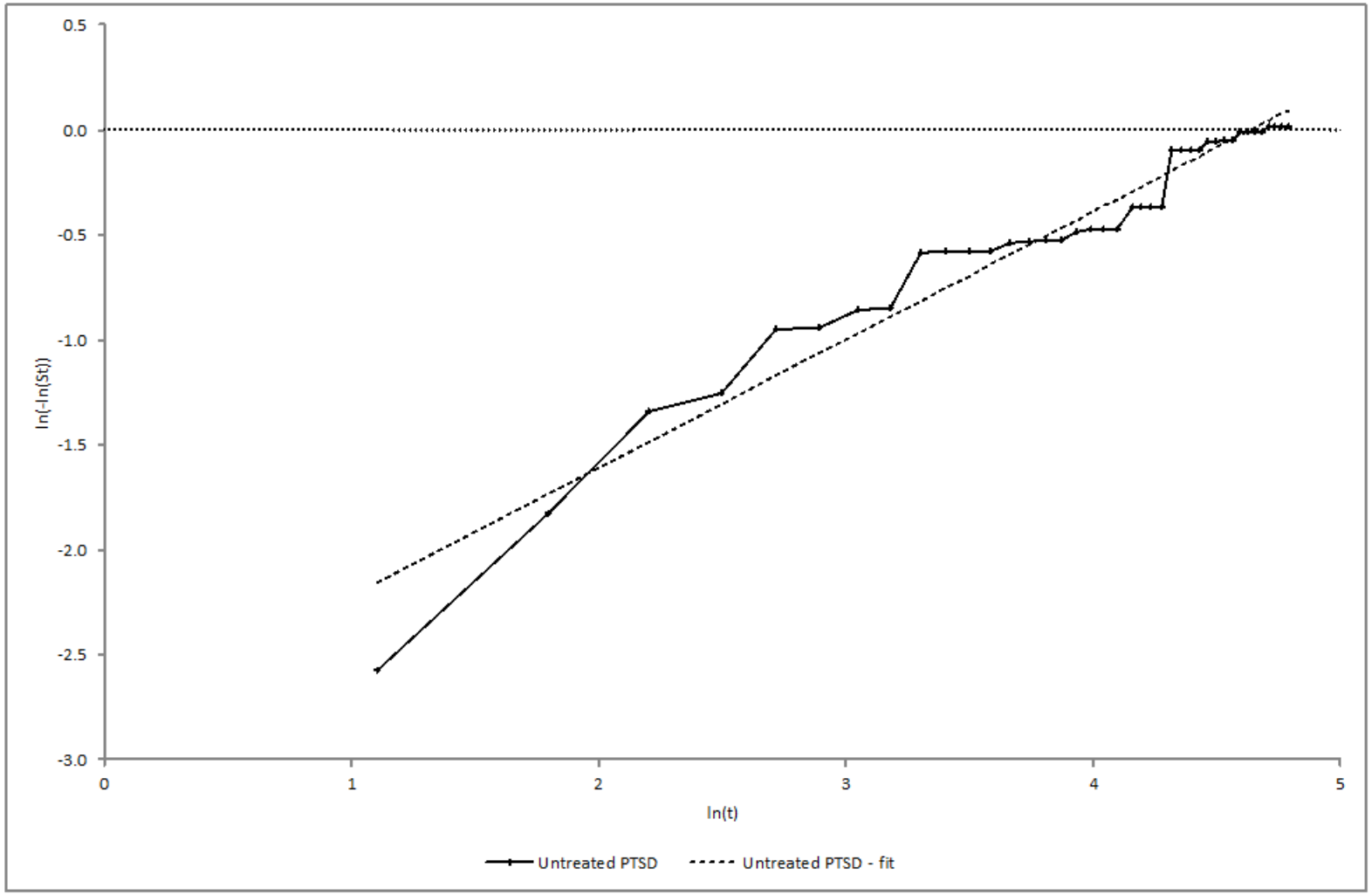
	log(λ)	γ
Weibull	-2.82786	0.61006

The mean duration of symptoms can be estimated using the following equation[154]

$$(1/\lambda)^{(1/\gamma)} \times \Gamma[1 + (1/\gamma)]$$

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 9.1 Transformed survival curve for duration of PTSD symptoms and linear fit



Appendix 11: Search strategy to identify utility values for PTSD

Specific PTSD and QOL search
Database: Ovid MEDLINE(R) <1948 to November Week 3 2011> 06/12/2011
Also run on MEIP, Science direct searched for HRQOL terms linked to posttraumatic stress disorder 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/ (17509)
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$.mp. (73060)
37 linear scal$.mp. (463)
38 linear analog$.mp. (776)
39 visual analog$.mp. (23714)

40 (categor$ adj2 scal$).mp. (1028)
41 or/1-40 (145653)
42 (letter or editorial or comment).pt. (1103139)
43 41 not 42 (141638)
44 Stress Disorders, Post-Traumatic/ (17609)
45 "posttraumatic stress".tw. (8436)
46 "post traumatic stress".tw. (4553)
47 PTSD.tw. (8895)
48 or/44-47 (20455)
49 43 and 48 (2564)
50 or/2-35 (47102)
51 48 and 50 (253)
52 HRQOL.tw. (4851)
53 "health related quality of life".tw. (15264)
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. (11857)
61 standard gamble*.tw. (577)
62 or/13-30 (11699)
63 or/52-61 (28623)
64 62 or 63 (32460)
65 48 and 64 (222)
66 51 or 65 (316)
67 (visual adj analogue adj scale*1).tw. (11051)
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 Pediatric 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.
Intra-operative depth of anaesthesia and influence on the incidence of post-operative 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 general anaesthesia with a planned duration of procedure greater than or equal to 1 hour	Unblinded BIS monitoring (anaesthetic not specified)	Blinded BIS monitoring (anaesthetic not specified)	**Primary:** Post-operative delirium incidence (DSM-IV); **Secondary:** device values; post-operative delirium (alternative Delirium scores); post-operative cognitive dysfunction; time to discharge (recovery room; hospital); length of stay (recovery room; hospital); quality of life (EQ-5D); organ dysfunction at hospital discharge; post-operative 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