Technology Assessment

Technology Assessment Program

Esophageal Doppler Ultrasound-Based Cardiac Output Monitoring for Real-Time Therapeutic Management of Hospitalized Patients

A Review

Agency for Healthcare Research and Quality
540 Gaither Road
Rockville, Maryland 20850

January 16, 2007

Esophageal Doppler Ultrasound-Based Cardiac Output Monitoring for Real-Time Therapeutic Management of Hospitalized Patients

A Review

This report is based on research conducted by the ECRI Evidence-based Practice Center (EPC) under contract to the Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. 290-02-0019). The findings and conclusions in this document are those of the author(s) who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ. Therefore, no statement in this report should be construed as an official position of the Agency for Healthcare Research and Quality or of the U.S. Department of Health and Human Services.

The information in this report is intended to help health care decision-makers; patients and clinicians, health system leaders, and policymakers, make well-informed decisions and thereby improve the quality of health care services. This report is not intended to be a substitute for the application of clinical judgment. Decisions concerning the provision of clinical care should consider this report in the same way as any medical reference and in conjunction with all other pertinent information, i.e., in the context of available resources and circumstances presented by individual patients.

This report may be used, in whole or in part, as the basis for development of clinical practice guidelines and other quality enhancement tools, or as a basis for reimbursement and coverage policies. AHRQ or U.S. Department of Health and Human Services endorsement of such derivative products may not be stated or implied.

January 16, 2007

Prepared for:
Agency for Healthcare Research and Quality
540 Gaither Road
Rockville, Maryland 20850

Table of Contents

TABLES

FIGURES

EXECUTIVE SUMMARY

The Centers for Medicare & Medicaid Services (CMS) requested that AHRQ commission an evidence report to inform CMS of the evidence regarding ultrasound-based monitoring of cardiac output (Doppler). This is currently listed as a Category II procedure (clinical reliability and efficacy not proven), which is "considered experimental and should not be covered at this time". Accordingly, on May 9th 2006, AHRQ, in consultation with CMS and ECRI, issued a Statement of Work (SOW) contracting ECRI to prepare an evidence report on this topic.

In commissioning this report, AHRQ, in consultation with CMS and ECRI, developed four Key Questions to be addressed. These four Key Questions are presented below.

Key Question 1: **What types of devices/techniques are currently used to assess cardiac output?**

Key Question 2: **Does therapeutic management based on esophageal Doppler ultrasound-based cardiac output monitoring during *surgery* lead to improved patient outcomes (fewer complications and shorter hospital stay), compared to:**

- ***a.*** **Pulmonary artery catheter-based measurement of cardiac output via thermodilution?**
- ***b.*** **Catheter-based measurement of central venous pressure?**
- ***c.*** **Conventional clinical assessment (physical examination, fluid input and output measurements)?**

Key Question 3: **Does therapeutic management based on esophageal Doppler ultrasound-based cardiac output monitoring during *hospitalization* lead to improved patient outcomes (fewer complications and shorter hospital stay), compared to:**

a. **Pulmonary artery catheter-based measurement of cardiac output via thermodilution?**

b. **Catheter-based measurement of central venous pressure?**

c. **Conventional clinical assessment (physical examination, fluid input and output measurements)?**

Key Question 4: **What complications, harms, and adverse events associated with esophageal Doppler ultrasound-based monitoring have been reported?**

Data Sources

We searched 17 external and internal databases, including PubMed and EMBASE, for clinical trials on the use of esophageal Doppler ultrasound for cardiac output monitoring. In addition, we routinely reviewed more than 1,600 journals and supplements maintained in ECRI's collections to determine if they contained relevant information. We also examined the bibliographies/reference lists from peer-reviewed and gray literature. (Gray literature includes reports and studies produced by local government agencies, private organizations, educational facilities, and corporations that do not appear in the peer-reviewed journal literature.) Although we examined gray literature sources to identify relevant information, we only consider published, peer-reviewed literature in this report.

Evidence Bases

Our searches identified 317 potentially relevant articles. Of these, we retrieved 75 full-length articles. We read each article in full to determine whether it met a set of general and question-specific *a priori* inclusion criteria. Twenty-seven of the retrieved articles met the inclusion criteria for at least one key question. Four of the 27 included articles addressed more than one of our four key questions. Key Question 1 is not evidence-based; instead, it requires a summary of the technologies currently used to measure cardiac output. The information addressing this question is derived primarily from review articles written by experts in the field. The evidence base for

Key Question 2 consisted of seven studies, the evidence base for Key Question 3 consisted of one study, and the evidence base for Key Question 4 consisted of 23 studies.

Main Findings and Conclusions

Key Question 1: **What types of devices/techniques are currently used to assess cardiac output?**

Several methods are currently used to monitor cardiac output in patients during surgery or intensive care. These methods include thermodilution, dye dilution, lithium dilution, methods using the Fick principle, pulse contour methods, thoracic electrical bioimpedance, transesophageal echocardiography, esophageal Doppler monitoring, and ultrasonic cardiac output monitoring.

Key Question 2: **Does therapeutic management based on esophageal Doppler ultrasound-based cardiac output monitoring during surgery lead to improved patient outcomes (fewer complications and shorter hospital stay), compared to:**

- ***a.*** **Pulmonary artery catheter-based measurement of cardiac output via thermodilution?**
- ***b.*** **Catheter-based measurement of central venous pressure (CVP)?**
- ***c.*** **Conventional clinical assessment (physical examination, fluid input and output measurements)?**

After searching the literature, retrieving articles, and applying the inclusion/exclusion criteria, we identified seven publications with 583 patients that addressed this question. None of these studies compared esophageal Doppler monitoring to thermodilution. Five studies compared esophageal Doppler with CVP plus conventional clinical assessment to CVP plus conventional clinical assessment. The median quality of these five studies was high, and the age generalizability to the Medicare population was fair. Two studies compared esophageal Doppler with conventional clinical assessment to

conventional clinical assessment alone (one of these studies also compared esophageal Doppler with conventional clinical assessment to CVP with conventional clinical assessment). The median quality of these two studies and their generalizability in terms of patient age to the Medicare population was high.

The addition of esophageal Doppler monitoring for guided fluid replacement to a protocol using CVP and conventional clinical assessment during surgery leads to a clinically significant reduction in the rate of major complications and total complications in surgical patients compared to CVP plus conventional clinical assessment. The strength of evidence supporting this finding is strong. Because only three of five studies separately reported major complications, and because of differences in the way total complications were reported, no quantitative conclusion is presented for these outcomes.

The addition of esophageal Doppler monitoring to the protocol described above also reduces the length of hospital stay for surgical patients (clinical significance uncertain). The strength of evidence supporting this finding is strong. The lack of a calculable precise effect size in some studies precluded a quantitative summary estimate of the reduction in length of stay.

Only one study compared esophageal Doppler plus conventional clinical assessment to CVP plus conventional clinical assessment. Because this was one small study with non-informative effect sizes, no evidence-based conclusions were possible for any of the outcomes of interest.

The addition of esophageal Doppler monitoring for guided fluid replacement to conventional clinical assessment during surgery leads to a clinically significant reduction in the length of hospital stay compared to that associated with conventional clinical assessment alone. The strength of evidence supporting this finding is weak. The low number of studies (two) precluded a quantitative estimate of the reduction in length of hospital stay. Because only a single study reported total complications, no conclusion was possible concerning this outcome.

The evidence was insufficient to allow conclusions to be reached concerning relative mortality rates for any of the comparisons in Key Question 2.

The conclusions for Key Question 2 only apply to patients undergoing surgical procedures with an expected substantial blood loss or fluid shifts requiring fluid replacement.

Key Question 3: **Does therapeutic management based on esophageal Doppler ultrasound-based cardiac output monitoring during hospitalization lead to improved patient outcomes (fewer complications and shorter hospital stay), compared to:**

a. **Pulmonary artery catheter-based measurement of cardiac output via thermodilution?**

b. **Catheter-based measurement of central venous pressure?**

c. **Conventional clinical assessment (physical examination, fluid input and output measurements)?**

After searching the literature, retrieving articles, and applying the inclusion/exclusion criteria, we identified one study that compared esophageal Doppler monitoring plus CVP plus conventional clinical assessment to CVP plus conventional clinical assessment for optimization of intravenous fluid replacement in patients admitted to intensive care following cardiac surgery. This study was judged to be of high quality based on ECRI ratings. Generalizability to the Medicare population was fair. However, this was a single small study without a demonstrably large treatment effect on the outcomes of interest. Therefore, no conclusions could be reached for this question.

Key Question 4: **What complications, harms, and adverse events associated with esophageal Doppler ultrasound-based monitoring have been reported?**

Currently, no serious adverse events associated with esophageal Doppler probes have been reported in the literature or in adverse event databases. The only minor events identified included two cases of incorrect probe placement in the left main bronchus,

one case of incorrect placement in the trachea, a tube displacement during probe removal, and an unspecified number of cases of minimal trauma in the buccal cavity during probe placement. Nineteen studies with a total of 654 patients specifically stated that esophageal Doppler probes did not cause any complications. The number of patients represented in these studies is relatively small. However, the available evidence suggests that esophageal Doppler probes are relatively low-risk devices, as reporting of even minor morbidity has been infrequent thus far.

SCOPE OF REPORT

The Centers for Medicare & Medicaid Services (CMS) requested that AHRQ commission an evidence report to inform CMS of the evidence regarding ultrasound-based monitoring of cardiac output (Doppler). Accordingly, on May 9th 2006, AHRQ, in consultation with CMS and ECRI, issued a Statement of Work (SOW) contracting ECRI to prepare an evidence report on this topic. AHRQ, in consultation with CMS and ECRI, developed four Key Questions to be addressed. These questions are as follows:

1. What types of devices/techniques are currently used to assess cardiac output?
2. Does therapeutic management based on esophageal Doppler ultrasound-based cardiac output monitoring during surgery lead to improved patient outcomes (fewer complications and shorter hospital stay), compared to:

    a. Pulmonary artery catheter-based measurement of cardiac output via thermodilution?

    b. Catheter-based measurement of central venous pressure?

    c. Conventional clinical assessment (physical examination, fluid input and output measurements)?
3. Does therapeutic management based on esophageal Doppler ultrasound-based cardiac output monitoring during hospitalization lead to improved patient outcomes (fewer complications and shorter hospital stay), compared to:

    a. Pulmonary artery catheter-based measurement of cardiac output via thermodilution?

    b. Catheter-based measurement of central venous pressure?

    c. Conventional clinical assessment (physical examination, fluid input and output measurements)?

4. What complications, harms, and adverse events associated with esophageal Doppler ultrasound-based monitoring have been reported?

The esophageal Doppler ultrasound devices evaluated in this report include the following: CardioQ, HemoSonic 100, and TECO. Earlier models of these devices (some with different names than the current models) were also included. These devices differ from transesophageal echocardiography (TEE) with Doppler in that they require less training to operate and are less expensive.(1) TEE systems are beyond the scope of this report.

As shown in the Key Questions, this report focuses on patient-oriented outcomes. Validation studies that compared the agreement between Doppler ultrasound cardiac output measurements and measurements obtained with comparable technologies (e.g., thermodilution) were beyond the scope of this report. Such comparisons are generally performed within the same patients, whereas any comparison of clinical outcomes requires that the compared technologies must be used to direct fluid replacement in different patients.

BACKGROUND

In this section, we provide background information on cardiac output monitoring and esophageal Doppler ultrasound. The purpose of this section is to provide context for the research syntheses presented later in this report. The information presented in this section may be based upon opinion, and we have not critically assessed its accuracy. This section is therefore not, in the strictest sense of the term, evidence-based. Consequently, no statement in this *Background* section should be interpreted as an endorsement or a criticism by ECRI.

Intravenous Fluid Management

For patients in surgery or intensive care units, optimization of intravenous fluid replacement (colloid or crystalloid solutions) is essential to achieve maintenance of adequate organ perfusion. Ideally, this requires measurement of blood pressure and flow. Blood pressure must be sufficient to maintain a patent (open) vessel lumen, and blood flow must be sufficient to deliver adequate oxygen and metabolites to every cell (as well as remove metabolic byproducts such as CO_2 and lactate).(1,2) If patients do not receive enough additional fluids, this can lead to hypovolemia (abnormally low levels of blood plasma) followed by hypotension and renal failure.(3) Conversely, addition of too much fluid may precipitate heart failure.

Methods of Intravenous Fluid Management

Conventional Clinical Assessment

Conventional clinical assessment usually refers to non-invasive assessment of various clinical markers. In some institutions, fluid management may be based only on assessment of hemodynamic variables such as heart rate, systolic blood pressure, and urinary output, with no measure of blood flow or central venous pressure (CVP).

Central Venous Pressure (CVP) Monitoring

In addition to the conventional clinical assessment described above, some institutions will monitor CVP via a central venous catheter. CVP is a measure of the pressure in the right atrium.(4) Loss of fluid leads to a drop in CVP, while addition of fluid tends to increase CVP. This provides another measure to aid the physician in deciding how much additional fluid is required for individual patients in surgery or intensive care. Although CVP is usually measured with a catheter, some institutions may monitor CVP using a non-invasive method,(5) which is less accurate and is often incorporated in conventional clinical assessment.

Cardiac Output Monitoring

Cardiac output refers to the amount of blood pumped by the heart per unit time, measured in liters per minute. It can be calculated by multiplying the stroke volume (the amount of blood pumped by the left ventricle in one contraction) and the heart rate.(6)

In theory, calculation of cardiac output may enable clinicians to more accurately titrate the level of additional fluids (colloid or crystalloid intravenous solutions) and vasoactive therapies to achieve adequate tissue perfusion. If the cardiac output does not increase after a fluid addition, this may indicate that the upper limit of fluid replacement has been achieved, and further fluid addition could lead to venous congestion and postoperative pulmonary edema.

Several methods are available for monitoring of cardiac output; the method generally used as a “reference” standard for other methods is thermodilution via a pulmonary artery catheter (see Key Question 1 for a detailed description of this and other methods for cardiac output monitoring). The use of pulmonary artery catheters carries a risk of serious complications, which has led some investigators to prefer less invasive methods of cardiac output measurement.

Esophageal Doppler Monitoring

Esophageal Doppler monitoring is a relatively non-invasive technique used to measure cardiac output. A small probe is inserted into the esophagus of mechanically-ventilated patients, usually during anesthesia. The probe is introduced orally and advanced gently until its tip is located approximately at the mid-thoracic level, and then rotated so that it faces the descending aorta. The tip of the probe contains a Doppler transducer which transmits an ultrasound beam (4 MHz continuous-wave or 5 MHz pulsed-wave). The change in frequency of this beam as it reflects off a moving object allows measurement of blood flow velocity in the descending aorta. This measurement, when combined with an estimate of the cross-sectional area of the aorta, allows calculation of hemodynamic variables including stroke volume and cardiac output.(2,7) Potential limitations of esophageal Doppler monitoring include operator dependency, occasional difficulties in probe placement, difficulty interpreting the signal during periods of arrhythmia,(8) and the lack of central venous access (which can be obtained when using a pulmonary artery catheter to measure cardiac output). However, some practitioners do not consider lack of central venous access to be a limitation, as central venous catheters carry a risk of infection and other complications.(9)

Currently, the two most widely-used esophageal Doppler monitors are the CardioQ (Deltex Medical, Chichester, UK) and the HemoSonic 100 (Arrow International, Reading, PA). Each uses a different method for determination of stroke volume. The CardioQ uses a disposable 6 mm probe to measure blood flow in the descending thoracic aorta. A proprietary nomogram (factoring in patient age, weight, and height) is used to estimate the cross-sectional area in the descending aorta. The HemoSonic 100 uses a 7 mm non-disposable probe that requires a disposable sheath for each use. The probe contains two transducers, one measuring aortic blood flow and the other (known as an M-mode echo transducer) measuring the cross-sectional area of the descending aorta. The M-mode (or motion mode) transducer is further used to confirm probe placement by providing visualization of the walls of the descending aorta. M-mode technology is also used in transesophageal echocardiography (TEE). For this reason, the HemoSonic monitor is sometimes referred to as an echo-ED monitor.(1,10)

Clinical Practice Guidelines

Our searches identified no clinical practice guidelines specifically focusing on the use of esophageal Doppler monitoring systems for optimization of fluid replacement in surgical or intensive care patients.

Target Population

Despite the lack of clinical practice guidelines, inclusion/exclusion criteria and patient characteristics described in studies of cardiac output monitoring suggest that the authors in these studies reserved guided fluid replacement during surgery or intensive care for relatively higher-risk patients. They tend to be older patients, including some with co-morbid conditions, who require major surgical procedures (such as bowel resection, hip fracture repair, and cardiac surgery) with a significant anticipated blood loss.(8) One study's inclusion criteria specified patients who were undergoing procedures where the anticipated blood loss was >500 ml.(11) Although most cardiac output monitoring studies do not specify an expected level of blood loss, they focus on procedures that are associated with high levels of blood loss necessitating fluid replacement. Cardiac output-guided fluid replacement is not generally considered for low-risk patients having ambulatory surgery.(8)

Previous Systematic Reviews

The Cochrane Collaboration has published a systematic review titled "Perioperative fluid volume optimization following proximal femoral fracture".(3) This review evaluated randomized controlled trials (RCTs) that compared different fluid optimization interventions, including esophageal Doppler monitoring. Two trials met the inclusion criteria; both compared esophageal Doppler monitoring to "usual care", and one of the trials also compared central venous pressure (CVP) monitoring to "usual care".
The authors focused on patient-oriented outcomes such as mortality, complications, length of hospital stay, and independence in activities of daily living. They concluded that "invasive methods of fluid optimization (they consider esophageal Doppler an

invasive method) during surgery may shorten hospital stay, but their effects on other important, patient-centred, longer-term outcomes are uncertain. Adverse effects on fatality cannot be excluded." They further concluded that "more research is needed".

Two additional systematic reviews have evaluated the agreement of esophageal Doppler monitoring and thermodilution via pulmonary artery catheter for measurement of cardiac output. One of these reviews was performed by investigators at the University of Manchester and University College London in the UK,(10) while the other was conducted by investigators at the University of Calgary (Alberta, Canada).(12) These reviews selected studies that compared agreement between measurements of the two techniques within the same patients. They did not evaluate the impact of these techniques on clinical outcomes (which would have required parallel control groups that each received a different monitoring method). The UK authors concluded that "the esophageal Doppler monitor has high validity (no bias and high clinical agreement with pulmonary artery thermodilution) for monitoring changes in cardiac output".
The Canadian authors concluded that esophageal Doppler "is a practical, reliable, and valid device for measuring cardiac output in perioperative and critically ill patients. Further studies with larger numbers of patients are needed to determine if the limited precision observed is inherent to the technique, the diagnoses of patients studied, or the small sample sizes."

Ongoing Trials

In a summary of preliminary results for the year 2005, Deltex Medical announced that the first multicenter RCT investigating the impact of CardioQ on death rates following emergency hip fracture repair would be initiated in 2006. The trial will involve more than 15 hospitals across France and is expected to be completed in three years.
Dr. Bernard Cholley is the lead investigator. The number of patients to be enrolled and the control intervention were not mentioned.(13)

Regulatory Issues

Manufacturers and U.S. Food and Drug Administration (FDA) Status

The CardioQ cardiac output and fluid status monitoring system is manufactured by Deltex Medical Ltd. (Chichester, West Sussex, UK). The earliest model of this system (originally known as EDM) received FDA approval for marketing under the 510(k) process in November, 1995.(14) The later model (renamed CardioQ) received FDA approval for marketing under the 510(k) process in August, 2003.(15)

The Hemosonic 100 cardiac output monitor is manufactured by Arrow International (Reading, PA). This device received FDA approval for marketing under the 510(k) process in February, 1998 (originally approved as the Somatec, Inc. DYNEMO 3000).(16)

The TECO cardiac output monitoring system was manufactured by Medicina Ltd. (Oak House, Cookham, Berkshire, UK). This device was never approved for marketing in the U.S., but has been marketed in the U.K., Ireland, India, and China. In February 2006, Deltex Medical announced that it had purchased the TECO monitor business from Medicina Ltd.(17) Deltex Medical has no plans to market the TECO system, so it is no longer commercially available.

Training and Credentialing

Our searches identified no formal guidelines for training and credentialing of esophageal Doppler system operators. However, a study by Lefrant et al. found that adequate training was achieved after the operator had used an esophageal Doppler device in 12 patients. Successful training was determined by the operator's ability to get "a loud and clear Doppler signal with a well-defined sharp waveform". Also, the correlation coefficient (r) for paired cardiac output measurements at different times in the same patients increased from 0.53 during the training period to 0.89 during the post-training period (r = 1 would be a perfect correlation).(18)

Current CMS Policy Regarding Cardiac Output Monitoring with Doppler Ultrasound

Current CMS policy appears in the *NCD for Ultrasound Diagnostic Procedures* (220.5).(19) The benefit category is listed as Diagnostic Tests. The indications and limitations of coverage section divides these procedures into two categories. The NCD states that “Medicare coverage is extended to the procedures listed in Category I” provided that the techniques are “medically appropriate and the general indications specified in these categories are met. Techniques in Category II are considered experimental and should not be covered at this time.” Monitoring of cardiac output (Doppler) is currently listed in Category II (clinical reliability and efficacy not proven) and is thus ineligible for coverage at present.

Third Party Payer Coverage

Our searches of ten company Web sites found three payers with coverage policies on topics related to esophageal Doppler monitoring, while seven payers had no coverage policy related to this technology. This is a representative but not comprehensive sample of coverage policies. Details about the payers with relevant coverage policies appear in Table 1; it is not clear that these policies include esophageal Doppler devices such as CardioQ. Payers that did not have relevant coverage policies include Blue Cross/ Blue Shield of Alabama, Blue Cross/ Blue Shield of Minnesota, Blue Cross/ Blue Shield of Tennessee, Blue Cross/ Blue Shield of Wisconsin, Cigna, Health Partners, and Humana. This does not necessarily mean that these companies do not provide reimbursement for esophageal Doppler monitoring; the procedure may be covered as part of a composite payment (DRG) for surgical or hospital services.

Table 1. Third Party Payer Coverage

Third Party Payer	Coverage Policy
Aetna(20) http://www.aetna.com/cpb/data/CPBA0008.html	The coverage policy is titled "Color-Flow Doppler Echocardiography in Adults". It includes transesophageal Doppler echocardiography; however, it is not clear that the policy includes esophageal Doppler devices such as CardioQ. Aetna considers color-flow Doppler in adults medically necessary for the following indications: • Evaluation of septal defects • Evaluation of the severity of valve stenosis or regurgitation • Evaluation of site of left-to-right or right-to-left shunts • Assessment of diseases of the aorta • Evaluation of prosthetic valves Aetna considers color-flow Doppler in adults experimental and investigational for all other indications.
Blue Cross/Blue Shield of Massachusetts(21) http://www.bluecrossma.com/common/en_US/medical_policies/108%20Echocardiography%20prn.pdf	The coverage policy is titled "Echocardiography (including transesophageal echo, stress echo, bubble echo, color Doppler echo, fetal cardiac echo)". It is not clear that esophageal Doppler devices are covered by this policy. Cardiac echocardiography in adults and children is covered for any indications, except when used as a screening test in the absence of signs or symptoms of a disease or condition.
Blue Cross/Blue Shield of North Carolina(22) http://www.bcbsnc.com/services/medical-policy/pdf/intraoperative_transesophageal_echocardiography.pdf	The coverage policy is titled "IntraoperativeTransesophageal Echocardiography". It is not clear that this policy covers use of esophageal Doppler systems. The policy provides a long list of surgical procedures for which intraoperative transesophageal echocardiography is covered, and a list of procedures for which use of this technology is not covered. The policy does not address non-operative usage of this technology.

METHODS

Key Questions Addressed

We address the following Key Questions in this report:

1. What types of devices/techniques are currently used to assess cardiac output?
2. Does therapeutic management based on esophageal Doppler ultrasound cardiac output monitoring during *surgery* lead to improved patient outcomes (fewer complications and shorter hospital stay), compared to:

    a. Pulmonary artery catheter-based measurement of cardiac output via thermodilution?

    b. Catheter-based measurement of central venous pressure?

    c. Conventional clinical assessment (physical examination, fluid input and output measurements)?
3. Does therapeutic management based on esophageal Doppler ultrasound cardiac output monitoring during *hospitalization* lead to improved patient outcomes (fewer complications and shorter hospital stay), compared to:

    d. Pulmonary artery catheter-based measurement of cardiac output via thermodilution?

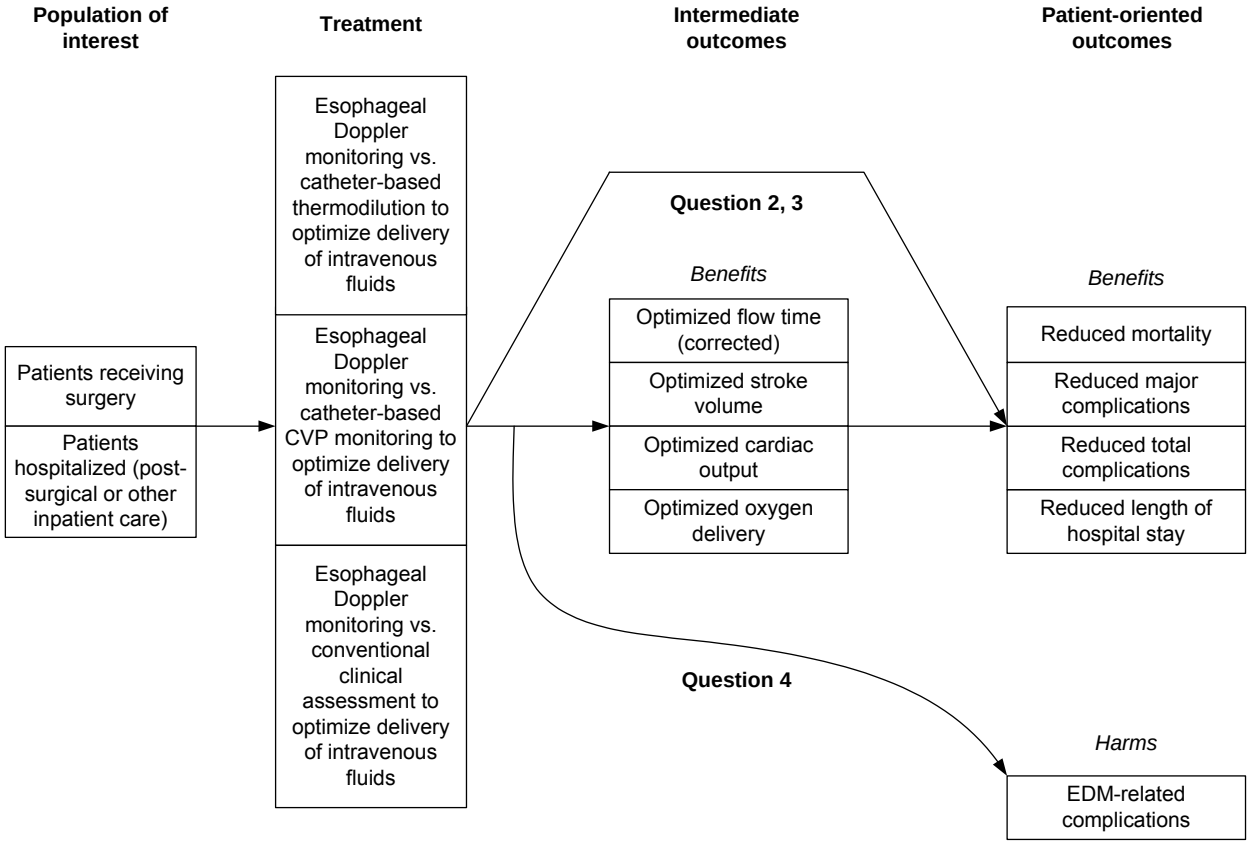
    e. Catheter-based measurement of central venous pressure?

    f. Conventional clinical assessment (physical examination, fluid input and output measurements)?
4. What complications, harms, and adverse events associated with esophageal Doppler ultrasound monitoring have been reported?

In assessing safety, we consider all reported complications that may be related to use of esophageal Doppler ultrasound devices.

Figure 1 illustrates the relationship between esophageal Doppler monitoring, the Key Questions, and the outcomes of interest. Because Key Question 1 is not evidence-based (it merely asks what are the alternative technologies used to measure cardiac output), it is not included in Figure 1. This report evaluates only patient-oriented outcomes, including total complications, major complications (generally defined as life-threatening or requiring intensive or high dependency care), mortality, and length of hospital stay. Intermediate outcomes such as increased cardiac output cannot be perceived by the patient and are beyond the scope of this report. They are presented in Figure 1 only to show their place within the analytic framework.

Figure 1. Analytic Framework



Literature Searches

Details of our literature searches, which included searches of 17 electronic databases, hand searches of the bibliographies of all retrieved articles, and searches of the gray literature, are presented in Appendix A.

Study Inclusion/Exclusion Criteria

General Inclusion/Exclusion Criteria

We used the following general criteria to determine which studies would be included in our analysis for Key Questions 1 through 4:

1. Studies must have been published in English. We recognize the possibility that requiring studies to be published in English could lead to bias, but we believe it is sufficiently unlikely that we cannot justify the additional time and expense for translation.(23,24)
2. Studies must have addressed one of the Key Questions.
3. Studies must have been published as full journal articles (no meeting abstracts). Meeting abstracts generally have insufficient description of methods to allow assessment of quality, and the reported results often contain discrepancies with results presented in later peer-reviewed publication of the same study.
4. If the same study is reported in multiple publications, only the most recent publication will be included. This serves to avoid duplication of data.
5. For controlled studies, 10 or more patients per treatment group must have been enrolled. This increases the likelihood that the studies contain a representative sampling of the patient population.

Question-Specific Inclusion/Exclusion Criteria

The following inclusion/exclusion criterion was specific to Key Question 1:

- Clinical guidelines, review articles, and FDA approvals will be used to identify other methods of cardiac output monitoring

The following inclusion/exclusion criteria were specific to Key Questions 2 and 3:

- Studies must include parallel control groups – controlled studies are required in situations where influences other than the technology of interest may be responsible for treatment outcomes. Comparison of esophageal ultrasound-based monitoring to a standard-of-care-monitored control group (ideally catheter-based measurements of cardiac output or CVP, but also including other conventional clinical assessments) is needed to sort out the influence of the monitoring technology from other potential influences. Trials that compare two monitoring technologies in the same patients cannot be used to determine which technology leads to better clinical outcomes. Therefore, only trials with head-to-head comparisons of esophageal ultrasound and standard-of-care monitoring in separate patients will be examined.

- Within a given trial, patients in both groups must have received comparable surgery (for Key Question 2) or must have had comparable diagnoses (for Key Question 3)

- Studies cannot perform a mixed analysis of surgical and non-surgical patients (such a study would not answer Key Question 2 or 3).

The following inclusion/exclusion criterion was specific to Key Question 4:

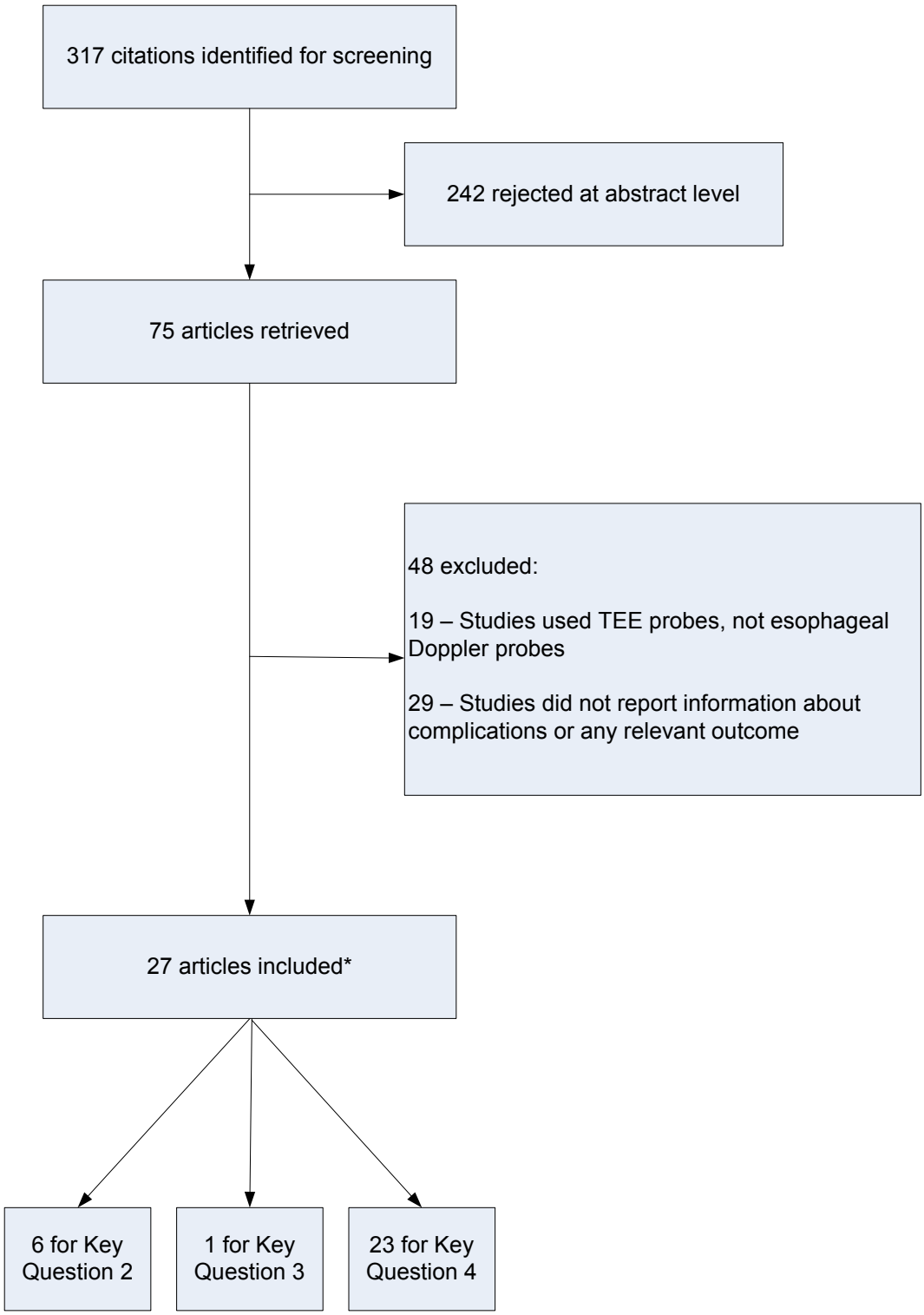
- Studies of any design (controlled trials, case series, case reports), ECRI's Health Device Alerts database and other adverse event databases will be examined for reports of complications, harms and adverse events. These sources cannot be used to determine causality or to estimate frequency of adverse events, but can be used to generate a list of adverse events possibly attributable to the technology.

Identification of Evidence Bases

The selection process used to identify the articles that comprise the evidence base for the key questions addressed in this report is presented in Figure 2. One relevant study published after the original search cutoff date (June 2006) was brought to our attention

by external reviewers. Accordingly, we performed an updated search (September 2006) to identify any additional relevant studies that may have been published since the initial search. Together, our searches identified 317 articles that potentially addressed Key Questions 1 through 4. Of these 317 articles, we retrieved 75. Key Question 1 did not involve an evaluation of evidence, but instead was a survey of current techniques used for cardiac output monitoring. As such, we do not include it as part of the selection process in Figure 2. Seven included articles addressed Key Question 2, one included study addressed Key Question 3, and 23 included articles addressed Key Question 4 (Four of these 23 also also addressed Key Question 2). The included studies are listed in the Evidence Synthesis section under each Key Question that they address.

Figure 2. Summary of Study Selection Process



* 3 articles addressed Key Questions 2 and 4

Data Extraction

Information extracted from the included studies is presented in Evidence Tables in Appendices C-F. These tables describe patient inclusion/exclusion criteria, design details (prospective, blinding status, etc.), information on enrolled patients (demographics, underlying risk, etc.), and study results. When study authors did not report dichotomous data as percentages, we computed percentages. We have only extracted outcome data relevant to the Key Questions in this report.

Evaluation of the Quality of the Evidence Base

ECRI's algorithm, which is presented in Appendix B, provides systematic, reproducible, transparent, and *a priori* decision rules for rating the strength of a body of evidence. In applying the algorithm, we draw a distinction between a qualitative conclusion (one which answers the question "Does it work?") and a quantitative conclusion (one which answers the question "How well does it work?"). Second, we utilize an algorithm that we developed to assign a strength rating to the evidence that supports our qualitative conclusions and a rating that defines how stable we believe any estimate of treatment effect to be.

Table 2 presents definitions of the strength of evidence and stability ratings that may be obtained using the algorithm. These definitions, which are similar to those proposed by the Grading of Recommendations Assessment, Development and Evaluation (GRADE) working group,(25) are intuitive. Qualitative conclusions that are supported by strong evidence are less likely to be overturned by the publication of new data than are conclusions supported by weak evidence. Likewise, quantitative estimates of treatment effect that are backed up by stable data (data with relatively narrow confidence intervals) are less likely to change significantly when new data are published than are estimates of treatment effect drawn from a less stable data set. For more information on the criteria used to rate studies, see the Quality of Included Studies section under each Key Question in the Evidence Synthesis section of the report.

Table 2. Interpretation of Strength of Evidence and Stability Ratings

Strength of Evidence Interpretation	
Qualitative Conclusion (Does it work?)	
Strong evidence	Evidence supporting the qualitative conclusion is convincing. It is highly unlikely that new evidence will lead to a change in this conclusion.
Moderate evidence	Evidence supporting the qualitative conclusion is somewhat convincing. There is a small chance that new evidence will overturn or strengthen our conclusion. ECRI recommends regular monitoring of the relevant literature at this time.
Weak evidence	Although some evidence exists to support the qualitative conclusion, this evidence is tentative and perishable. There is a reasonable chance that new evidence will overturn or strengthen our conclusions. ECRI recommends frequent monitoring of the relevant literature at this time.
Inconclusive	Although some evidence exists, this evidence is not of sufficient strength to warrant drawing an evidence-based conclusion from it. ECRI recommends frequent monitoring of the relevant literature at this time.
Quantitative Conclusion (How well does it work?)	
High stability	The estimate of treatment effect included in the conclusion is stable. It is highly unlikely that the magnitude of this estimate will change substantially as a result of the publication of new evidence.
Moderate stability	The estimate of treatment effect included in the conclusion is somewhat stable. There is a small chance that the magnitude of this estimate will change substantially as a result of the publication of new evidence. ECRI recommends regular monitoring of the relevant literature at this time.
Low stability	The estimate of treatment effect included in the conclusion is likely to be unstable. There is a reasonable chance that the magnitude of this estimate will change substantially as a result of the publication of new evidence. ECRI recommends frequent monitoring of the relevant literature at this time.
Unstable	Estimates of the treatment effect are too unstable to allow a quantitative conclusion to be drawn at this time. ECRI recommends frequent monitoring of the relevant literature.

We apply each kind of rating to the *body* of evidence that addresses each outcome, not to individual studies. We also rate on an outcome-by-outcome basis. Four primary factors determine our ratings for both strength and stability; the quality, quantity, robustness, and consistency of the evidence. Under certain circumstances, the size of the treatment's effect, and whether mega-trials (trials with ≥ 1000 patients) are available also influence our ratings of the evidence underlying qualitative conclusions.

We estimated the generalizability of each study to the U.S. Medicare population using study enrollment criteria and the reported characteristics of the patients who were actually enrolled in the study.

Statistical Methods

We calculated individual study effect sizes from dichotomous data using the log odds ratio (summary log odds ratios were converted to odds ratios in the text and conclusion statements). If there were no events in one of the study groups, the Peto log odds and odds ratios were used, as this method is appropriate for rare events in studies with no substantial imbalance in the number of patients in each comparison group. An alternative method used to calculate effect size (Cohen's h, the arcsin transform of the difference between proportions) from dichotomous data was described by Snedecor and Cochran.(26) Effect sizes for continuous data (e.g., length of hospital stay) were calculated in the original metric (the weighted mean difference in days).

In some instances, certain studies presented data for a continuous outcome (length of hospital stay) in a form that did not allow calculation of a precise effect size. Whereas accurate calculation of an effect requires means and standard deviations (SDs) or 95% confidence intervals, some studies report length of stay as medians and ranges. Because in some studies the mean length of hospital stay may be markedly skewed by an outlier, the median becomes a superior measure of the average patient experience. For this reason, we used imputation methods to estimate effect sizes from the median, range, and sample size when possible as described by Pudar Hozo et al.(27) If studies reported medians and interquartile ranges (IQRs), we assumed the distance between the median and the 25^{th} percentile was 0.67 SDs. If medians were not available, effect sizes were calculated from means and SDs.

Whenever relevant data from three or more studies were available and could be combined (and assuming that the studies used similar enough clinical methods that combining was considered appropriate), we summarized the results using meta-analysis. Meta-analysis allows one to pool data from different studies to obtain an average estimate of the treatment effect. It also provides a means for formally identifying and exploring important differences among the results of different studies (consistency). For a complete description of when studies can be combined in a meta-analysis, see Appendix B under Strength of Evidence Algorithm.

In brief, we first tested the available data to determine whether the results of the studies included in the meta-analysis differed from one another by more than that expected by chance (heterogeneity testing) using the I^2 statistic ($I^2 \geq 50\%$ indicates notable unexplained inconsistency).(28) If study results did not differ in this manner (i.e., the data were consistent), we next pooled the study results in a fixed-effects model to obtain a summary estimate.(29) Random effects meta-analysis was performed to enable a qualitative conclusion if $I^2 \geq 50\%$ or if fewer than 80% of studies reported the outcome of interest or had calculable effect sizes.

Having obtained a summary estimate of the results, we then tested the robustness of our findings using sensitivity analyses as recommended by Olkin.(30) This involved the removal of each individual study separately to determine whether any one study had a substantial influence on the meta-analytic findings. We also performed the systematic addition of each study (cumulative meta-analysis) to determine the study's effect on the summary result. Studies were added in order, beginning with the highest-weighted study and ending with the lowest-weighted study (we also added them in reverse order, from lowest-weighted to highest-weighted). These sensitivity analyses were used for testing both quantitative and qualitative robustness. As a further test of qualitative robustness, we re-calculated summary effects in a different metric (Cohen's h in place of the log odds ratio, Hedges' g in place of the weighted mean difference) to see if this overturned the qualitative conclusions. Because of the assumptions used in meta-analysis of length of stay, additional sensitivity analyses were undertaken on this outcome (described under findings for Key Question 2).

In instances where the evidence base consisted of two studies and the median quality of the studies was high, we combined the studies in a meta-analysis in an attempt to reach a qualitative (but not quantitative) conclusion.

EVIDENCE SYNTHESIS

Key Question 1: What Types of Devices/Techniques Are Currently Used to Assess Cardiac Output?

This question requires a summary of the technologies currently used to measure cardiac output. The information described in this section is derived primarily from review articles written by experts in the field, and is not truly evidence-based.

Summary of Technologies Used to Measure Cardiac Output

A variety of invasive and non-invasive methods are currently used to measure cardiac output. The ideal technique would be easy to apply, operator independent, without morbidity, accurate, reproducible, continuous use, and cost-effective. None of the available techniques meets all of these criteria.(31) Each method and its advantages and limitations are described below.

Thermodilution (via a Pulmonary Artery Catheter)

Thermodilution refers to the measurement of blood flow based on induction of a known change in the intravascular heat content of flowing blood at one point of the circulation, and detection of the resultant change in temperature at a point downstream.(32) Measurement requires insertion of a pulmonary artery catheter (PAC) through the right atrium of the heart with the tip placed in the pulmonary artery. For many years, the standard thermodilution technique involved introduction of a fluid bolus (colder than the patient's blood) through the proximal port of the PAC into the right atrium. The injected fluid mixes with blood passing through the tricuspid valve into the right ventricle. As the cooler blood passes the catheter tip located in the pulmonary artery, a thermistor within the catheter senses the temperature change. A computer attached to the catheter calculates a curve for change in temperature over time and converts it into a measurement of cardiac output.(33) This method allows intermittent but not continuous cardiac output measurement.(34)

A more recent alternative to bolus thermodilution is continuous thermodilution, so named because it allows continuous measurement of cardiac output. This technique uses a modified PAC containing a thermal filament (maintained in the right ventricle) that continuously transfers heat directly into the blood. As in bolus thermodilution, the temperature change is detected downstream by the thermistor located in the catheter tip, and cardiac output is calculated by a computer. However, in this instance the computer continuously displays cardiac output readings that are updated every 30 seconds to provide an average flow over the previous three to five minutes.(34)

Although often used as the “reference” method when evaluating other methods of cardiac output measurement, thermodilution is not a true “gold standard” technique. Under ideal circumstances, the fluid bolus method has a 10% error rate related to instrument inaccuracies, operator error, and temperature transduction.(33) The relative accuracy of the continuous method is uncertain. Furthermore, the invasiveness of these methods carries a risk of serious complications, which has led some investigators to prefer less invasive methods of cardiac output measurement.

Dye Dilution

An indicator dilution method similar to bolus thermodilution is injection of a known concentration of a colored dye (such as indocyanine green) into the pulmonary artery. The concentration of dye after equilibration is measured at a downstream site (usually the femoral or radial artery). The moment-to-moment dye concentration measured spectrophotometrically by a densitometer produces a curve. This allows calculation of cardiac output, which equals 60x the dose of injected dye divided by the area under the curve (average dye concentration x time).(35) Dye dilution appears to be less widely used than thermodilution or lithium dilution (described below).

Lithium Dilution

Lithium dilution is used to measure cardiac output using a principle similar to bolus thermodilution and dye dilution. Unlike thermodilution or dye dilution, Lithium chloride (LiCl) can be injected into a central or peripheral vein, which may decrease the risk of serious complications. An arterial catheter with an attached lithium sensor records the arterial lithium concentration time curve; this data are transmitted to a hemodynamic

monitor that calculates cardiac output based on the following formula: cardiac output = LiCl dose x 60area x (1-haematocrit) $lmin^{-1}$. LiCl is administered in a bolus of 0.15 to 0.3 mmol (for an average adult) and has no known pharmacological side effects at this dose range. Nevertheless, it is recommended that not more than 10 to 20 boluses should be administered to a single patient.(7) Like bolus thermodilution, this method allows only intermittent, rather than continuous, measurement of cardiac output. However, the commercially available lithium dilution cardiac output (LiDCO^{TM}) system (LiDCO Ltd., London, UK) can be used as an indicator method in conjunction with another method that does allow continuous cardiac output measurement (see below under Pulse Contour Cardiac Output).(36)

Pulse Contour Cardiac Output

Analysis of the arterial pulse pressure waveform, measured by an arterial catheter, can be used to calculate cardiac output. The pulse pressure waveform results from the interaction between stroke volume and the mechanical characteristics of the arterial tree. Pulse contour methods use the pressure waveform to predict stroke volume. An independent technique is required to provide initial calibration of the continuous cardiac output analysis. The two major commercially available devices use different methods for calibration. PiCCO (PULSION Medical Systems AG, Munich, Germany) uses transpulmonary thermodilution measured from a central venous line (PAC not needed) to a central arterial line (femoral or axillary) for calibration purposes, while PulseCO (LiDCO Ltd., London, U.K.) uses lithium dilution for this purpose.(37)

A more recent modification of pulse contour analysis, known as the pressure recording analytical method (PRAM), can derive cardiac output from the pressure waveform without requiring an independent calibration method.(38) However, this method has not yet been widely evaluated.

Methods Using The Fick Principle (Direct Oxygen and Inert Gas Rebreathing)

The Fick principle enables calculation of cardiac output as the ratio between the consumption of any gas diffusing through the lungs (e.g., carbon dioxide) and the difference between arterial and venous blood levels of the gas. Commercially available devices use the Fick principle to measure either oxygen or carbon dioxide (CO_2). Devices that measure oxygen have the drawback of requiring invasive central venous and arterial catheters for mixed arterial and venous blood samples, and they cannot be used in patients ventilated with a fractional inspired oxygen (FiO2) greater than 60%; thus, they often cannot be used in critically ill patients.(37)

In contrast, devices that use the Fick principle to measure CO_2 (inert gas rebreathing) allow non-invasive determination of cardiac output. The NICO monitor (Novametrix Medical Systems, Inc., Wallingford, CT) calculates CO_2 production from minute ventilation and arterial CO_2 is estimated from end-tidal CO_2. Partial rebreathing reduces CO_2 elimination and increases end-tidal CO_2. Combining measurements under normal and rebreathing conditions allows omission of venous CO_2 from the Fick equation, eliminating the need for a central venous catheter.(2,37) This method only measures blood flow that participates in gas exchange. However, it does provide an estimate of the amount of blood that bypasses the lungs (the intrapulmonary shunt); this amount added to the amount involved in gas exchange comprises the total cardiac output.(34) This method requires the patient to be under fully controlled mechanical ventilation, and arterial blood samples are required for shunt estimation.(2) Also, poor hemodynamic instability and increased intrapulmonary shunt may decrease the precision of cardiac output estimation with this method.(37)

Thoracic Electrical Bioimpedance

Thoracic electrical bioimpedance (TEB, also known as impedance plethysmography) uses four electrodes attached to the neck and thorax to provide continuous assessment of cardiac output. A high frequency, low amplitude electric current is passed across the thorax. The electrodes measure changes in electrical impedance (resistance to flow) of

the thoracic cavity as aortic blood flow increases and decreases in response to the beating heart.(39) Changes in impedance correlate with stroke volume and allow calculation of stroke volume (and ultimately, cardiac output). Although this is a completely non-invasive method, the degree of methodological diversity (including different available electrode arrays and equations that produce differing results), difficulties in placing electrodes, questions about accuracy and other methodological issues have limited the diffusion of this technique for cardiac output measurement.(7)

Transesophageal Echocardiography

Transesophageal echocardiography (TEE) for measuring cardiac output has some features in common with esophageal Doppler monitoring. Like the latter technique, TEE uses an esophageal probe that can employ Doppler monitoring to calculate blood flow velocity, and an M-mode transducer to measure the cross-sectional area of the aorta (although the M-mode transducer is used in only one commercially available esophageal Doppler probe). Unlike esophageal Doppler, TEE can employ multiplane imaging to more accurately visualize cardiovascular anatomy. Other differences include the higher cost of TEE and the higher level of training required to use TEE.(7) Another limitation of TEE is a small risk of pharyngeal or esophageal perforation.(40) Such events have not yet been recorded for esophageal Doppler probes, possibly because these latter probes are smaller than TEE probes (see Key Question 4 for more information).

Esophageal Doppler Monitoring

See background section for description of this technology.

Ultrasonic Cardiac Output Monitoring (USCOM)

USCOM is a new system that employs continuous-wave Doppler ultrasound to measure cardiac output.(31) Unlike TEE and esophageal Doppler systems, USCOM does not require insertion of an esophageal probe. Instead, a transducer is placed on the patient's chest and positioned to measure either transpulmonary or transaortic blood flow. Thus, this method is completely non-invasive. The cross-sectional area of the aorta is estimated by the Nidorf equation or measured directly via another imaging

method (e.g., echocardiography).(41) Because this is a relatively new device, it has not been widely evaluated at this time.

Subsection Summary

Several methods are currently used to monitor cardiac output in patients during surgery or intensive care. These methods include thermodilution, dye dilution, lithium dilution, methods using the Fick principle, pulse contour methods, thoracic electrical bioimpedance, transesophageal echocardiography, esophageal Doppler monitoring, and ultrasonic cardiac output monitoring.

Key Question 2: Does Therapeutic Management Based on Esophageal Doppler Ultrasound Cardiac Output Monitoring During Surgery Lead to Improved Patient Outcomes (Fewer Complications and Shorter Hospital Stay), Compared to Catheter-Based Measurement of Cardiac Output (Thermodilution) or Central Venous Pressure, or Conventional Clinical Assessment?

Evidence Base

Our searches found eight randomized controlled trials (RCTs) that potentially addressed this question. On retrieval, one of the eight articles was found not to meet our inclusion criteria for this question. This study addressed Key Question 3 and is evaluated later in this report.

Seven RCTs with a total of 583 patients remained to address Key Question 2. These studies are listed in Table 3. No included studies compared the efficacy of esophageal Doppler monitoring to thermodilution with a pulmonary artery catheter for optimization of intravenous fluid replacement. Five studies with a total of 453 patients compared the efficacy of esophageal Doppler monitoring plus central venous pressure (CVP) monitoring plus conventional clinical assessment to CVP plus conventional clinical assessment to optimize intravenous fluid replacement. These studies were basically asking whether Doppler monitoring is an effective complementary procedure to CVP plus conventional protocol. Two studies with a total of 130 patients compared the efficacy of esophageal Doppler monitoring plus conventional clinical assessment to conventional clinical assessment to optimize intravenous fluid replacement. One of these two studies also compared esophageal Doppler monitoring to CVP as a competing, rather than complementary, procedure. Details of these studies are presented in Tables D-1 to D-3, Appendix D. Although all of the studies except Venn et al. reported placing Doppler probes in the control group patients during surgery, Doppler monitoring was not used in the fluid maintenance protocol in control patients (at a minimum, the anesthesiologist was blinded to the esophageal Doppler readings). Six of

the seven studies used either the CardioQ esophageal Doppler monitoring system or an earlier model of this system. The remaining study (Conway et al.) used the TECO esophageal Doppler monitoring system.

Table 3. Evidence Base for Key Question 2

Randomized Controlled Trials	EDM + CVP + conventional clinical assessment vs. CVP + conventional clinical assessment	Noblett et al. 2006(42) Wakeling et al. 2005(43); Conway et al. 2002[a](44); Gan et al. 2002[a](11); Mythen and Webb 1995(45)
	EDM + conventional clinical assessment vs. CVP + conventional clinical assessment	Venn et al. 2002(46)
	EDM + conventional clinical assessment vs. Conventional clinical assessment	Venn et al. 2002(46); Sinclair et al. 1997(47)

CVP Central Venous Pressure Assessment
EDM Esophageal Doppler Monitoring

[a] The studies by Conway et al. and Gan et al. did not use CVP as part of the algorithm for guiding additional fluid challenges in the EDM group. However, CVP was used as part of routine fluid management in most patients in the EDM and control groups.

The types of surgery performed in these studies appear in Table D-1, Appendix D. Three of the studies comparing esophageal Doppler monitoring plus CVP plus conventional clinical assessment to CVP plus conventional clinical assessment included patients undergoing elective bowel surgery. Of the remaining two studies in this group, one included patients undergoing elective general, urologic, or gynecologic surgery, and the other included patients undergoing elective cardiac surgery. Within each study, the type of surgery performed in the Doppler-monitored and control groups was identical. Although the types of surgery differ across studies, they all have one thing in common: an anticipated major loss of blood or significant fluid shifts requiring fluid replacement. This is the key factor that makes esophageal Doppler monitoring applicable to these surgeries. Doppler monitoring is expected to provide similar benefits to these differing types of surgeries because the need for fluid replacement is similar across these

procedures. Therefore, it is reasonable to combine data from these studies in a meta-analysis.

Two other studies, both including patients who received surgery for hip fracture repair, were analyzed as a separate group because they performed a different clinical comparison (Doppler monitoring plus conventional clinical assessment vs conventional clinical assessment alone).

Quality of Included Studies

As shown in Table 4, five studies that compared esophageal Doppler monitoring with CVP monitoring plus conventional clinical assessment to CVP plus conventional clinical assessment had a median quality score of 8.9 on the ECRI quality scale for controlled trials (a study with no flaws would score 10 on our 0-10 scale). Thus, the quality rating for this evidence base is high (for more details on the quality scale, see Appendix B).

Two studies comparing esophageal Doppler monitoring plus conventional clinical assessment to conventional clinical assessment had a median quality score of 9.0, so the quality rating for this evidence base is high. One of these studies (Venn et al.) also compared esophageal Doppler monitoring to CVP.

Tables D-4 and D-5 in Appendix D show the individual study scores based on the answers to each question in the ECRI quality scale.

Table 4. Quality of Included Studies Addressing Key Question 2

Reference	Year	ECRI Quality Score (Rating)
Trials comparing EDM + CVP + conventional protocol to CVP + conventional protocol		
Noblett et al.(42)	2006	9.7 (High)
Wakeling et al.(43)	2005	9.0 (High)
Conway et al.(44)	2002	8.5 (High)
Gan et al.(11)	2002	8.1 (Moderate)
Mythen and Webb(45)	1995	8.9 (High)
Median quality score		8.9 (High)
Trials comparing EDM + conventional protocol to CVP + conventional protocol		
Venn et al.(46)	2002	9.0 (High)

Reference	Year	ECRI Quality Score (Rating)
Trials comparing EDM + conventional protocol to conventional protocol		
Venn et al.(46)	2002	9.0 (High)
Sinclair et al.(47)	1997	8.9 (High)
Median quality score		9.0 (High)

Details of Study Enrollees and Study Generalizability

Patient enrollment criteria appear in Table D-1, Appendix D, and characteristics of included patients appear in Table D-2, Appendix D. Of the factors determining generalizability to the Medicare population, age is a key factor. For example, in a study where 100% of the patients are age 65 or older, all of these patients belong to some subset of the Medicare population. The same cannot be said of studies which include patients below age 65, regardless of what other clinical characteristics these patients may have. Thus, for each study we rated the age generalizability of enrolled patients based on the percentage of patients in the age range 65 or above.

We note that certain of the exclusion criteria of these studies may be relevant for assessment of generalizability to the Medicare population. Some studies excluded patients with heart failure, renal dysfunction, esophageal disease, or patients requiring emergency surgery. However, many of the criteria listed are criteria of surgical eligibility; these criteria may be applied to anyone otherwise eligible for surgery and are not particularly useful for evaluating generalizability to the Medicare population. For some of these procedures, only a fraction of the Medicare population would be considered surgical candidates or would be likely to undergo the procedure. Determination of generalizability based on surgical eligibility criteria would require knowledge of the percentage and characteristics of Medicare patients who would be candidates and/or likely to undergo a given procedure. Although some exclusion criteria are not related to surgical eligibility and therefore more useful for determining generalizability, it is often difficult to determine which category each individual criterion belongs to. Due to the lack

of such information, we did not attempt to determine an overall generalizability rating based on study exclusion criteria.[1]

The five studies that compared esophageal Doppler monitoring plus central venous pressure (CVP) monitoring plus conventional clinical assessment to CVP plus conventional clinical assessment had some overlap with the Medicare population, but the age ranges were generally large, ranging from young adult to elderly. Therefore, the median age generalizability of these studies to the Medicare population is "Fair."[2] Table 5 provides the estimated percentage of patients (calculated by ECRI) in each study with age ≥65 years.

The two studies that compared esophageal Doppler monitoring plus conventional clinical assessment to conventional clinical assessment alone were more age generalizable to the Medicare population. Venn et al. excluded patients with age less than 65 years,(46) while Sinclair et al. excluded patients with age less than 55 years.(47) The average age of patients in these studies was in the mid-seventies (Sinclair et al.) and the mid-eighties (Venn et al.), with. Therefore, the age generalizability of these studies to the Medicare population is "High."[2] Table 5 provides more specific information for each study. Females represented the majority of patients (about 80%) in the study by Venn et al. (Sinclair et al. did not report information on patient gender). However, the high percentage of women is not unusual given the advanced age of the patients and the increased risk of hip fractures among women.

Further details of the patients enrolled in these studies are presented in Table D-1 and D-2 in Appendix D.

---

[1] A similar problem arises when considering patient characteristics such as gender. For example, one study of older hip fracture patients is comprised of about 80% women (Venn et al. 2002).(46) Although the overall Medicare population is less than 80% women, this does not mean the study is not generalizable; women are more likely to have osteoporosis and in turn more likely to suffer hip fractures. Thus, one would expect a higher percentage of women in studies of hip fracture surgery. The study is generalizable to the subgroup of Medicare patients most likely to require this procedure.

[2] High = Characteristics of ≥75% of enrolled patients typical of Medicare population; Fair = Characteristics of ≥33% to <75% of enrolled patients typical of Medicare population; Poor = Characteristics of <33% of enrolled patients typical of Medicare population (or enrolled patients represent a subgroup of Medicare population).

Table 5. Age Generalizability of Studies to the Medicare Population

Reference	Year	Percentage of patients age 65 and older	Age generalizability rating
Trials comparing EDM + CVP + conventional protocol to CVP + conventional protocol			
Noblett et al.[b](42)	2006	50%	Fair
Wakeling et al.[a](43)	2005	70%	Fair
Conway et al.[b](44)	2002	57%	Fair
Gan et al.[b](11)	2002	27%	Poor
Mythen and Webb[C](45)	1995	44%	Fair
Median age generalizability		49.5%	Fair
Trials comparing EDM + conventional protocol to CVP + conventional protocol			
Venn et al.[d](46)	2002	100%	High
Trials comparing EDM + conventional protocol to conventional protocol			
Venn et al.[d](46)	2002	100%	High
Sinclair et al.[a](47)	1997	87%	High
Median age generalizability		93.5%	High

[a] The study reported only the medians and interquartile ranges (IQRs), so to estimate the percentage, ECRI assumed a normal distribution and that the 25th and 75th percentiles were each 0.675 SDs from the mean.

[b] The study reported only the means and SDs, so ECRI estimated this percentage by assuming a normal distribution of age.

[c] The study reported only the means and ranges, so to estimate the percentage, ECRI assumed a normal distribution and that the minimum and maximum were each 2.5 SD from the mean.

[d] The study only included patients who were 65 years or older.

Findings of Included Studies

Studies Comparing Esophageal Doppler Monitoring Plus CVP Plus Conventional Assessment to CVP Plus Conventional Assessment

Three of the five studies that addressed this question used CVP on a discretionary basis (i.e., some but not all patients received CVP monitoring).(11,44) Since patients that did not receive CVP appeared to be distributed equally in both groups, we considered it acceptable to group these studies with the two studies wherein all patients received CVP monitoring. These same three studies did not employ CVP in the actual algorithm used to guide fluid challenges in the Doppler-monitored group. However, CVP

was used as part of the routine management of patients in both the Doppler-monitored and control groups.

A brief discussion concerning analysis of complications is warranted. Ideally, analysis would be performed not only on total complications but also on individual complications. However, because different studies reported complications in different ways, the evidence did not permit analysis of most individual complications. The one exception was death, which was reported separately in each study. In any event, a comparison of individual minor complications (e.g. nausea) does not capture the experience of patients with multiple complications, some of which are more serious than others.

Three of the five studies reported “major” complications (see definition under Major Complications below) separate from total complications, thus allowing a separate analysis of these more severe complications. The specific types of major complications were relatively similar among the relevant studies (for further details, see Table D-6, Appendix D).

Total complications (including major and less severe complications) were reported by four of the five studies. Some complications were more specific to the type of surgery being performed, and other complications (such as infections and pulmonary complications) were reported across studies (see Table D-6, Appendix D). However, many of these complications may be triggered or exacerbated by hypovolemia and tissue hypoxia.(11) Thus, it is reasonable to assume that Doppler-guided fluid replacement would have a similar impact on total complications regardless of the type of surgery performed in these studies. Therefore, a combined analysis of total complications from studies of different surgical procedures is appropriate.

Mortality

Five studies with a total of 453 patients reported the number of deaths in each treatment group (see Table D-7, Appendix D). None of these studies reported any deaths occurring during surgery. Three studies reported one death each in the control group within 30 days following surgery, and one study reported a death in the control group within 60 days following surgery. The total number of deaths is too low to allow pooling

of data in a meta-analysis, and the effect sizes of the individual studies are non-informative (none of the studies showed a statistically significant between-group difference and the confidence intervals are too large to demonstrate equivalence). Therefore, no conclusion is possible regarding mortality.

Major Complications

Three of five studies reported only major complications (Mythen and Webb) or reported these complications separately from lesser complications (Noblett et al., Conway et al.). Major complications were generally defined as life-threatening or requiring intensive or high dependency care (for further details, see Table D-6, Appendix D). Table D-8 in Appendix D presents the individual study results. All studies showed a statistically significant reduction in major complications in the Doppler-monitored group.

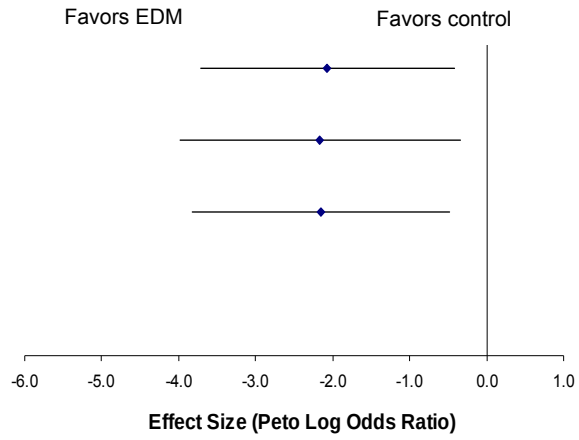
Our test for between-study differences revealed no substantial differences among study results ($I^2 = 0\%$), indicating that the study results could be combined in a meta-analysis. Because each study reported no major complications in the Doppler-monitored groups, we used the Peto method for calculating log odds ratios and odds ratios in our analyses. However, because only three of five available studies presented separate data on major complications, the possibility of selective reporting of positive data remains (although the three studies show highly consistent results). Under such circumstances, the rules of our algorithm prevent a quantitative conclusion.

We then proceeded to pool the studies in a random-effects meta-analysis to arrive at a qualitative conclusion. Pooling these studies resulted in a Peto odds ratio that was statistically significant ($p = 0.00002$) and showed a reduction in major complications associated with Doppler-monitored fluid replacement (see Table 6 below). The next step involved testing with multiple sensitivity analyses to determine the strength of the qualitative conclusion. The most rigorous sensitivity analysis involved inclusion of the two studies that did not separately report major complications, with the assumption that the major complication rates were equal in the Doppler and control groups in these two studies. None of the sensitivity analyses overturned that qualitative findings, indicating that the summary effect was qualitatively robust (Table D-12, Appendix D). Therefore,

the strength of evidence supporting the qualitative conclusion (that Doppler-guided fluid replacement during surgery leads to a clinically significant reduction in major complications compared to the control protocol) is strong. This conclusion applies only to patients undergoing surgical procedures with an expected substantial blood loss or fluid compartment shifts requiring fluid replacement.

Table 6. Meta-analysis – Major Complications (EDM + CVP + Conventional Assessment vs. CVP + Conventional Assessment)

Study	N =	Effect Size	Lower 95% CI	Upper 95% CI	p-value	I^2	Forest Plot
Noblett et al.(42)	103	-2.08	-3.72	-0.44	0.01	NA	
Conway et al.(44)	57	-2.19	-4.01	-0.37	0.02	NA	
Mythen and Webb(45)	60	-2.19	-3.86	-0.51	0.01	NA	
Random-effects summary Peto log odds ratio	220	NC	-3.14	-1.17	0.00002	0%	
Random-effects summary Peto odds ratio	220	NC	0.04	0.31	0.00002	0%	



NA Not Applicable
NC Not Calculated

Total Complications

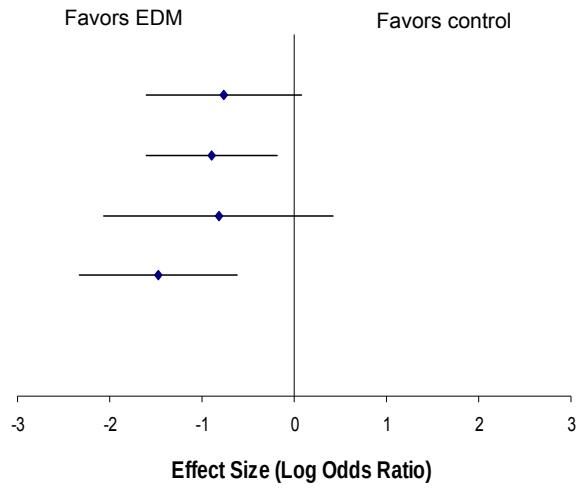
Four of five studies reported total complications (the remaining study by Mythen and Webb reported only major complications and was excluded from this analysis). Table D-9 in Appendix D presents the individual study results. Two of the four studies showed a statistically significant difference indicating fewer total complications in the Doppler-monitored group (the remaining studies also showed fewer complications in the Doppler group, but the difference was not significant). Three of these studies (Noblett et

al., Wakeling et al., Conway et al.) reported the total number of patients with complications in each group, while Gan et al. reported the number of complications rather than number of patients with complications. Our test for between-study differences revealed no substantial differences among study results (I^2 = 0%), indicating that the study results could be combined in a meta-analysis. However, because we combined studies that reported complications somewhat differently (patients vs events), this involves an assumption that the two types of data will produce a similar effect. Although no heterogeneity was detected, we chose not to present a quantitative summary estimate for this outcome.

The studies were then pooled in a random-effects meta-analysis to reach a qualitative conclusion. Pooling these studies showed a statistically significant reduction in total complications associated with Doppler-monitored fluid replacement (see Table 7 below). Our sensitivity analysis for qualitative robustness indicated that the summary effect size is qualitatively robust (i.e., the effect of Doppler monitoring is greater than zero) (Table D-13, Appendix D). Therefore, the strength of evidence supporting the qualitative conclusion (that Doppler-guided fluid replacement during surgery leads to a clinically significant reduction in the total number of complications compared to the control protocol) is strong. This conclusion only applies to patients undergoing surgical procedures with an expected substantial blood loss or fluid compartment shifts requiring fluid replacement.

Table 7. Meta-analysis – Total Complications (EDM + CVP + Conventional Assessment vs. CVP + Conventional Assessment)

Study	N =	Effect Size	Lower 95% CI	Upper 95% CI	p-value	I^2	Forest Plot
Noblett et al.(42)	103	-0.76	-1.60	0.07	0.07	NA	
Wakeling et al.(43)	128	-0.89	-1.60	-0.18	0.01	NA	
Conway et al.(44)	57	-0.82	-2.07	0.43	0.20	NA	
Gan et al.(11)	100	-1.48	-2.33	-0.62	0.0008	NA	
Random-effects log odds ratio	388	NC	-1.43	-0.57	0.000006	0%	



NA Not Applicable

Length of Hospital Stay

All of the included studies reported length of hospital stay. Individual study data appear in Table D-10, Appendix D. The mean or median length of stay varied somewhat among the studies. Gan et al. reported the shortest mean hospital stay (average five to seven days), while Conway et al. reported the longest (12 to 18 days). This could have been due to differences in age and/or surgical procedures in different trials (e.g., the patients in Gan's trial had the lowest average age). Four of the five studies found a statistically significant reduction in length of stay (based on either medians or means) associated with Doppler-monitored fluid replacement (Table D-10, Appendix D). Because a precise effect size could not be calculated from three studies (data were not reported as means with standard deviations or 95% confidence intervals), we could not combine the results of these studies to obtain a quantitative summary estimate of the length of stay (we required that at least 80% of studies in the evidence base for a given outcome must provide data that allow calculation of a precise effect size). Therefore, no quantitative conclusion is possible for this outcome.

However, even if all of the studies had reported means and SDs, there is a legitimate concern that means may not be the best measure of central tendency (and the average patient experience) in these studies. In some studies the mean may be considerably skewed by outliers (as appears to be the case in the Conway study, discussed in more detail below). Therefore, the best measure of central tendency for length of stay appears to be the median. Although calculation of an effect size from medians and ranges (or interquartile ranges) requires certain assumptions, we decided to base our primary analysis of this outcome on medians and ranges when available (means and SDs were used otherwise).

We were able to impute an estimated effect size for studies that reported medians and ranges using the methods described by Hozo et al.(27) This allowed us to combine the results in a random effects meta-analysis to reach a qualitative conclusion. Because these imputation methods are conservative, the individual study effect sizes based on this method were not statistically significant in four out of five studies (in contrast to the p-values reported in the original studies, which were mostly based on non-parametric tests of the medians). However, the summary effect size was statistically significant and qualitatively robust as determined by multiple sensitivity analyses; the 95% CI of the summary effect never overlapped zero (Table D-14, Appendix D). The summary effect size also was qualitatively robust with respect to the line of clinical significance. In the primary meta-analysis, the 95% CI overlapped with the clinically significant level of one day (-0.57 days). Sensitivity analyses did not alter this finding, indicating uncertainty as to whether the true effect size was ≥ or < one day.

The study by Conway et al. was unique in that it was the only study that did not use a version of the CardioQ system. However, the apparent discrepancy between the results of this study and other studies may have been due predominantly to one outlying patient in the Doppler group. Although the means had not been reported in the original study (which had reported medians and ranges), the lead author sent us the unpublished means and standard deviations. Because one patient in the Doppler group stayed in the hospital 103 days (not because of complications, but because the hospital could not find social/community placement for the patient),(8) the difference between means in the

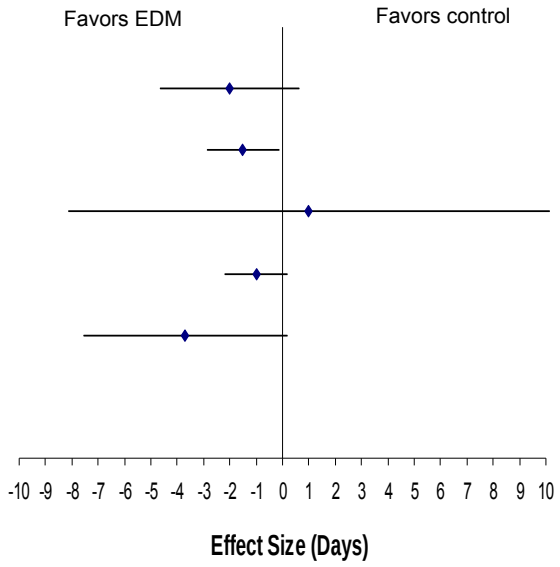
two groups is substantially skewed in favor of the control group (the difference between medians was much smaller, only one day).

To further examine this issue, we performed a sensitivity analysis wherein we used the means from the two studies that reported both means and medians (in other words, means were used whenever possible). This analysis (shown in Table D-14, Appendix D) did not overturn the qualitative findings of the primary meta-analysis. Additional sensitivity analyses using more conservative assumptions for imputing standard deviations from ranges or IQRs also did not overturn the findings (Table D-14).

In summary, the strength of evidence supporting the qualitative conclusion that Doppler-monitored fluid replacement leads to a reduction in hospital stay (clinical significance uncertain) is strong. This conclusion only applies to patients undergoing surgical procedures with an expected substantial blood loss or fluid compartment shifts requiring fluid replacement.

Table 8. Individual Study Effect Sizes – Length of Hospital Stay (EDM + CVP + Conventional Assessment vs. CVP + Conventional Assessment)

Study	N =	Effect Size	Lower 95% CI	Upper 95% CI	p-value	I^2	Forest Plot
Noblett et al.(42)	103	-2.00	-4.61	0.61	0.13	NA	
Wakeling et al.(43)	128	-1.50	-2.86	-0.14	0.03	NA	
Conway et al.(44)	57	1.00	-8.14	10.14	0.83	NA	
Gan et al.(11)	100	-1.00	-2.18	0.18	0.10	NA	
Mythen and Webb(45)	60	-3.70	-7.56	0.16	0.06	NA	
Random effects 95% CI	448	NC	-2.21	-0.57	0.0009	NA	



NA – Not applicable
NC – Not calculated

Studies Comparing Esophageal Doppler Monitoring Plus Conventional Assessment to CVP Plus Conventional Assessment

One study (Venn et al.) compared esophageal Doppler-guided fluid replacement with CVP-guided fluid replacement in patients receiving hip surgery. This study differs from those in the previous section in that none of the patients in the Doppler-monitored group received CVP monitoring. Also, patients in this study received additional fluid challenges in both the Doppler and CVP groups, whereas patients in the control group in the previous section's studies did not receive additional fluid challenges beyond that dictated by conventional protocol. The previous studies were basically asking whether Doppler monitoring is an effective complementary procedure to CVP plus conventional protocol, whereas the study by Venn et al. treats Doppler monitoring as a competing technology to CVP. However, the previous studies do not attempt fluid challenges in the CVP groups, whereas the Venn study presents a protocol for fluid challenges based solely on changes in CVP. An ideal study might be a three-armed study comparing a fluid challenge protocol based on Doppler plus CVP to fluid challenge protocols based on CVP alone and esophageal Doppler alone.

Mortality

Although the study by Venn et al. had half as many deaths in the Doppler group compared to the CVP group, the difference was not statistically significant (3 vs. 6, $p = 0.30$, see Table D-7, Appendix D). Because the difference between the upper and lower 95% confidence intervals around the effect size exceeded 0.8 (using the log odds ratio), this result was not informative. Therefore, no conclusion could be reached for this outcome.

Major Complications

This study did not separate major complications from total complications, so no conclusion was possible for this outcome.

Total Complications

The difference in the rate of total complications (comprised mostly of infections and cardiovascular events) between the Doppler-monitored and CVP-monitored groups was not statistically significant (46.7% vs. 51.6%, p = 0.70, see Table D-9, Appendix D) and not informative for the reason cited above. Thus, no conclusion was possible for this outcome.

Length of Hospital Stay

The mean length of hospital stay did not differ significantly between the Doppler-monitored and CVP-monitored groups (13.5 vs. 13.3 days, p = 0.96, see Table D-10, Appendix D), and the effect size was non-informative. Therefore, no conclusion could be reached for this outcome.

Studies Comparing Esophageal Doppler Monitoring Plus Conventional Assessment to Conventional Assessment

Mortality

Both studies (Venn et al and Sinclair et al) reported total mortality rates (for individual study data, see Table D-7, Appendix D). Six of the eight deaths occurred during the early post-operative period (within 30 days following surgery), while two deaths occurred after this period (within three months following surgery). Neither study showed a statistically significant difference in mortality rates between the two treatment groups. The pooled mortality difference derived from a random-effects meta-analysis was not informative (not statistically significant and having very large confidence intervals), so no conclusion was possible for this outcome.

Major Complications

One study (Sinclair et al) did not report complications, and the other (Venn et al) did not separate major complications from total complications. Therefore, no conclusion was possible for this outcome.

Total Complications

Only one of the studies (Venn et al.) reported total complication rates; data are presented in Table D-9, Appendix D. The between-group difference was statistically significant and favored fewer complications in the Doppler-monitored group (46.7% vs. 79.3%, p = 0.015). However, the difference in the percentage of patients with complications was not quite statistically significant (33.3% vs 55.2%, p = 0.09), although the percentage was lower in the Doppler group. This was a high-quality study, but the magnitude of effect was not large enough to meet our predetermined definition of a large effect (see Appendix B, Decision Point 10), and therefore we do not draw a conclusion about the strength of the evidence based on this single study.

Length of Hospital Stay

Both studies reported length of hospital stay (see Table D-10 in Appendix D for individual study results). Both studies reported a shorter length of stay among patients receiving Doppler-monitored fluid replacement, but the difference was statistically significant only in the study by Sinclair et al. Because the information in this study did not allow calculation of an accurate effect size, we imputed an effect size based on the reported medians and interquartile ranges. This allowed us to pool the results of both studies in a random-effects meta-analysis. The summary effect size was statistically significant (p = 0.008) and favored a shorter hospital stay among patients in the Doppler-monitored group (Table D-11, Appendix D). Because the lower 95% confidence interval of the summary effect was above one day (1.75 days), the difference between groups was clinically significant by our definition. The low number of studies precluded a quantitative conclusion but allowed us to reach a qualitative conclusion. Since the quality of these studies is high, and the magnitude of effect is not large, the strength of evidence supporting the qualitative conclusion that esophageal Doppler monitoring without CVP leads to a clinically significant reduction in hospital stay is weak. This conclusion only applies to patients undergoing surgical procedures with an expected substantial blood loss or significant fluid shifts requiring fluid replacement.

Subsection Summary

After searching the literature, retrieving articles, and applying the inclusion/exclusion criteria, we identified seven studies that addressed Key Question 2. Five compared

esophageal Doppler monitoring plus CVP monitoring plus conventional clinical assessment to CVP plus conventional clinical assessment for optimization of fluid replacement during surgery. The median quality of these five studies was high, and the age-applicability to the Medicare population was fair. The remaining two studies compared the efficacy of esophageal Doppler monitoring plus conventional clinical assessment to conventional clinical assessment to optimize intravenous fluid replacement (one of these studies also compared esophageal Doppler to CVP plus conventional clinical assessment). The median quality of these two studies and their age-applicability to the Medicare population was high.

The addition of esophageal Doppler monitoring for guided fluid replacement to a protocol using CVP and conventional clinical assessment during surgery leads to a clinically significant reduction in the rate of major complications and total complications in surgical patients compared to CVP plus conventional clinical assessment alone. The strength of evidence supporting this conclusion is strong. Because only three of five studies separately reported major complications, and because of differences in the way total complications were reported, no quantitative conclusion is presented for these outcomes.

The addition of esophageal Doppler monitoring to CVP plus conventional assessment also reduces the length of hospital stay for surgical patients (clinical significance uncertain). The strength of evidence supporting this conclusion is strong. The lack of a calculable precise effect size in three studies precluded a quantitative summary estimate of the reduction in length of stay.

Only one study compared esophageal Doppler plus conventional clinical assessment to CVP plus conventional clinical assessment. Because this was one small study with non-informative effect sizes, no conclusions were possible for any of the outcomes of interest.

The addition of esophageal Doppler monitoring for guided fluid replacement to conventional clinical assessment during surgery leads to a clinically significant reduction in the length of hospital stay compared to that associated with conventional clinical

assessment alone. The strength of evidence supporting this conclusion is weak. The low number of studies precluded a quantitative estimate of the reduction in length of hospital stay. Because only a single study reported total complications, no conclusion was possible concerning this outcome.

No conclusion could be reached concerning relative mortality rates for any of the comparisons in Key Question 2.

The conclusions in this subsection only apply to patients undergoing surgical procedures with an expected substantial blood loss or significant fluid compartment redistribution requiring fluid replacement.

Key Question 3: Does Therapeutic Management Based on Esophageal Doppler Ultrasound Cardiac Output Monitoring During Hospitalization Lead to Improved Patient Outcomes (Fewer Complications and Shorter Hospital Stay), Compared to Catheter-Based Measurement or Conventional Clinical Assessment?

Evidence Base

Our searches identified one study (an RCT) that potentially met our *a priori* inclusion criteria and was therefore retrieved. This study of 174 patients was included and appears in Table 9 below. Details of this study are presented in Tables E-1 through E-3, Appendix E.

Table 9. Evidence Base for Key Question 3

Study Design	Treatment Comparison	References
Randomized controlled trials	EDM + CVP + conventional clinical assessment vs. CVP + conventional clinical assessment	McKendry et al. 2004(48)

EDM – Esophageal Doppler Monitoring
CVP – Central Venous Pressure Assessment

Quality of Included Studies

The results of our analysis of the quality of the study by McKendry et al. are summarized in Table 10. We based the quality ratings for this study on the criteria and information presented in Table E-4 of Appendix E.

Table 10. Quality of Included Studies Addressing Key Question 3

Reference	Year	ECRI Quality Score (Rating)
McKendry et al. 2004(48)	2004	8.5 (high)

Details of Study Enrollees and Study Generalizability

Details about the patients enrolled by McKendry et al. are presented in Table E-1 to E-2 of Appendix E. This study allowed inclusion of patients 18 years or older, and the average age of included patients was 66 years. Therefore, some but not all patients were in the age range applicable to the Medicare population, so the age generalizability of this study was considered to be "Fair."[3] The estimated percentage of patients ≥65 years in this study appears in Table 11 below.

Table 11. Age-Generalizability of Study to the Medicare Population

Reference	Year	Percentage of patients age 65 years and older	Age generalizability rating
McKendry et al.[a] 2004(48)	2004	54%	Fair

[a] The study reported only the means and SDs, so ECRI estimated this percentage by assuming a normal distribution of age.

Findings of Included Studies

The only study that met our inclusion criteria (McKendry et al.) compared the efficacy of esophageal Doppler monitoring plus CVP plus conventional clinical assessment to CVP plus conventional clinical assessment for optimization of intravenous fluid replacement in patients admitted to cardiac intensive care following cardiac surgery. The specific individual complications reported are listed in Table E-5, Appendix E.

[3] High = Characteristics of all enrolled patients typical of Medicare population; Fair = Characteristics of some enrolled patients typical of Medicare population; Poor = Characteristics of only a few enrolled patients typical of Medicare population or enrolled patients represent a subgroup of Medicare population.

Mortality

McKendry et al. reported four deaths in the Doppler-monitored group and two deaths in the control group. The causes of death “were not considered directly attributable to early post-operative care”.(48) Although the difference in mortality rates was not statistically significant (p = 0.43), the 95% confidence intervals were so large that the finding could not be considered informative (Table E-6, Appendix E). Thus, no conclusion can be drawn concerning this outcome.

Major Complications

This study did not specifically separate major complications from total complications reported, although the authors stated that there was “a trend towards fewer major postoperative complications and deaths” in the Doppler-monitored group. It is possible that the authors considered all reported complications as major, but some of these complications (such as atrial fibrillation) would not have been considered major as defined by the trials in Key Question 2. Thus, no conclusion is possible for this outcome.

Total Complications

Although this study reported fewer patients with postoperative complications in the Doppler-monitored group (19.1% vs. 30.6% in control group), the difference was not statistically significant (p = 0.08). Furthermore, the 95% confidence intervals around the difference were too large to be considered informative (Table E-6, Appendix E). Therefore, no conclusion can be reached for this outcome.

Length of Hospital Stay

McKendry et al. reported a statistically significant reduction in median length of hospital stay in the Doppler-monitored group (7 vs. 9 days in control group, p = 0.02) (Table E-5, Appendix E). However, the mean difference in hospital stay (11.4 vs. 13.9 days) was apparently not statistically significant. Furthermore, although this is a high-quality study, the mean between-treatment difference of 2 to 2.5 days means that the treatment effect cannot be judged as “large” (this would have required a minimum difference of at least 3 days). Therefore, no conclusion can be reached for this outcome.

Subsection Summary

After searching the literature, retrieving articles, and applying the inclusion/exclusion criteria, we identified one study that compared esophageal Doppler monitoring plus CVP plus conventional clinical assessment to CVP plus conventional clinical assessment for optimization of intravenous fluid replacement in patients admitted to intensive care following cardiac surgery. This study was judged to be of high quality based on ECRI ratings. Generalizability to the Medicare population was fair. However, this was a single small study without a demonstrably large treatment effect on the outcomes of interest. Therefore, no conclusions could be reached for this question.

Key Question 4: What Complications, Harms, and Adverse Events Associated with Esophageal Doppler Ultrasound Monitoring Have Been Reported?

Evidence Base

Our searches identified 75 studies that potentially addressed this question. Upon retrieval, 52 studies were found to either use TEE systems (rather than esophageal Doppler systems) or to contain no relevant information on complications and were therefore excluded. The remaining 23 studies that addressed this question are listed in Table 8.

Table 12. Evidence Base for Key Question 4

Study Design	References
Randomized controlled trials	Noblett et al. 2006(42); Conway et al. 2002(44); Venn et al.(46); Sinclair et al.(47)
Case series[a]	Cipolla et al. 2006(49); Collins et al. 2005(50); Koliopanos et al. 2005(51); Sawai et al. 2005(52); Sharma et al. 2005(53); Bein et al. 2004(54); Feldman et al. 2004(55); Iregui et al. 2003(56); Moxon et al. 2003(57); Seoudi et al. 2003(58); Su et al. 2002(59); Odenstedt et al. 2001(60); Madan et al. 1999(61); Elliott et al. 1998(62); Lefrant et al. 1998(18); Valtier et al. 1998(63); Singer et al. 1989(64)
Case reports	Chandan and Hull 2004(65)

[a] Some of these studies compared the accuracy of esophageal Doppler monitoring to other methods (e.g., thermodilution) within the same patient. However, for the purposes of evaluating esophageal probe-related complications, these studies are equivalent to case series (because every patient received esophageal Doppler probes).

Quality of Included Studies

Because the intent of this question is simply to list reported harms of esophageal Doppler ultrasound probes from any available data sources, we have not formally evaluated the quality of the evidence for this question. Uncontrolled studies or case reports cannot be used to determine causality or to estimate frequencies of adverse events; they can only be used to generate a list of adverse events possibly attributable to the device.

Details of Study Enrollees and Study Generalizability

For the reason described under quality of included studies, we do not present details of study enrollees or make judgments about study generalizability for the studies addressing this question.

Findings of Included Studies

Our searches identified one publication that reported two case reports of incorrect placement of an esophageal Doppler probe in the left main bronchus.(65) This led to adverse symptoms in only one of these patients. The cases are described in more detail in Table 9. An additional study of 106 critically ill patients reported accidental removal of an orogastric tube during esophageal probe removal. A study of 13 patients reported incorrect placement of an esophageal Doppler probe in one patient's trachea, but this did not cause any adverse effect. Finally, a study of 60 patients mentioned "occasional minimal trauma in the buccal cavity during placement of the esophageal probe" but did not state the number of patients who experienced this problem.

Our searches also identified 19 studies (4 RCTs and 15 case series) with a total of 654 patients that specifically stated that esophageal Doppler probes led to no complications in any of the patients included in these studies. The studies are listed in Table F- 1, Appendix F.

Another 33 studies that were retrieved did not report whether any complications were associated with esophageal Doppler probes. Without a statement to the effect that no complications occurred, one cannot determine whether the authors simply did not report complications or whether no complications occurred in these studies.

Our search of the FDA's Manufacturer and User Facility Device Experience (MAUDE) database identified only one report of a mechanical problem with the CardioQ Doppler probe. While a nurse was cleaning one of these probes with a tissue, the probe boot (distal end of the probe) separated from the rest of the probe body. However, this particular probe did not cause any complication in a patient.

It is noteworthy that no case of an esophageal or pharyngeal perforation has yet been reported in association with an esophageal Doppler probe. As noted in Key Question 1, such perforations have occasionally occurred during use of TEE probes. This may be because traditional TEE probes are larger than esophageal Doppler probes and tend to undergo more manipulation during monitoring.

Table 13. Studies Reporting Complications Associated with Esophageal Doppler Probes

References	Esophageal Doppler Probe (manufacturer)	Description of cases
Chandan and Hull 2004(65)	CardioQ (Deltex Medical)	A 68-year old man had increasing airway resistance and oxygen requirement several hours after insertion of the esophageal probe. A chest x-ray showed the probe to be in the left main bronchus, and it was promptly removed. The x-ray also showed increased pulmonary shadowing in the lower zones consistent with aspiration of regurgitated gastric fluid. It is likely that the esophageal Doppler probe caused the tracheal tube cuff to become incompetent. A 73-year old man received a chest x-ray for respiratory system review. The esophageal Doppler probe was found in the left main bronchus and promptly removed. There were no other changes to alert the staff to the incorrect placement of the probe.
Iregui et al. 2003(56)	ODM I (Deltex Medical) (earlier model of CardioQ)	In this study of 106 critically ill ICU patients, 1 patient had unintentional removal of an orogastric tube during removal of the esophageal Doppler probe. Forty patients received additional sedation during probe placement, but no side effects were reported.
Moxon et al. 2003(57)	HemoSonic 100 (Arrow International)	In this study of 13 patients, the esophageal Doppler probe was accidentally placed in the trachea in 1 patient. This was immediately recognized and corrected by the clinicians, and did not cause an adverse effect. No other complications were reported.
Singer et al. 1989(64)	Prototype esophageal transducer	In this study of 60 patients, no complications occurred other than occasional minimal trauma in the buccal cavity during placement of the esophageal probe (number of patients with minimal trauma not reported).

Subsection Summary

Currently, no serious adverse events associated with esophageal probes have been reported in the literature or the MAUDE database. The only minor events identified included two cases of incorrect probe placement in the left main bronchus, one case of incorrect placement in the trachea, a tube displacement during probe removal, and an unspecified number of cases of minimal trauma in the buccal cavity during probe placement. Nineteen studies with a total of 654 patients specifically stated that esophageal Doppler probes did not cause any complications. The number of patients represented in these studies is relatively small. However, the available evidence suggests that esophageal Doppler probes are relatively low-risk devices, as reporting of even minor morbidity has been infrequent thus far.

Conclusions

Conclusions for Key Question 1

Several methods are currently used to monitor cardiac output in patients during surgery or intensive care. These methods include thermodilution, dye dilution, lithium dilution, methods using the Fick principle, pulse contour methods, thoracic electrical bioimpedance, transesophageal echocardiography, and esophageal Doppler monitoring.

Conclusions for Key Question 2

The addition of esophageal Doppler monitoring for guided fluid replacement to a protocol using CVP and conventional clinical assessment during surgery leads to a clinically significant reduction in the rate of major and total complications in surgical patients compared to CVP plus conventional clinical assessment alone. The strength of evidence supporting this conclusion is strong. Because only three of five studies separately reported major complications, and because of differences in the way total complications were reported, no quantitative conclusion is presented for these outcomes.

The addition of esophageal Doppler monitoring to the protocol described above also reduces the length of hospital stay for surgical patients (clinical significance uncertain). The strength of evidence supporting this conclusion is strong. The lack of a calculable precise effect size in some studies precluded a quantitative summary estimate of the reduction in length of stay.

Only one study compared esophageal Doppler plus conventional clinical assessment to CVP plus conventional clinical assessment. Because this was one small study with non-informative effect sizes, no conclusions were possible for any of the outcomes of interest.

The addition of esophageal Doppler monitoring for guided fluid replacement to conventional clinical assessment during surgery leads to a clinically significant reduction

in the length of hospital stay compared to conventional clinical assessment alone. The strength of evidence supporting this conclusion is weak. The low number of studies precluded a quantitative estimate of the reduction in length of hospital stay. Because only a single study reported total complications, no conclusion was possible concerning this outcome.

No conclusion could be reached concerning relative mortality rates for any of the comparisons in Key Question 2.

The conclusions for Key Question 2 only apply to patients undergoing surgical procedures with an expected substantial blood loss or significant fluid compartment shifts requiring fluid replacement.

Conclusions for Key Question 3

The evidence base contained only one small study that was insufficient to allow conclusions to be reached about the effectiveness of esophageal Doppler monitoring in hospitalized patients in nonoperative settings.

Conclusions for Key Question 4

Currently, no serious adverse events associated with esophageal probes have been reported in the literature or the MAUDE database. The only minor events identified included two cases of incorrect probe placement in the left main bronchus, one case of incorrect placement in the trachea, a tube displacement during probe removal, and an unspecified number of cases of minimal trauma in the buccal cavity during probe placement. Nineteen studies with a total of 654 patients specifically stated that esophageal Doppler probes did not cause any complications. The number of patients represented in these studies is relatively small. However, the available evidence suggests that esophageal Doppler probes are relatively low-risk devices, as reporting of even minor morbidity has been infrequent thus far.

Bibliography

1. Cholley BP, Singer M. Esophageal Doppler: noninvasive cardiac output monitor. Echocardiography 2003 Nov;20(8):763-9.

2. Cholley BP, Payen D. Noninvasive techniques for measurements of cardiac output. Curr Opin Crit Care 2005 Oct;11(5):424-9.

3. Price JD, Sear JW, Venn RM. Perioperative fluid volume optimization following proximal femoral fracture. In: The Cochrane Database of Systematic Review 2004, Issue 1 [database online]. Hoboken (NJ): John Wiley & Sons, Ltd.[Art No.: CD003004]. Available: DOI: 10.1002/14651858.CD003004.pub2.

4. Hocking G. Central venous access and monitoring. Update Anaesthesia 2000;(12):article 13. Also available: http://www.nda.ox.ac.uk/wfsa/html/u12/u1213_01.htm.

5. Vinayak AG, Levitt J, Gehlbach B, Pohlman AS, Hall JB, Kress JP. Usefulness of the external jugular vein examination in detecting abnormal central venous pressure in critically ill patients. Arch Intern Med 2006 Oct 23;166(19):2132-7.

6. Cardiac output. [internet]. St. Petersburg (FL): Wikimedia Foundation Inc.; 2006 May 13 [accessed 2006 Jun 2]. [5 p]. Available: http://en.wikipedia.org/wiki/Cardiac_output.

7. Hett DA, Jonas MM. Non-invasive cardiac output monitoring. Intensive Crit Care Nurs 2004 Apr;20(2):103-8.

8. Conway D. (Consultant in Anaesthesia & Intensive Care. Manchester Royal Infirmary. Manchester, UK). Personal communication. 2006. 6 p.

9. Venn R. (Consultant in Anesthesia & Intensive Care. Worthing Hospital, Worthing UK). Personal communication. 2006. 4 p.

10. Dark PM, Singer M. The validity of trans-esophageal Doppler ultrasonography as a measure of cardiac output in critically ill adults. Intensive Care Med 2004 Nov;30(11):2060-6.

11. Gan TJ, Soppitt A, Maroof M, el-Moalem H, Robertson KM, Moretti E, Dwane P, Glass PS. Goal-directed intraoperative fluid administration reduces length of hospital stay after major surgery. Anesthesiology 2002 Oct;97(4):820-6.

12. Laupland KB, Bands CJ. Utility of esophageal Doppler as a minimally invasive hemodynamic monitor: a review. Can J Anaesth 2002;49(4):393-401.

13. Deltex Medical Group plc, preliminary results for year ended 31 December 2005. Chichester, UK: Deltex Medical Group; 2005. 14 p.

14. 510(k) premarket notification database [database online]. Rockville (MD): U.S. Food and Drug Administration, Center for Devices and Radiological Health; 1995 Nov 7 [updated 2006 Oct 6]; [accessed 2006 Oct 9]. Flowmeter, blood, cardiovascular. K951369. [1 p]. Available: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMN/pmn.cfm?ID=95874.

15. 510(k) summary. CardioQ cardiac output and fluid status monitoring system. K031706. Chichester, UK: Deltex Medical Limited; 2003 Aug 6. 8 p.

16. Sometic, Inc. 510(k) premarket notification database. Flowmeter, blood, cardiovascuar. Dynemo 3000. K972798. [internet]. Rockville (MD): U.S. Food and Drug Administration (FDA); 1998 [2 p]. Available: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfPMN/pmn.cfm?ID=114313.

17. Deltex Medical Group. Acquisition of TECO. Chichester (UK): Deltex Medical Group plc; 2006 Feb 7. 2 p.

18. Lefrant JY, Bruelle P, Aya AG, Saissi G, Dauzat M, de La Coussaye JE, Eledjam JJ. Training is required to improve the reliability of esophageal Doppler to measure cardiac output in critically ill patients. Intensive Care Med 1998 Apr;24(4):347-52.

19. NCD for ultrasound diagnostic procedures (220.5). [internet]. Baltimore (MD): Centers for Medicare & Medicaid Services (CMS); [accessed 2005 Sep 23]. [3 p].

20. Aetna. Clinical policy bulletin number: 0008. Color-flow doppler echocardiography in adults . [internet]. Hartford (CT): Aetna, Inc.; 2006 Jun 9 [accessed 2006 Jun 20]. [6 p]. Available: http://www.aetna.com/cpb/data/CPBA0008.html.

21. Blue Cross and Blue Shield of Massachusetts. Echocardiography. Policy no. 108. [internet]. Rockland (MA): Blue Cross and Blue Shield of Massachusetts; 2006 May 16 [accessed 2006 Jun 20]. [7 p]. Available: http://www.bluecrossma.com/common/en_US/medical_policies/108%20Echocardiography%20prn.pdf.

22. Blue Cross and Blue Shield of North Carolina. Intraoperative transesophageal echocardiography. Policy No. SUR6410. [internet]. Durham (NC): Blue Cross and Blue Shield of North Carolina; 1996 Apr 1 [updated 2005 Oct 8]; [accessed 2006 Jun 20]. [5 p]. Available: http://www.bcbsnc.com/services/medical-policy/pdf/intraoperative_transesophageal_echocardiography.pdf.

23. Moher D, Pham, Klassen TP, Schulz KF, Berlin JA, Jadad AR, Liberati A. What contributions do languages other than english make on the results of meta-analyses? J Clin Epidemiol 2000 Sep;53(9):964-72.

24. Juni P, Holenstein F, Sterne J, Bartlett C, Egger M. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. Int J Epidemiol 2002 Feb;31(1):115-23.

25. Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schunemann HJ, Edejer TT, Varonen H, Vist GE, Williams JW Jr, Zaza S. Grading quality of evidence and strength of recommendations. BMJ 2004 Jun 19;328(7454):1490. Also available: http://bmj.bmjjournals.com/cgi/reprint/328/7454/1490.

26. Snedecor GW, Cochran WG. Statistical methods. 8th ed. Ames (IA): Iowa State University Press; 1989. 503 p.

27. Pudar Hozo S, Djulbegovic B, Hozo I. Estimating the mean and variance from the median, range, and the size of a sample. BMC Med Res Methodol 2005 Apr 20;5(1):13.

28. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-58.

29. Hedges LV. Fixed effects models. In: Cooper H, Hedges LV, editors. The handbook of research synthesis. New York (NY): Russell Sage Foundation; 1994. p. 285-99.

30. Olkin I. Diagnostic statistical procedures in medical meta-analysis. Stat Med 1999 Sep 15;18(17-18):2331-41.

31. Alberta Heritage Foundation for Medical Research. USCOM for cardiac output monitoring. Edmonton, Alberta: Alberta Heritage Foundation for Medical Research; 2006 May. 17 p. (TechNote Series; no. 56).

32. Thermodilution. [internet]. Biology-Online.org; 2005 Oct [accessed 2006 Jun 1]. [1 p]. Available: http://www.biology-online.org.

33. Gawlinski A. Measuring cardiac output: intermittent bolus thermodilution method. [internet]. Aliso Viejo (CA): American Association of Critical-Care Nurses (AACN); 2004 Oct [accessed 2006 Jun 1]. [5 p]. Available: http://www.findarticles.com/articles/mi_m0NUC/is_5_24/ai_n6332747/print.

34. Ott K, Johnson K, Ahrens T. New technologies in the assessment of hemodynamic parameters. J Cardiovasc Nurs 2001 Jan;15(2):41-55.

35. Cardiac output measurement by indicator dilution. [internet]. Victoria (AU): Australasian Society of Cardio-Vascular Perfusionists; [accessed 2006 Jun 5]. [3 p]. Available: http://www.perfusion.com.au.

36. LiDCO plus frequently asked questions & features and benefits. [internet]. Cambridge (UK): LiDCO Ltd.; [accessed 2006 Jun 5]. [8 p]. Available: http://www.lidco.com/html/technology/faqs.asp.

37. Berton C, Cholley B. Equipment review: new techniques for cardiac output measurement--oesophageal Doppler, Fick principle using carbon dioxide, and pulse contour analysis. Crit Care 2002 Jun;6(3):216-21.

38. Scolletta S, Romano SM, Biagioli B, Capannini G, Giomarelli P. Pressure recording analytical method (PRAM) for measurement of cardiac output during various haemodynamic states. Br J Anaesth 2005 Aug;95(2):159-65.

39. Humana. Thoracic electrical bioimpedance (impedance cardiography). [internet]. Louisville (KY): Humana, Inc.; 2004 Oct 21 [accessed 2006 Jun 7]. [5 p]. Available: http://www.humana.com.

40. Min JK, Spencer KT, Furlong KT, DeCara JM, Sugeng L, Ward RP, Lang RM. Clinical features of complications from transesophageal echocardiography: a single-center case series of 10,000 consecutive examinations. J Am Soc Echocardiogr 2005 Sep;18(9):925-9.

41. Tan HL, Pinder M, Parsons R, Roberts B, van Heerden PV. Clinical evaluation of USCOM ultrasonic cardiac output monitor in cardiac surgical patients in intensive care unit. Br J Anaesth 2005 Mar;94(3):287-91.

42. Noblett SE, Snowden CP, Shenton BK, Horgan AF. Randomized clinical trial assessing the effect of doppler-optimized fluid management on outcome after elective colorectal resection. Br J Surg 2006 Sep;93(9):1069-76.

43. Wakeling HG, McFall MR, Jenkins CS, Woods WG, Miles WF, Barclay GR, Fleming SC. Intraoperative oesophageal Doppler guided fluid management shortens postoperative hospital stay after major bowel surgery. Br J Anaesth 2005 Nov;95(5):634-42.

44. Conway DH, Mayall R, Abdul-Latif MS, Gilligan S, Tackaberry C. Randomised controlled trial investigating the influence of intravenous fluid titration using oesophageal Doppler monitoring during bowel surgery. Anaesthesia 2002 Sep;57(9):845-9.

45. Mythen MG, Webb AR. Perioperative plasma volume expansion reduces the incidence of gut mucosal hypoperfusion during cardiac surgery. Arch Surg 1995 Apr;130(4):423-9.

46. Venn R, Steele A, Richardson P, Poloniecki J, Grounds M, Newman P. Randomized controlled trial to investigate influence of the fluid challenge on duration of hospital stay and perioperative morbidity in patients with hip fractures. Br J Anaesth 2002 Jan;88(1):65-71.

47. Sinclair S, James S, Singer M. Intraoperative intravascular volume optimisation and length of hospital stay after repair of proximal femoral fracture: randomised controlled trial. BMJ 1997 Oct 11;315(7113):909-12.

48. McKendry M, McGloin H, Saberi D, Caudwell L, Brady AR, Singer M. Randomised controlled trial assessing the impact of a nurse delivered, flow monitored protocol for optimisation of circulatory status after cardiac surgery. BMJ 2004 Jul 31;329(7460):258.

49. Cipolla J, Stawicki S, Spatz D. Hemodynamic monitoring of organ donors: a novel use of the esophageal echo-Doppler probe. Am Surg 2006 Jun;72(6):500-4.

50. Collins S, Girard F, Boudreault D, Chouinard P, Normandin L, Couture P, Caron M-J, Ruel M. Esophageal doppler and thermodilution are not interchangeable for determination of cardiac output. Can J Anaesth 2005;52(9):978-85.

51. Koliopanos A, Zografos G, Skiathitis S, Stithos D, Voukena V, Karampinis A, Papastratis G. Esophageal Doppler (ODM II) improves intraoperative hemodynamic monitoring during laparoscopic surgery. Surg Laparosc Endosc Percutan Tech 2005 Dec;15(6):332-8.

52. Sawai T, Nohmi T, Ohnishi Y, Takauchi Y, Kuro M. Cardiac output measurement using the transesophageal Doppler method is less accurate than the thermodilution method when changing PaCO2. Anesth Analg 2005 Dec;101(6):1597-601.

53. Sharma J, Bhise M, Singh A, Mehta Y, Trehan N. Hemodynamic measurements after cardiac surgery: transesophageal Doppler versus pulmonary artery catheter. J Cardiothorac Vasc Anesth 2005 Dec;19(6):746-50.

54. Bein B, Worthmann F, Tonner PH, Paris A, Steinfath M, Hedderich J, Scholz J. Comparison of esophageal Doppler, pulse contour analysis, and real-time pulmonary artery thermodilution for the continuous measurement of cardiac output. J Cardiothorac Vasc Anesth 2004 Apr;18(2):185-9.

55. Feldman LS, Anidjar M, Metrakos P, Stanbridge D, Fried GM, Carli F. Optimization of cardiac preload during laparoscopic donor nephrectomy: a preliminary study of central venous pressure vs esophageal Doppler monitoring. Surg Endosc 2004;18(3):412-6.

56. Iregui MG, Prentice D, Sherman G, Schallom L, Sona C, Kollef MH. Physicians' estimates of cardiac index and intravascular volume based on clinical assessment versus transesophageal Doppler measurements obtained by critical care nurses. Am J Crit Care 2003 Jul;12(4):336-42.

57. Moxon D, Pinder M, van Heerden PV, Parsons RW. Clinical evaluation of the HemoSonic monitor in cardiac surgical patients in the ICU. Anaesth Intensive Care 2003 Aug;31(4):408-11.

58. Seoudi HM, Perkal MF, Hanrahan A, Angood PB. The esophageal Doppler monitor in mechanically ventilated surgical patients: does it work? J Trauma 2003 Oct;55(4):720-5; discussion 725-6.

59. Su NY, Huang CJ, Tsai P, Hsu YW, Hung YC, Cheng CR. Cardiac output measurement during cardiac surgery: esophageal Doppler versus pulmonary artery catheter. Acta Anaesthesiol Sin 2002;40(3):127-33.

60. Odenstedt H, Aneman A, Oi Y, Svensson M, Stenqvist O, Lundin S. Descending aortic blood flow and cardiac output: a clinical and experimental study of continuous oesophageal echo-Doppler flowmetry. Acta Anaesthesiol Scand 2001 Feb;45(2):180-7.

61. Madan AK, UyBarreta VV, Aliabadi-Wahle S, Jesperson R, Hartz RS, Flint LM, Steinberg SM, Dulchavsky SA, Kuwagata Y, Bergstein J. Esophageal doppler ultrasound monitor versus pulmonary artery catheter in the hemodynamic management of critically III surgical patients. J Trauma Inj Infect Crit Care 1999;46(4):607-12.

62. Elliott S, Savill P, Eckersall S. Cardiovascular changes during laparoscopic cholecystectomy: a study using transoesophageal Doppler monitoring. Eur J Anaesthesiol 1998 Jan;15(1):50-5.

63. Valtier B, Cholley BP, Belot JP, de la Coussaye JE, Mateo J, Payen DM. Noninvasive monitoring of cardiac output in critically ill patients using transesophageal Doppler. Am J Respir Crit Care Med 1998 Jul;158(1):77-83.

64. Singer M, Clarke J, Bennett ED. Continuous hemodynamic monitoring by esophageal Doppler. Crit Care Med 1989 May;17(5):447-52.

65. Chandan GS, Hull JM. Incorrectly placed oesophageal Doppler probe. Anaesthesia 2004;59:723.

66. West S, King V, Carey TS, Lohr KN, McKoy N, Sutton SF, Lux L. Systems to rate the strength of scientific evidence. (Prepared by Research Triangle Institute - University of North Carolina Evidence-based Practice Center under Contract no. 290-97-0011). AHRQ Publication no. 02-E016. Rockville (MD): Agency for Healthcare Research and Quality (AHRQ); 2002 Apr. 199 p. (Evidence report/technology assessment; no. 47). Also available: http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=hstat1.chapter.70996.

67. Higgins JP, Thompson SG. Controlling the risk of spurious findings from meta-regression. Stat Med 2004 Jun 15;23(11):1663-82.

68. Harbord R, Statistical Software Components, Boston College Department of Economics. METAREG: Stata module to perform meta-analysis regression. [software download]. Stockholm, Sweden: EconPapers; 2004 Sep 22 [accessed 2004 Nov 3]. Available: http://econpapers.hhs.se/software/bocbocode/s446201.htm.

69. Cochrane Collaboration. Cochrane reviewers' handbook. Version 4.2.2. Oxford (UK): Cochrane Collaboration; 2004 Mar. 234 p. Also available: http://www.cochrane.org/resources/handbook/index.htm.

70. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale (NJ): Lawrence Erlbaum Associates; 1988. 567 p.

71. Krishnamurthy B, McMurray TJ, McClean E. The peri-operative use of the oesophageal Doppler monitor in patients undergoing coronary artery revascularisation. A comparison with the continuous cardiac output monitor. Anaesthesia 1997 Jul;52(7):624-9.

APPENDICES: SUPPORTING DOCUMENTATION AND EVIDENCE TABLES

Appendix A. Literature Searches

Electronic Database Searches

To obtain information for this report, we searched the following databases for relevant information:

Database	Date limits	Platform/provider
CINAHL (Cumulative Index to Nursing and Allied Health Literature)	1982 through September 11, 2006	OVID
The Cochrane Central Register of Controlled Trials (CENTRAL)	Inception through 2006, Issue 3	www.thecochranelibrary.com
The Cochrane Database of Methodology Reviews (Methodology Reviews)	Inception through 2006, Issue 3	www.thecochranelibrary.com
The Cochrane Database of Systematic Reviews (Cochrane Reviews)	Inception through 2006, Issue 3	www.thecochranelibrary.com
Database of Abstracts of Reviews of Effects (DARE)	Inception through 2006, Issue 3	www.thecochranelibrary.com
ECRI Health Devices Alerts	1977 through June 7, 2006	ECRI
ECRI International Health Technology Assessment (IHTA)	Inception through June 7, 2006	ECRI
ECRI Library Catalog	Inception through March 2006	ECRI
Embase (Excerpta Medica)	1974 through September 11, 2006	OVID
Health Technology Assessment Database (HTA)	Inception through 2006, Issue 3	www.thecochranelibrary.com
MEDLINE	1966 through September 11, 2006	OVID
metaRegister of Controlled Trials (mRCT)	Searched June 14, 2006	http://www.controlled-trials.com/mrct/
PubMed (PreMEDLINE, Publisher)	Through September 11, 2006	www.pubmed.gov
U.K. National Health Service Economic Evaluation Database (NHS EED)	Inception through 2006, Issue 3	www.thecochranelibrary.com
U.S. Centers for Medicare & Medicaid (CMS) Web site	Inception through June 19, 2006	www.cms.gov Mediregs (www.coverageandpayment.com)
U.S. Food and Drug Administration (FDA) (adverse event reports)	1977 through June 7, 2006	www.fda.gov www.ecri.org
U.S. National Guideline Clearinghouse™ (NGC™)	Through June 14, 2006	www.ngc.gov

Hand Searches of Journal and Nonjournal Literature

Journals and supplements maintained in ECRI's collections were routinely reviewed. Nonjournal publications and conference proceedings from professional organizations, private agencies, and government agencies were also screened. Other mechanisms used to retrieve additional relevant information included review of bibliographies/reference lists from peer-reviewed and gray literature. (Gray literature consists of reports, studies, articles, and monographs produced by federal and local government agencies, private organizations, educational facilities, consulting firms, and corporations. These documents do not appear in the peer-reviewed journal literature).

Search Strategies

The search strategies employed combinations of freetext keywords as well as controlled vocabulary terms including (but not limited to) the following concepts. The strategy below is presented in OVID syntax; the search was simultaneously conducted across Embase, Medline, and PsycINFO. A parallel strategy was used to search the databases comprising the Cochrane Library.

Medical Subject Headings (MeSH), Emtree, PsycINFO and Keywords

Conventions:

OVID

$ = truncation character (wildcard)

exp = "explodes" controlled vocabulary term (e.g., expands search to all more specific related terms in the vocabulary's hierarchy).

.de. = limit controlled vocabulary heading

.fs. = floating subheading

.hw. = limit to heading word

.md. = type of methodology (PsycINFO)

.mp. = combined search fields (default if no fields are specified)

.pt. = publication Type

.ti. = limit to title

.tw. = limit to title and abstract fields

PubMed

[mh] = MeSH heading

[majr] = MeSH heading designated as major topic

[pt] = Publication Type

[sb] = Subset of PubMed database (PreMedline, Systematic, OldMedline)

[sh] = MeSH subheading (qualifiers used in conjunction with MeSH headings)

[tiab] = keyword in title or abstract

[tw] = Text word

CINAHL/Embase/Medline

(English language, human)

Set Number	Concept	Search statement
1	Cardiac output	exp heart output/ or exp cardiac output/
2	Hemodynamic monitoring	(hemodynamic monitoring or haemodynamic monitoring or hemodynamic assessment or haemodynamic assessment)
3	Esophageal doppler	((esophag$ or oesophag$ or trans?esoph$) and (doppler or ultrasound or son$ or ultrason$ or echocard$)).
4	Esophageal doppler (controlled vocabulary)	exp transesophageal echocardiography/ or exp echocardiography, transesophageal/
5	Combine sets	1 or 2
6	Combine sets	3 or 4
7	Device names	HemoSonic or CardioQ or ODM II or Waki.ti. or Waki.ab. or Dynemo 3000
8	Combine sets	or/5-7
9	Limit by publication type	8 not ((letter or editorial or news or comment or case reports or review or note or conference paper).de. or (letter or editorial or news or comment or case reports or review).pt.)
10	Limit by study type	9 and ((Randomized controlled trials or random allocation or double-blind method or single-blind method or placebos or cross-over studies or crossover procedure or double blind procedure or single blind procedure or placebo or latin square design or crossover design or double-blind studies or single-blind studies or triple-blind studies or random assignment or exp controlled study/ or exp clinical trial/ or exp comparative study/ or cohort analysis or follow-up studies.de. or intermethod comparison or parallel design or control group or prospective study or retrospective study or case control study or major clinical study).de. or random$.hw. or random$.ti. or placebo$.mp. or ((singl$ or doubl$ or tripl$ or trebl$) and (dummy or blind or sham)).mp. or latin square.mp. or ISRTCN.mp.)

PreMedline (PubMed)

(English language)

Set Number	Concept	Search statement
1	Doppler cardiac output monitoring	(“cardiac output” OR “hemodynamic assessment” OR “haemodynamic assessment”) AND (doppler OR ultrasound OR ultrason* OR echocard*)
2	transesophageal	#1 AND (esophag* OR oesophag* OR transesophag* OR transoesophag*)
3	Combine sets	#1 AND #2
4	Device names	HemoSonic OR CardioQ OR “ODM II” OR “Waki” [tiab] OR “Dynemo 3000”
5	Combine sets	#3 OR #4
6	Limit to Premedline subfile	#5 AND premedline [sb]

Appendix B. Quality of Literature and Evidence Strength Rating

Study Quality Scale

A poorly designed study may contain biases that may make a treatment look more or less effective than it actually is. In well-designed studies, the outcomes can be definitively attributed to the treatment of interest.

In order to grade the quality of studies, we use a quality rating scale. This scale allows us to calculate an evidence quality score based on *a priori* quality criteria. The questions in the scale are worded so that study design aspects that provide evidence with good internal validity result in "Yes" answers, design aspects that create potential for bias result in "No", and design aspects that are inadequately described result in an answer of "NR" (not reported).

The 25-item quality assessment instrument used to assess the quality of the three studies that addressed Key Question 3 is presented below:

Comparability of Groups at Baseline

1. Were patients randomly assigned to the study's groups?
2. Did the study employ stochastic randomization?
3. Were any methods other than randomization used to make the patients in the study's groups comparable?
4. Were patients assigned to groups based on factors other than patient or physician preference?
5. Were the characteristics of patients in the different study groups comparable at the time they were assigned to groups?
6. Did patients in the different study groups have similar levels of performance on all of the outcome variables at the time they were assigned to groups?
7. Was the comparison of interest prospectively planned?
8. Did ≥85% of the patients compete they study?
9. Was there a ≤15% difference in completion rates in the study's groups?

10. Were all of the study's groups concurrently treated?
11. Was compliance with treatment ≥85% in both of the study's groups?
12. Was there concealment of allocation?

Blinding

13. Were subjects blinded to the treatment they received?
14. Did the authors perform any tests after completing the study to ensure that the integrity of the blinding of patients was maintained throughout the study?
15. Was the treating physician blinded to the groups to which the patients were assigned?
16. Were those who assessed the patient's outcomes blinded to the group to which the patients were assigned?

Measurement/Instrument

17. Was the outcome measure of interest objective and was it objectively measured?
18. Were the same laboratory tests, clinical findings, psychological instruments, etc., used to measure the outcomes in all of the study's groups?
19. Was the instrument used to measure the outcome standard?
20. Were the follow-up times in all of the study's relevant groups approximately equal?

Treatment

21. Was the same treatment given to all patients enrolled in the experimental group?
22. Was the same treatment given to all patients enrolled in the control group?
23. Were all of the study's groups treated at the same center?

Investigator Bias

24. Was the funding for this study derived from a source that does not have a financial interest in its results?

25. Were the author's conclusions, as stated in the abstract or the article's discussion section, supported by the data presented in the article's results section?

We used these items to compute a summary score, which ranges from 0 to 10, where 10 indicates an ideal study and 0 indicates a study of the poorest possible quality. To compute this summary score, we made the following calculations. We first converted the individual item answers to numeric scores by counting 1 for each Yes answer, -1 for each No, and -0.5 for each NR. We then added the numeric scores for all 25 items, added 25 to the total, divided by 50, and multiplied by 10. These calculations yield the 0-10 summary scale described above. Studies that scored less than 5 were considered unacceptable quality, greater than 5 but less than or equal to 6.7 were considered low quality, greater than 6.7 but less than or equal to 8.4 were considered moderate quality, and 8.5 or greater were considered high quality.

Strength of Evidence Algorithm

After grading the body of evidence for a particular question on each of several decision points (listed in the next sections), we apply the grades to an algorithm that divides the strength of the evidence supporting each conclusion into one of four categories: strong, moderate, weak, or inconclusive. Table B-1 illustrates how these categories relate to qualitative and quantitative conclusions.

Table B-1. Interpretation of Different Categories of Strength of Evidence Supporting Conclusion

Strength of Evidence	Interpretation of Qualitative Conclusion
Strong	Evidence supporting the qualitative conclusion is convincing. It is highly unlikely that new evidence will lead to a change in this conclusion.
Moderate	Evidence supporting the qualitative conclusion is somewhat convincing. There is a small chance that new evidence will overturn or strengthen our conclusion. ECRI recommends regular monitoring of the relevant literature at this time.
Weak	Although some evidence exists to support the qualitative conclusion, this evidence is tentative and perishable. There is a reasonable chance that new evidence will overturn or strengthen our conclusions. ECRI recommends frequent monitoring of the relevant literature at this time.

Strength of Evidence	Interpretation of Qualitative Conclusion
Inconclusive	Although some evidence exists, this evidence is not of sufficient strength to warrant drawing an evidence-based conclusion from it. ECRI recommends frequent monitoring of the relevant literature at this time.
Stability of Evidence	Interpretation of Quantitative Conclusion
High	The estimate of treatment effect included in the conclusion is stable. It is highly unlikely that the magnitude of this estimate will change substantially as a result of the publication of new evidence.
Moderate	The estimate of treatment effect included in the conclusion is somewhat stable. There is a small chance that the magnitude of this estimate will change substantially as a result of the publication of new evidence. ECRI recommends regular monitoring of the relevant literature at this time.
Low	The estimate of treatment effect included in the conclusion is likely to be unstable. There is a reasonable chance that the magnitude of this estimate will change substantially as a result of the publication of new evidence. ECRI recommends frequent monitoring of the relevant literature at this time.
Inconclusive	Estimates of the treatment effect are too unstable to allow a quantitative conclusion to be drawn at this time. ECRI recommends frequent monitoring of the relevant literature.

To arrive at these strength-of-evidence categories, we applied the ECRI Strength of Evidence Algorithm. This algorithm, which appears in Figure 3 through Figure 6 below, involves 10 decision points. The methods we used to resolve these 10 decision points appear next.

Decision Point #1: Acceptable quality

The above section entitled Study Quality Scale describes our approach to determining whether each study was of acceptable quality.

Decision Point #2: Overall quality

After assigning quality scores to each individual study, we then classified the overall quality of the evidence base by taking the median of the Overall quality scores. Quality scores were converted to categories as defined in Table B-2. For example, if the evidence base consists of four studies with overall scores of 6.5, 7, 7.8, and 9, then the median is 7.4 and the overall evidence base is considered moderate quality.
The definitions for what constitutes low, moderate, or high quality evidence were determined *a priori* by a committee of three methodologists. If the median quality was on the border between categories, we took the lower quality category as the overall quality.

Table B-2. Categorization of Quality

	High quality	Moderate quality	Low quality
Median quality	>8.4	>6.7 but ≤8.4	>5 but ≤6.7

Decision Point #3: Does Reporting Allow Quantitative Analysis to be Performed?

The answer to Decision Point 3 depends upon the adequacy of reporting in available studies as well as the number of available studies. In order to conduct a quantitative analysis of a given outcome, the data for that outcome must be reported in at least three studies in a manner that allows the data to be pooled in a meta-analysis. If less than three studies are available, no quantitative conclusion is usually possible regardless of reporting. Another situation that does not allow a quantitative conclusion is when three or more studies are available, but fewer than 80% of them permit determination of the effect size and its dispersion, either by direct reporting from the trial or calculations based on reported information. Finally, no quantitative conclusion is possible if fewer than 80% of available studies report a given outcome. If no quantitative conclusion is possible, then one moves directly to Decision Point 8 to begin a qualitative analysis.

Decision Point #4: Are Data Quantitatively Consistent (Homogeneous)?

This decision point is used only if the answer to Decision Point 3 was Yes. Consistency refers to the extent to which the results of studies in an evidence base agree with each other.(66) The more consistent the evidence, the more precise a summary estimate of treatment effect derived from the evidence base. Quantitative consistency refers to consistency tested in a meta-analysis using the Q statistic and Higgins and Thompson's I^2 statistic.(28) We consider the evidence base to be quantitatively consistent when I^2 <50% and the p-value of Q is ≥0.10 (both criteria must be met).

If the studies are homogeneous, we combine the results in a fixed-effects meta-analysis (FEMA). We then determine whether the summary effect size is informative or non-informative. The summary effect is considered informative if it meets any one of the following three criteria:

1) The summary effect is statistically significant.
2) If the minimum boundary of clinical significance is greater than 0, the 95% confidence intervals of the summary effect must exclude the possibility of a clinically significant effect.
3) If the boundary of clinical significance equals 0 (clinical significance = statistical significance), the 95% confidence intervals of the summary effect must not overlap with -0.2 or +0.2 (this assumes one is using Hedges' d or Cohen's h as the meta-analytic summary statistic; for the log odds ratio, the interval is -0.4 to +0.4).

Criteria 2) and 3) require definitions of the minimum difference between treatments (or between baseline and post-treatment measurements) that is considered clinically significant. The definitions that we used appear in Table B-3.

Table B-3. Definitions of clinical significance

Outcome	Minimum effect considered to be clinically significant
Key Question 2	
Complications (including mortality)	Any statistically significant difference
Hospital stay	1 day (difference between treatment groups)
Key Question 3	
Complications (including mortality)	Any statistically significant difference
Hospital stay	1 day (difference between treatment groups)

We did not identify any clinical consensus in the literature regarding what constitutes a clinically significant difference in the length of hospital stay (either as an absolute number of days or as a fraction of the expected length of stay). Therefore, our choice of one day as the minimum clinically significant difference for hospital stay is subjective. Our assumption is that if the average duration of stay is less than two weeks, a difference of at least one day would be important to patients. Also, every additional day a patient stays in a hospital increases their risk of acquiring a nosocomial infection. However, other reviewers might consider a shorter or longer difference in length of stay to be clinically significant.

If the summary effect is informative, we then test the stability of the findings in decision point 5.

Decision Point #5: Are Findings Stable (Quantitatively Robust)?

Stability of findings refers to the likelihood that a summary effect estimate will be substantially altered by changing the conditions of the analysis. This was tested by first removing each individual study separately to see if any single study had a substantial influence on the summary result. Secondly we performed cumulative meta-analysis.

A pre-requisite of an analysis of quantitative robustness is that the 95% confidence interval around a meta-analytic effect size should not exceed a certain range. If the difference between the upper and lower bounds is GREATER than 0.4, then the

estimate is automatically considered not robust (no analysis necessary). If it is less than or equal to 0.4, then perform an analysis of quantitative robustness. This number (0.4) is based on the use of 0.2 as the minimum clinically important effect. Thus, if the confidence interval width is less than 0.4, then the point estimate must be within 1 unit of clinical significance, which would pass this initial pre-requisite. This number also assumes that one is using either Hedges' d or Cohen's h as the measure of effect size. If a different effect size measure is being used, one would change the number accordingly. When using the log odds ratio, the interval becomes 0.8. We refer to the point estimate of the meta-analytic summary statistic as $SES_{full.}$

1) Compute SES_{full} +0.2 and SES_{full} -0.2. These will denote the two horizontal dashed lines in your cumulative meta-analysis plot, to represent the range of acceptable deviation from SES_{full}.
2) Determine which study had the lowest weight in the meta-analysis (usually this will be the smallest study). Remove that study, and recompute the SES and its confidence bounds.
3) If EITHER the upper bound is greater than SES_{full} +0.2, OR the lower bound is less than SES_{full} -0.2, then the estimate is not robust. On the plot, this is equivalent to whether the CI crosses EITHER dashed line. If neither line is crossed, continue to the next step.
4) Determine which study had the next lowest weight in the meta-analysis (usually this will be the next-to-smallest study). Remove that study also, and recompute the SES and its confidence bounds. This meta-analysis is based on k-2 studies.
5) If EITHER the upper bound is greater than SES_{full} +0.2, OR the lower bound is less than SES_{full} -0.2, then the estimate is not robust. On the plot, this is equivalent to whether the CI crosses EITHER dashed line.
6) If EITHER of the above meta-analyses changes the summary effect estimate by ≥ 10%, then the estimate is not robust.
7) If EITHER of the above meta-analyses shows significant heterogeneity (either an I^2 ≥50% or a Q with a p-value <0.10), then the estimate is not robust.

Decision Point #6: Meta-regression Explains Heterogeneity?

Meta-analyses with heterogeneity are further evaluated with meta-regression. Meta-regression is not performed on a low quality evidence base. Heterogeneity was assessed by meta-regression using the permutation test method of Higgins and Thompson (2004)(67) and the meta-regression module in the Stata software package.(68) Meta-regression was only performed if there were 10 or more studies in an evidence base with an average quality that was moderate or high, and if ≥80% of the studies had data allowing effect sizes to be calculated without imputation.(69) Because none of the evidence bases evaluated in this report had 10 or more studies, meta-regression could not be performed.

Decision Point #7: Meta-regression Model Robust?

If heterogeneity can be explained with meta-regression does the model hold through sensitivity testing? Testing would involve removal of each individual study from the meta-regression to determine whether removal of any single study changes the results of the meta-regression. However, no meta-regressions were performed in this report.

Decision Point #8: Qualitatively Robust?

If the evidence base for an outcome had three or more studies, we determined whether the qualitative findings could be overturned by removal of any single study, changing the measure of effect, or cumulative meta-analysis. We considered findings to be overturned only when a study removal altered the conclusion (i.e., a statistically significant finding becomes non-significant as studies are added to the evidence base). However, the analysis differs somewhat depending upon which of the following four bolded questions is being addressed.

Is the qualitative finding that the effect size is different from 0 robust?

The steps below are taken only if the all-study summary SES is statistically significant, because if it were not, then one must be testing the conclusion that there is no clinically significant difference, which is addressed in the 4th bolded question below.

1) Compute the meta-analysis based on the separate removal of each of the studies (if there are four studies, this will be four separate meta-analyses). If all of the meta-analyses are statistically significant in the same direction as the full MA, then the conclusion is qualitatively robust.
2) Compute the meta-analysis based on the cumulative removal of studies (the last two from smallest to largest and from largest to smallest). If all of the meta-analyses are statistically significant in the same direction as the full MA, then the conclusion is qualitatively robust.

Is the qualitative finding that the effect size is clinically significant robust?

The steps below are taken only if the all-study summary SES lies fully above the line of clinical significance (i.e., the lower bound of the CI is greater than the line of clinical significance).

1) Compute the meta-analysis based on the separate removal of each of the studies. If the lower bound of the SES is above the line of clinical significance for each meta-analysis, then the conclusion is qualitatively robust.
2) Compute the meta-analysis based on the cumulative removal of studies (the last two from smallest to largest and from largest to smallest). If the lower bound of the SES is above the line of clinical significance for each meta-analysis, then the conclusion is qualitatively robust.

Is the qualitative finding that the effect size is not clinically significant robust?

The steps below are taken only if the all-study summary SES lies fully below the line of clinical significance (i.e., the upper bound of the CI is lower than the line of clinical significance).

1) Compute the meta-analysis based on the separate removal of each of the studies. If the upper bound of the SES is below the line of clinical significance, then the conclusion is qualitatively robust.

2) Compute the meta-analysis based on the cumulative removal of studies (the last two from smallest to largest and from largest to smallest). If the upper bound of the SES is below the line of clinical significance, then the conclusion is qualitatively robust.

Is the qualitative finding that the effect size is not substantial robust?

The steps below are taken only if the cutoff for clinical significance is identical to the cutoff for statistical significance (zero) and if the confidence intervals around the all-study summary SES overlap with zero. This cutoff for clinical significance is generally used for severe adverse events (for example, when the outcome is death, even a miniscule difference is clinically important, thus zero is taken as the cutoff for clinical significance). Note: for this bolded question, if the upper or lower bound of the SES overlaps with -0.2 or 0.2 (again assuming the effect size measure is Hedges' d or Cohen's h; for the log odds ratio, the interval is -0.4 to 0.4), then no qualitative conclusion is possible (the evidence base is inconclusive). Otherwise, follow the steps below.

1) Compute the meta-analysis based on the separate removal of each of the studies. If the upper or lower bound of the SES still overlaps zero and both are between -0.2 and 0.2, then it is robust.
2) Compute the meta-analysis based on the cumulative removal of studies (the last two from smallest to largest and from largest to smallest). If the upper or lower bound of the SES still overlaps zero and both are between -0.2 and 0.2, then it is robust.

Decision Point #9: Qualitatively Consistent?

This Decision Point is used only when the evidence base for an outcome consists of two studies. For a given outcome, studies were considered qualitatively consistent if both studies had a statistically significant effect in the same direction, or if both studies did not have a statistically significant effect.

Decision Point #10: Magnitude of Effect Extremely Large?

When considering the strength of evidence supporting a qualitative conclusion based on only one or two studies, magnitude of effect becomes very important. If a single study finds a large effect with a narrow confidence interval, then new evidence is unlikely to overturn the qualitative conclusion. To resolve this decision point, we consulted the effect size and the 95% confidence interval around the effect size for the study (with two studies, we consulted the interval around the random effects summary statistic). If this interval was fully above +0.5 (or if it was fully below -0.5) and the effect size was ≥0.8 (or ≤-0.8), we considered the effect to be large. Otherwise, we considered it to be not large. For example, an interval from +0.6 to +1.1 would be considered a large effect, whereas an interval from +0.4 to +1.3 would not be considered a large effect.
Another effect that would be considered large is an interval from -1.1 to -0.6 (large in the negative direction). The choice of 0.5 and 0.8 is based on Cohen,(70) who stated that an effect size of 0.5 was “moderate” and 0.8 was “large”; thus the decision rule required that the effect be statistically significantly larger than “moderate”. The use of 0.5 and 0.8 applies to both Hedges’ d and Cohen’s h as measures of effect size. When using the log odds ratio, small, moderate, and large effects are 0.4, 0.9, and 1.5, respectively. For length of hospital stay, we judged small, moderate, and large effects to be 1 day, 2 days, and 3 days, respectively.

Figure 3. General Section of Strength-of-Evidence Algorithm

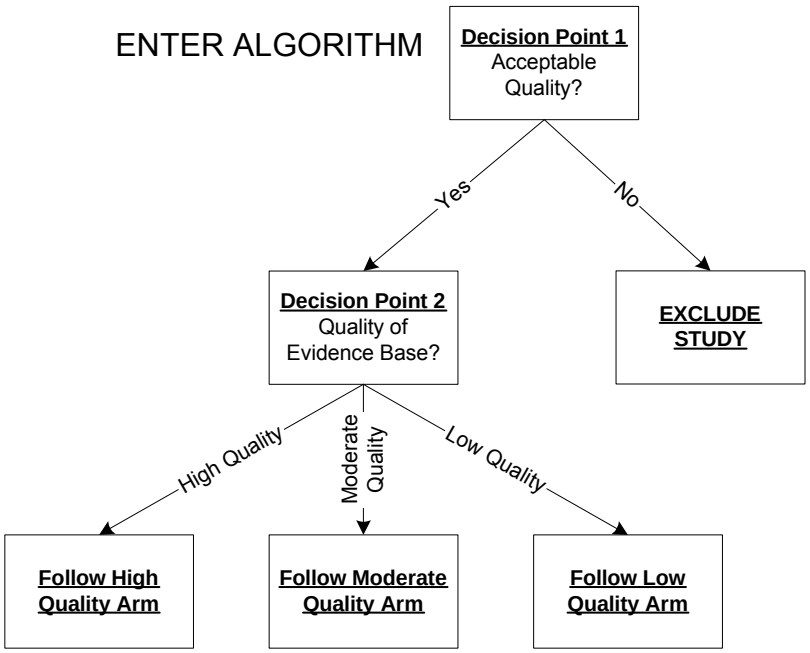


Figure 4. High Quality Arm of Strength-of-Evidence Algorithm

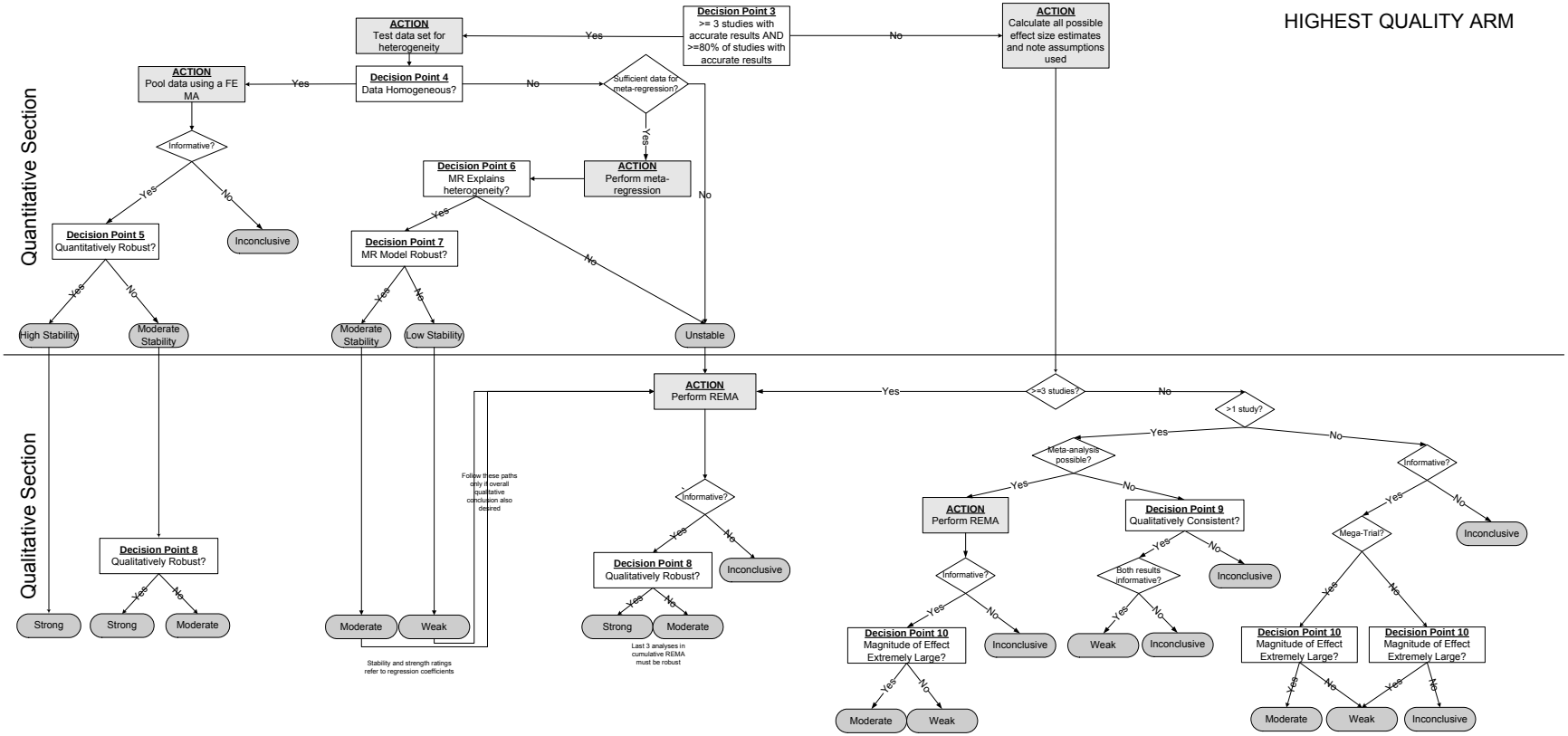


Figure 5. Moderate Quality Arm of Strength-of-Evidence Algorithm

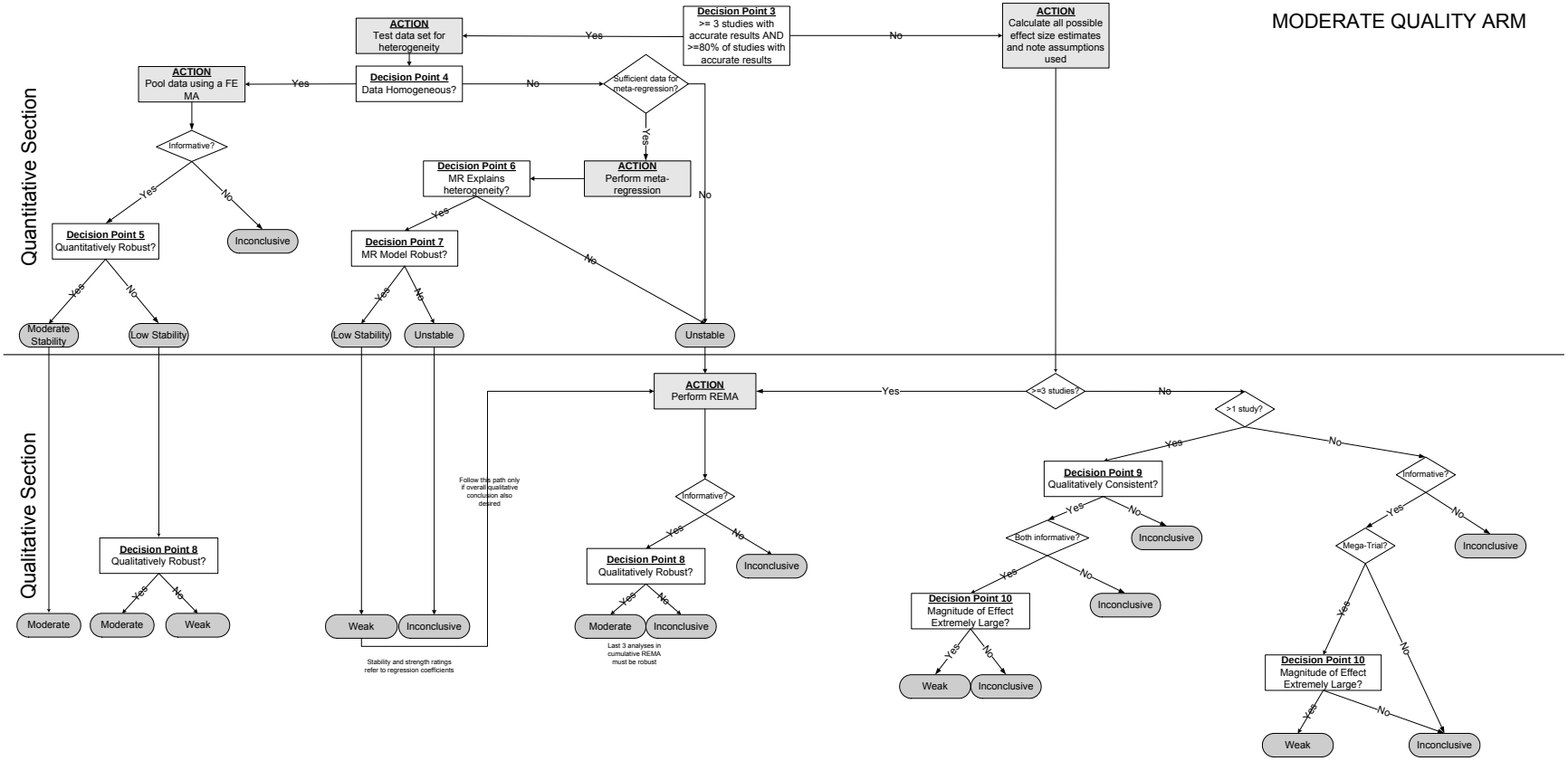
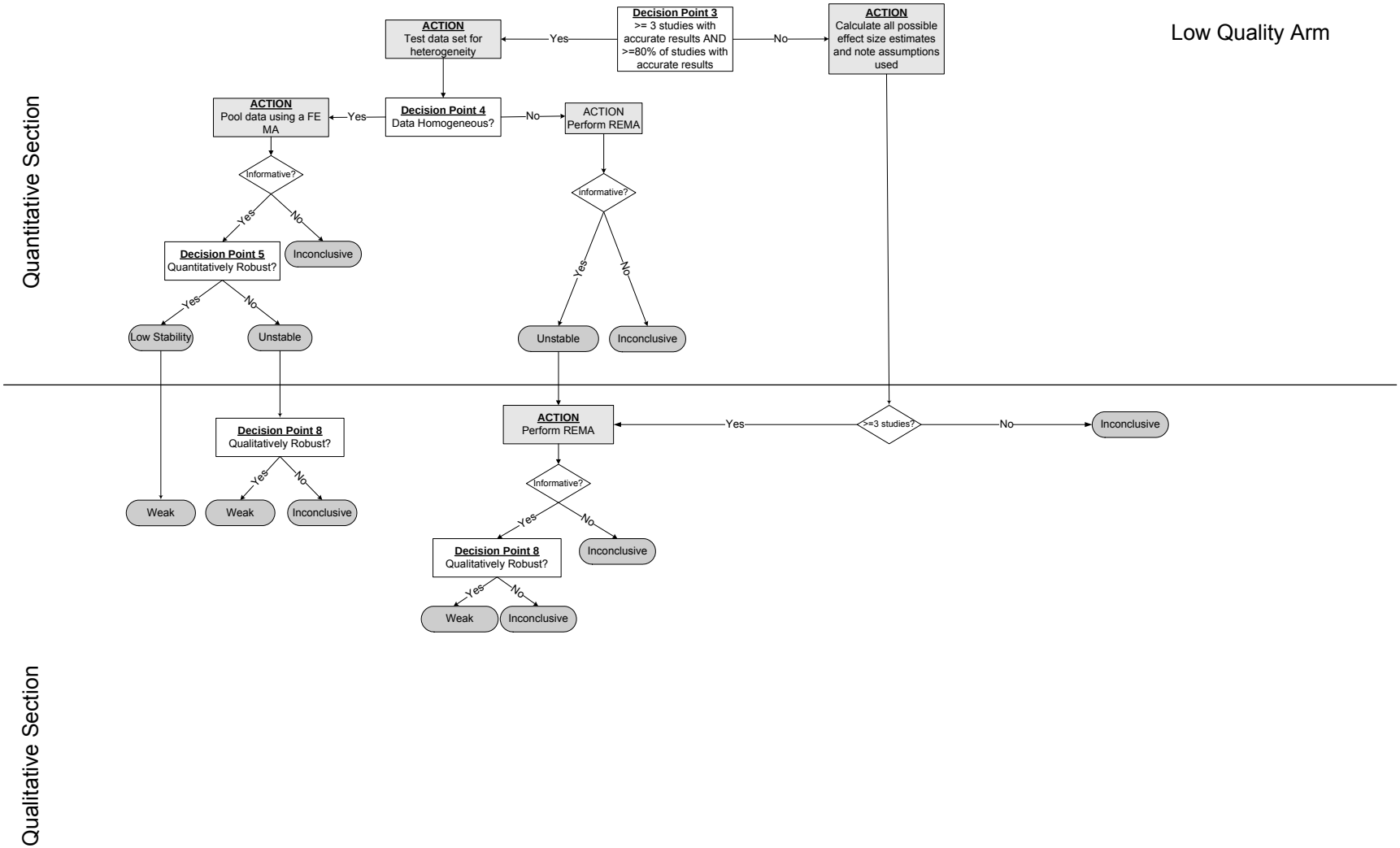


Figure 6. Low Quality Arm of Strength-of-Evidence Algorithm



Appendix C. Summary Evidence Tables

Table C-1. Summary of Included Studies (Key Questions 2 and 3)

Author/Year	Study Design/Purpose	Intervention/Outcomes	Demographics	Results
Noblett et al. 2006(42)	Design: RCT Purpose: to assess the effect of optimizing hemodynamic status, using a protocol-driven intraoperative fluid regimen, on outcome following elective colorectal resection. ECRI Quality Score (Rating): 9.1 (High)	Treatment Intervention: EDM (CardioQ) + CVP + conventional protocol Control Intervention: CVP + conventional protocol	Total Enrolled: 108 54 in treatment group 54 in control group Age (median ± IQR) Treatment: 62.3 ± 14.0 Control: 67.6 ± 15.2 % female: Treatment: NR Control: NR Type of surgery: Elective bowel surgery Inclusion Criteria: Patients requiring elective bowel surgery Exclusion Criteria: Severe esophageal disease, recent esophageal or upper airway surgery, systemic steroid medication, moderate or severe aortic valve disease, bleeding diathesis and patient choice.	Mortality: Treatment: 0% (0/51) Control: 1.9% (1/52) p-value: 1.0 Major complications: Treatment: 0% (0/30) Control: 11.5% (6/30) p-value: 0.01 Total complications: Treatment: 25.5% (13/51) Control: 42.3% (22/52) p-value: 0.07[a] Length of hospital stay (median days) Treatment: 7 (IQR: 3 to 35) Control: 9 (IQR: 4 to 45) p-value: 0.005

Author/Year	Study Design/Purpose	Intervention/Outcomes	Demographics	Results
Wakeling et al. 2005(43)	Design: RCT Purpose: to determine whether using intraoperative esophageal Doppler guided fluid management to minimize hypovolemia would reduce post-operative hospital stay and the time before return to gut function after colorectal surgery. ECRI Quality Score (Rating): 9.0 (High)	Treatment Intervention: EDM (CardioQ) + CVP + conventional protocol Control Intervention: CVP + conventional protocol	Total Enrolled: 128 64 in treatment group 64 in control group Age (median ± IQR) Treatment: 69.1 ± 12.3 Control: 69.6 ± 10.2 % female: Treatment: 40.6 Control: 46.9 Type of surgery: Elective or semi-elective bowel surgery Inclusion Criteria: Patients requiring elective or semi-elective bowel surgery Exclusion Criteria: Patients with age <18 years, hepatic pathology, perforated viscus, esophageal pathology, and coagulopathy	Mortality: Treatment: 0% (0/64) Control: 1.6% (1/64) p-value: 1.0 Major complications: Not reported separate from total complications Total complications: Treatment: 37.5% (24/64) Control: 59.4% (38/64) p-value: 0.01 Length of hospital stay (median days) Treatment: 10 (IQR 5.75) Control: 11.5 (IQR 4.75) p-value: 0.03

Author/Year	Study Design/Purpose	Intervention/Outcomes	Demographics	Results
Conway et al. 2002(44)	Design: RCT Purpose: to examine the effect of esophageal Doppler guided fluid administration during colorectal resection on hemodynamic performance, hospital stay, and post-operative complications ECRI Quality Score (Rating): 8.5 (High)	Treatment Intervention: EDM (TECO 2) + CVP + conventional protocol Control Intervention: CVP + conventional protocol	Total Enrolled: 57 29 in treatment group 28 in control group Age (mean ± SD) Treatment: 66.5 ± 12.5 Control: 67.5 ± 10.1 % female: Treatment: NR Control: NR Type of surgery: Major non-emergent bowel surgery Inclusion Criteria: Patients undergoing major bowel resections Exclusion Criteria: Patients undergoing emergency, intrathoracic, or esophageal surgery, patients with known sensitivity to starch-based colloid or history of esophageal disease	Mortality Treatment: 0% (0/29) Control: 3.6% (1/28) p-value: 0.49 Major complications: Treatment: 0% (0/30) Control: 17.9% (5/30) p-value: 0.02 Total complications Treatment: 17.2% (5/29) Control: 32.1% (9/28) p-value: 0.23[a] Length of hospital stay: Mean ±SD in days Treatment: 18.7 ±20.2 Control: 12.7 ±6.0 Median (range) Treatment: 12 (7-103) Control: 11 (7-30) p-value: NR

Author/Year	Study Design/Purpose	Intervention/Outcomes	Demographics	Results
Gan et al. 2002(11)	Design: RCT Purpose: to investigate whether goal-directed intraoperative plasma volume expansion guided by the EDM would shorten the length of hospital stay and improve post-operative outcomes in patients undergoing moderate risk surgery ECRI Quality Score (Rating): 8.1 (Moderate)	Treatment Intervention: EDM (EDM™) + CVP + conventional protocol Control Intervention: CVP + conventional protocol	Total Enrolled: 100 50 in treatment group 50 in control group Age (mean ± SD) Treatment: 56 ± 13 Control: 59 ± 12 % female: Treatment: 38 Control: 48 Type of surgery: Major elective general, urologic, or gynecologic surgery Inclusion Criteria: Patients with American Society of Anesthesiologists (ASA) physical status I, II, and III who were to undergo major elective general, urologic, or gynecologic surgery with an anticipated blood loss of >500 ml. Exclusion Criteria: Patients with age <18 years, emergency surgery, preoperative bowel obstruction, coagulopathy, significant renal and hepatic dysfunction (creatinine >50% or liver enzymes >50% upper limit of normal values), congestive heart failure, and esophageal pathology (avoid potential complications of the esophageal probe), and those undergoing gastric or esophageal surgery or who were on antiemetic medication within 3 days of surgery	Mortality: Treatment: 0% (0/50) Control: 0% (0/50) p-value: 1.0 Major complications: Not reported separate from total complications Total complications: Treatment: 42% (21/50) Control: 76% (38/50) p-value: 0.001[a] Length of hospital stay: Mean ± SD in days Treatment: 5 ± 3 Control: 7 ± 3 Median Treatment: 6 Control: 7 p-value: 0.03

Author/Year	Study Design/Purpose	Intervention/Outcomes	Demographics	Results
Venn et al. 2002(46)	Design: RCT Purpose: to investigate whether repeated colloid fluid challenges to optimize the circulation intraoperatively, guided by CVP or esophageal Doppler ultrasonography, would benefit high-risk patients admitted with fractured hips to a London teaching hospital ECRI Quality Score (Rating): 9.0 (High)	Treatment Intervention: EDM (CardioQ) + conventional protocol Control Intervention: CVP + conventional protocol Other Intervention: Conventional protocol alone	Total Enrolled: 90 30 in treatment group 31 in control group 29 in other group Age (mean ± SD) Treatment: 82 ± 8.7 Control: 85 ± 6.2 Other: 84.5 ± 9.3 % female: Treatment: 80 Control: 87.1 Other: 79.3 Type of surgery: Proximal femoral fracture repair Inclusion Criteria: Patients admitted with fractured hips Exclusion Criteria: Patients with age <65 years, esophageal pathology, central venous cannula already *in situ*, pathological fracture of femur, refusal of informed consent, and those undergoing regional anesthesia	Mortality: Treatment: 10% (3/30) Control: 19.4% (6/31) Other: 6.9% (2/29) p-value: 0.31 Major complications: Not reported separate from total complications Total complications: Treatment: 46.7% (14/30) Control: 51.6% (16/31) Other: 79.3% (23/29) p-value: 0.24 Length of hospital stay: Mean days (95% CI) Treatment: 13.5 (10.9 to 17.5) Control: 13.3 (10.3 to 19.2) Other: 17.5 (13.9 to 24.4) p-value: 0.035

Author/Year	Study Design/Purpose	Intervention/Outcomes	Demographics	Results
Sinclair et al. 1997(47)	Design: RCT Purpose: to examine the possible benefits of intraoperative circulatory optimization using EDM in patients with fractured neck of femur ECRI Quality Score (Rating): 8.9 (High)	Treatment Intervention: EDM (ODM 2) + conventional protocol Control Intervention: Conventional protocol	Total Enrolled: 40 20 in treatment group 20 in control group Age (Median and IQR) Treatment: 74 (70.5-82) Control: 75.5 (69-80) % female: Treatment: NR Control: NR Type of surgery: Proximal femoral fracture repair Inclusion Criteria: Patients with fractures of the femoral neck of the hip Exclusion Criteria: Patients with age <55 years, fracture secondary to neoplasm, fractures occurring during hospitalization for an acute illness, fracture through the site of a previous surgical correction or associated with instabiliy of a previous prosthesis, planned regional anesthesia (precluding placement of Doppler probe), and refusal of consent or inability to contact next of kin in the case of patietns unable to give consent themselves	Mortality: Treatment: 5% (1/20) Control: 10% (2/20) p-value: 1.0 Major and Total complications: Treatment: NR Control: NR p-value: NR Length of hospital stay: Median and range in days: Treatment: 11 (3 to 23) Control: 20 (5 to 220) p-value: <0.05

Author/Year	Study Design/Purpose	Intervention/Outcomes	Demographics	Results
Mythen and Webb 1995(45)	Design: RCT Purpose: to test the hypothesis that perioperative plasma volume expansion with colloid (guided by EDM + CVP) would maintain pHi during elective cardiac surgery ECRI Quality Score (Rating): 8.9 (High)	Treatment Intervention: EDM (ODM 1) + CVP + conventional protocol Control Intervention: CVP + conventional protocol	Total Enrolled: 60 30 in treatment group 30 in control group Age (mean and range) Treatment: 63 (42-89) Control: 64 (44-86) % female: Treatment: NR Control: NR Type of surgery: Elective cardiac surgery (CABG or single valve replacement) Inclusion Criteria: Patients scheduled for elective coronary artery bypass grafts (CABG) or single heart valve replacement) who had a preoperative left ventricular ejection fraction (LVEF) estimated to be ≥50% and graded by the anesthesiologist in charge as ASA grade III. Exclusion Criteria: Patients with age <18 years, pregnancy, coagulopathies or perforated viscus, esophageal or gastric abnormalities, nonpulsatile cardiopulmonary bypass, administration of aprotinin, and prior heart surgery or preexisting respiratory, hepatic, or renal disease.	Mortality: Treatment: 0% (0/30) Control: 3.3% (1/30) p-value: 1.0 Major complications: Treatment: 0% (0/30) Control: 20% (6/30) p-value: 0.01 Total complications: Treatment: NR Control: NR p-value: NR Length of hospital stay: Mean and range in days Treatment: 6.4 (5 to 9) Control: 10.1 (5 to 48) p-value: 0.01

Author/Year	Study Design/Purpose	Intervention/Outcomes	Demographics	Results
McKendry et al. 2004(48)	Design: RCT Purpose: to assess whether a nurse led, flow monitored protocol for optimizing circulatory status in patients after cardiac surgery reduces complications and shortens stay in intensive care and hospital ECRI Quality Score (Rating): 8.5 (High)	Treatment Intervention: EDM (CardioQ) + conventional protocol Control Intervention: Conventional protocol	Total Enrolled: 174 89 in treatment group 85 in control group Age (mean ± SD) Treatment: 65.6 ± 10.3 Control: 66.7 ± 10.2 % female: Treatment: 37 Control: 33 Type of surgery: Cardiopulmonary bypass surgery Inclusion Criteria: Patients undergoing cardiopulmonary bypass surgery who had provided informed consent Exclusion Criteria: Patients with age <18 years, those undergoing off-pump surgery, those who declined consent, or those with relative contraindications to use of the esophageal Doppler probe (esophageal varices or other oropharyngeal and esophageal disease).	Mortality: Treatment: 4.5% (4/89) Control: 2.4% (2/85) p-value: 0.68 Major complications: Not reported separate from total complications Total complications: Treatment: 19.1% (17/89) Control: 30.6% (26/85) p-value: 0.08[a] Length of hospital stay: Mean days: Treatment: 11.4 Control: 13.9 p-value: NR Median days: Treatment: 7 Control: 9 p-value: 0.02

[a] Calculated by ECRI

Appendix D. Evidence Tables for Key Question 2

Table D-1. Patient Enrollment Criteria for Studies Addressing Key Question 2

Reference	Year	Inclusion Criteria	Exclusion Criteria
Noblett et al.(42)	2006	Patients requiring elective bowel surgery	Severe esophageal disease, recent esophageal or upper airway surgery, systemic steroid medication, moderate or severe aortic valve disease, bleeding diathesis and patient choice.
Wakeling et al.(43)	2005	Patients requiring elective or semi-elective bowel surgery	Patients with age <18 years, hepatic pathology, perforated viscus, esophageal pathology, and coagulopathy
Conway et al.(44)	2002	Patients undergoing major bowel resections	Patients undergoing emergency, intrathoracic, or esophageal surgery, patients with known sensitivity to starch-based colloid or history of esophageal disease
Gan et al.(11)	2002	Patients with American Society of Anesthesiologists (ASA) physical status I, II, and III who were to undergo major elective general, urologic, or gynecologic surgery with an anticipated blood loss of >500 ml.	Patients with age <18 years, emergency surgery, preoperative bowel obstruction, coagulopathy, significant renal and hepatic dysfunction (creatinine >50% or liver enzymes >50% upper limit of normal values), congestive heart failure, and esophageal pathology (avoid potential complications of the esophageal probe), and those undergoing gastric or esophageal surgery or who were on antiemetic medication within 3 days of surgery
Venn et al.(46)	2002	Patients admitted with fractured hips	Patients with age <65 years, esophageal pathology, central venous cannula already *in situ*, pathological fracture of femur, refusal of informed consent, and those undergoing regional anesthesia
Sinclair et al.(47)	1997	Patients with fractures of the femoral neck of the hip	Patients with age <55 years, fracture secondary to neoplasm, fractures occurring during hospitalization for an acute illness, fracture through the site of a previous surgical correction or associated with instabiliy of a previous prosthesis, planned regional anesthesia (precluding placement of Doppler probe), and refusal of consent or inability to contact next of kin in the case of patietns unable to give consent themselves

Reference	Year	Inclusion Criteria	Exclusion Criteria
Mythen and Webb(45)	1995	Patients scheduled for elective coronary artery bypass grafts (CABG) or single heart valve replacement who had a preoperative left ventricular ejection fraction (LVEF) estimated to be ≥50% and graded by the anesthesiologist in charge as ASA grade III.	Patients with age <18 years, pregnancy, coagulopathies or perforated viscus, esophageal or gastric abnormalities, nonpulsatile cardiopulmonary bypass, administration of aprotinin, and prior heart surgery or preexisting respiratory, hepatic, or renal disease.

NR Not Reported

Table D-2. Characteristics of Patients Receiving Cardiac Output Monitoring

Author/ year	Year	Monitoring protocol	N	Age (mean ±SD)	% female	ASA grade	Goldman cardiac risk index	POSSUM score	Hemoglobin level at baseline (g/dl)	Type of surgery
Noblett et al.(42)	2006	EDM (CardioQ) + CVP + conventional protocol	54	62.3 ± 14.0	NR	2.1 ± 0.6	NR	15.4 ± 4.2	NR	Elective bowel surgery (colorectal resection)
		CVP + conventional protocol	54	67.6 ± 15.2	NR	2.2 ± 0.6	NR	16.1 ± 3.7	NR	
Wakeling et al.(43)	2005	EDM (CardioQ) + CVP + conventional protocol	64	69.1 ± 12.3	40.6	2 ± 1	NR	17 ± 6.5	NR	Elective or semi-elective bowel surgery
		CVP + conventional protocol	64	69.6 ± 10.2 Median ± IQR	46.9	2 ± 1 Median ± IQR	NR	18 ± 7 Median ± IQR	NR	

Author/ year	Year	Monitoring protocol	N	Age (mean ±SD)	% female	ASA grade	Goldman cardiac risk index	POSSUM score	Hemoglobin level at baseline (g/dl)	Type of surgery
Conway et al.(44)	2002	EDM (Teco 2) + CVP + conventional protocol CVP + conventional protocol	29 28	66.5 (12.5) 67.5 (10.1)	NR NR	I (I-III) II (I-III) Median (range)	3 (3-11) 3 (3-29) Median (range)	NR NR	12.8 (1.8) 12.8 (1.5)	Major non-emergent bowel surgery
Gan et al.(11)	2002	EDM (EDM™) + CVP + conventional protocol CVP + conventional protocol	50 50	56 (13) 59 (12)	38 48	II (I-III) II (I-III) Median (range)	NR NR	NR NR	13.4 (1.9) 12.9 (1.7)	Major elective general, urologic, or gynecologic surgery
Venn et al.(46)	2002	EDM (CardioQ) + conventional protocol CVP + conventional protocol Conventional protocol	30 31 29	82 (8.7) 85 (6.2) 84.5 (9.3)	80 87.1 79.3	3 (2.5-3) 3 (3-4) 3 (3-4) Median (IQR)	NR NR NR	35 (32-40) 40 (35-42) 38 (34-40) Median (IQR)	NR NR NR	Proximal femoral fracture repair
Sinclair et al.(47)	1997	EDM (ODM 2) + conventional protocol Conventional protocol	20 20	74 (70.5-82) 75.5 (69-80) Median (IQR)	NR NR	2 (2-3) 2 (2-3) Median (IQR)	9 (9-13) 9 (8-12) Median (IQR)	NR NR	12.5 (11.1-13.7) 12.7 (10.7-14) Median (IQR)	Proximal femoral fracture repair

Author/ year	Year	Monitoring protocol	N	Age (mean ±SD)	% female	ASA grade	Goldman cardiac risk index	POSSUM score	Hemoglobin level at baseline (g/dl)	Type of surgery
Mythen and Webb(45)	1995	EDM (ODM 1) + CVP + conventional protocol CVP + conventional protocol	30 30	63 (42-89) 64 (44-86) Mean (range)	NR NR	III III (All patients)	NR NR	NR NR	NR NR	Elective cardiac surgery (CABG or single valve replacement)

CVP Central Venous Pressure Monitoring
EDM Esophageal Doppler Monitoring
IQR Interquartile Range
NR Not Reported

Table D-3. Fluid Management Protocols in Studies Addressing Key Question 2

Reference	Year	Type of Surgery	Fluid management protocol for EDM group	Fluid management protocol for control group
Noblett et al.(42)	2006	Elective bowel surgery	In addition to routine fluid management, patients received colloid fluid boluses to maintain a descending aortic corrected flow time of > 0.35 seconds, and further boluses were given to optimize the stroke volume. Once achieved, further fluid boluses were given only if the stroke volume altered by >10% or the corrected flow time fell below 0.35 seconds.	Patients managed using routine fluid management at the discretion of the anesthiologist. Only some patients had CVP lines. All patients had EDM probes inserted, but they were not used to guide fluid administration according to a specific protocol. The anesthetist and those in charge of postoperative patient care were blinded to the EDM readings.
Wakeling et al.(43)	2005	Elective or semi-elective bowel surgery	In addition to routine fluid management, patients received 250 ml boluses of colloid solution. If stroke volume increased by 10% or more but CVP did not rise by 3 mm Hg or more, the fluid challenge was repeated. Fluid challenges were repeated until the stroke volume	Patients managed using routine cardiovascular monitoring and CVP measurements. CVP was used to guide i.v. fluid administration and was kept between 12 and 15 mm Hg. Anesthetist was blinded to esophageal Doppler measurements, which were taken before, after laparotomy,

Reference	Year	Type of Surgery	Fluid management protocol for EDM group	Fluid management protocol for control group
			failed to rise by 10% and/or the CVP rose by 3 mm Hg or more. No further colloid fluid boluses were given until a 10% decrease in stroke volume occurred. Esophageal Doppler measurements were performed continuously.	and at the end of the operation.
Conway et al.(44)	2002	Major non-emergent bowel surgery	In addition to routine fluid management, patients received additional colloid fluid boluses of 3 ml/kg according to an algorithm based on esophageal Doppler measurements. The algorithm was designed to optimize stroke volume (until SV did not increase by 10%) and maintain the corrected flow time >0.35 seconds. The anesthetist was blinded to EDM readings (but if the stroke volume fell in response to a fluid challenge and flow time and aortic velocity waveform indicated that the patient was volume overloaded, the anesthetist was unblinded).	Routine fluid management based on monitoring of heart rate, blood pressure, and (at the discretion of the anesthesiologist) CVP. EDM probe was inserted in all patients, and the anesthetist was blinded to EDM readings (but if the stroke volume fell in response to a fluid challenge and flow time and aortic velocity waveform indicated that the patient was volume overloaded, the anesthetist was unblinded).
Gan et al.(11)	2002	Major elective general, urologic, or gynecologic surgery	Patients received colloid fluid boluses of 200 ml according to an algorithm based on esophageal Doppler measurements. The algorithm was designed to optimize stroke volume (until SV did not increase by 10%) and maintain the corrected flow time >0.35 seconds. If the latter increased above 0.4 seconds with no change in stroke volume, further fluid was not administered until the stroke volume decreased by 10% of the last value. In addition, patients received fluid equivalent to that lost as a result of surgical hemorrhage. When 20 ml/kg of 6% hydroxyethy starch in saline had been given, lactated Ringer's solution was used for further fluid boluses as required. Crystalloid was used in a 3:1 ratio for replacement of surgical blood loss.	The esophageal Doppler monitor was turned away from the anesthetist and the screen covered. Hemodynamic variables triggering fluid administration involved a urinary output <0.5 ml/kg/h, an increase in heart rate >20% above baseline or >110 beats/min, a decrease in mean systolic blood pressure <20% below baseline or <90 mm Hg, or CVP <20% of baseline. Boluses of 200 ml of fluid were administered until the above target was restored. The anesthesiologist would also administer additional fluid if deemed clinically indicated.

Reference	Year	Type of Surgery	Fluid management protocol for EDM group	Fluid management protocol for control group
Venn et al.(46)	2002	Proximal femoral fracture repair	In addition to routine fluid management, patients received additional 200 ml colloid fluid challenges guided by Doppler measurements of stroke volume and corrected flow time. Similar to Conway et al. protocol, except if the corrected flow time rose above 0.4 seconds and the stroke volume remained the same, a further 100 ml colloid fluid was given. If after this fluid challenge the stroke volume still remained the same, no further fluid was given until the stroke volume fell by 10%. CVP was not recorded.	Control (no CVP) group: I.v. fluid administered as appropriate according to physician judgment (crystalloid and/or colloid). Although CVP was recorded, the clinician was blinded to measurements and unable to use them to guide therapy. No additional fluid boluses were given in this group. CVP group: in addition to routine fluid management, patients received additional 200 ml colloid fluid challenges guided by the response of CVP to a fluid challenge.
Sinclair et al.(47)	1997	Proximal femoral fracture repair	In addition to routine fluid management, patients received additional colloid fluid boluses of 3 ml/kg according to an algorithm based on esophageal Doppler measurements. The algorithm was designed to optimize stroke volume (until SV did not increase by 10%) and maintain the corrected flow time >0.35 seconds. If the latter increased above 0.4 seconds with no change in stroke volume, further fluid was not administered until the stroke volume decreased by 10% of the last value. CVP was not recorded. The anesthetist was blind to EDM readings, but aware of the fluid volumes given as fluid challenges to the protocol group.	Patients received crystalloid, colloid, or blood to replace estimated fluid losses and to maintain heart rate and blood pressure. CVP was not recorded. EDM probes were inserted in all patients, but the anesthetist was blinded to EDM readings.
Mythen and Webb(45)	1995	Elective cardiac surgery (CABG or single valve replacement)	In addition to routine fluid management, patients received additional 200 ml colloid fluid boluses to obtain a maximum stroke volume (when stroke volume failed to rise by 10%) and a rise in CVP >3 mm Hg. Stroke volume was monitored until a 10% decrease occurred, at which time another bolus was given. This procedure was repeated every 15 minutes until the end of surgery except when the patients underwent cardiopulmonary bypass.	Patients received crystalloid or colloid solutions based on judgment of the anesthetist. EDM probe was inserted in all patients, but operating room personnel were blinded to EDM readings.

CABG Coronary Artery Bypass Grafting
CVP Central Venous Pressure
EDM Esophageal Doppler Monitoring

Table D-4. Study Quality Evaluation – Studies Comparing Esophageal Doppler Monitoring Plus CVP Plus Conventional Protocol to CVP Plus Conventional Protocol

ECRI study quality scale - questions	Study				
	Noblett et al.	Wakefield et al.	Conway et al.	Gan et al.	Mythen and Webb
1. Were patients randomly assigned to groups?	Yes	Yes	Yes	Yes	Yes
2. Did the study employ stochastic randomization?	Yes	NR	Yes	Yes	No
3. Were any methods used to make the groups comparable- randomization, matching, etc.?	Yes	Yes	Yes	Yes	Yes
4. Were patients assigned to groups based on factors other than patient or physician preference?	Yes	Yes	Yes	Yes	Yes
5. Were the characteristics of the patients in different groups comparable?	Yes	Yes	No	Yes	Yes
6. Did the patients in the different study groups have similar levels of performance on outcomes at baseline?	Yes	Yes	Yes	Yes	Yes
7. Was the study prospectively planned?	Yes	Yes	Yes	Yes	Yes
8. Did 85% or more of the patients complete the study?	Yes	Yes	Yes	Yes	Yes
9. Was there a less than 16% difference in completion rates in the study's groups?	Yes	Yes	Yes	Yes	Yes
10. Were all of the study's groups concurrently treated?	Yes	Yes	Yes	Yes	Yes
11. Was compliance with treatment greater than or equal to 85% in both of the groups?	Yes	Yes	Yes	Yes	Yes
12. Were both groups treated at the same centers?	Yes	Yes	Yes	Yes	Yes
13. Were subjects blinded to treatment?	Yes	Yes	Yes	Yes	Yes
14. Did the authors test and confirm that blinding of patients was maintained?	NR	NR	NR	NR	NR
15. Was the treating physician blinded to group assignment?	Yes	No	No	No	No
16. Were the outcome assessors blinded to group assignment?	Yes	Yes	No	Yes	Yes
17. Was there concealment of allocation?	Yes	Yes	Yes	Yes	Yes
18. Was the outcome of interest objective and was it objectively measured?[a]	Yes	Yes	Yes	Yes	Yes
19. Were the same methods used to measure outcomes in all of the study's groups?[a]	Yes	Yes	Yes	Yes	Yes

ECRI study quality scale - questions	Study				
	Noblett et al.	Wakefield et al.	Conway et al.	Gan et al.	Mythen and Webb
20. Was the instrument used to measure the outcome standard?[a]	Yes	Yes	Yes	Yes	Yes
21. Was the same treatment given to all of the patients enrolled in the experimental group?	Yes	Yes	Yes	No	Yes
22. Was the same treatment given to all of the patients enrolled in the control group?	Yes	Yes	Yes	No	Yes
23. Were the follow-up times in all of the study's relevant groups approximately equal?	Yes	Yes	Yes	Yes	Yes
24. Was the funding for this study derived from a source that does not have a financial interest in its results?	Yes	Yes	Yes	No	Yes
25. Were the author's conclusions supported by the data in the results section?	Yes	Yes	Yes	Yes	Yes
Quality score	9.7	9.0	8.5	8.1	8.9
Quality rating	High	High	High	Moderate	High

Table D-5. Study Quality Evaluation – Studies Comparing Esophageal Doppler Monitoring Plus Conventional Protocol to Conventional Protocol

ECRI study quality scale - questions	Study	
	Venn et al.	Sinclair et al.
1. Were patients randomly assigned to groups?	Yes	Yes
2. Did the study emply stochastic randomization?	Yes	No
3. Were any methods used to make the groups comparable- randomization, matching, etc.?	Yes	Yes
4. Were patients assigned to groups based on factors other than patient or physician preference?	Yes	Yes
5. Were the characteristics of the patients in different groups comparable?	NR	Yes
6. Did the patients in the different study groups have similar levels of performance on outcomes at baseline?	Yes	Yes
7. Was the study prospectively planned?	Yes	Yes
8. Did 85% or more of the patients complete the study?	Yes	Yes
9. Was there a less than 16% difference in completion rates in the study's groups?	Yes	Yes
10. Were all of the study's groups concurrently treated?	Yes	Yes
11. Was compliance with treatment greater than or equal to 85% in both of the groups?	Yes	Yes
12. Were both groups treated at the same centers?	Yes	Yes
13. Were subjects blinded to treatment?	Yes	Yes
14. Did the authors test and confirm that blinding of patients was maintained?	NR	NR
15. Was the treating physician blinded to group assignment?	No	No
16. Were the outcome assessors blinded to group assignment?	Yes	Yes
17. Was there concealment of allocation?	Yes	Yes
18. Was the outcome of interest objective and was it objectively measured?[a]	Yes	Yes
19. Were the same methods used to measure outcomes in all of the study's groups?[a]	Yes	Yes
20. Was the instrument used to measure the outcome standard?[a]	Yes	Yes

ECRI study quality scale - questions	Study	
	Venn et al.	Sinclair et al.
21. Was the same treatment given to all of the patients enrolled in the experimental group?	Yes	Yes
22. Was the same treatment given to all of the patients enrolled in the control group?	Yes	Yes
23. Were the follow-up times in all of the study's relevant groups approximately equal?	Yes	Yes
24. Was the funding for this study derived from a source that does not have a financial interest in its results?	Yes	Yes
25. Were the author's conclusions supported by the data in the results section?	Yes	Yes
Quality score	9.0	8.9
Quality rating	High	High

Table D-6. Specific Complications Reported in Included Studies

Reference	Year	Type of Surgery	Reported complications
Trials comparing EDM + CVP + conventional protocol to CVP + conventional protocol			
Noblett et al.(42)	2006	Elective bowel surgery	Major complications: Life-threatening complications requiring HDU or ICU care. Included death, pneumonia plus multiple organ dysfunction syndrome (MODS), anastomotic breakdown and MODS, sepsis and intra-abdominal collection, pneumonia requiring non-invasive ventilatory support. Intermediate complications: Complications requiring surgical, endoscopic, or radiological intervention. (Note: The authors reported the total number of intermediate or major complications for each treatment group.) Other complications: complications requiring pharmacological treatment, or deviations from normal postoperative course not requiring intervention. Specific complications mentioned were ileus, nausea, and vomiting.
Wakeling et al.(43)	2005	Elective or semi-elective bowel surgery	Death, pulmonary/thrombotic, infectious, renal, gastrointestinal, cardiovascular, neurological, wound, hemotological, pain, social. They did not separate major complications from total complications.
Conway et al.(44)	2002	Major bowel resections	Death, chest infection, pulmonary embolus, cardiac failure, arrhythmias, surgical problems requiring reoperation, delirium Major complications: complications requiring critical care. (Note: the authors did not report how many patients in each group had each type of complication, they only reported the total number of patients with complications in each group.)

Reference	Year	Type of Surgery	Reported complications
Gan et al.(11)	2002	Major elective general, urologic, or gynecologic surgery with an anticipated blood loss of >500 ml.	Acute renal dysfunction, cardiovascular (hypotension, pulmonary edema, arrhythmia), chest infection, wound infection, coagulopathy, severe nausea/vomiting requiring emetic, respiratory support > 24 h. (Note: authors reported each of these complications separately for each treatment group. They did not report the total number of patients who had complications, they reported only the total number of events.)
Mythen and Webb(45)	1995	Elective coronary artery bypass grafts (CABG) or single heart valve replacement	Major complications: Death, multiple organ failure, chest infection with pleural effusion and disorientation, respiratory failure with nausea and vomiting, cerebrovascular accident, paralytic ileus with pericardial effusion and disorientation. Minor complications: wound infection, dyspnea, disorientation, persistent nausea and vomiting beyond third post-operative day. (Note: the authors only reported the number of patients who had major complications. Because they did not even report the number of minor complications, no estimate of total complications was possible).
Trials comparing EDM + conventional protocol to conventional protocol			
Venn et al.(46)	2002	Hip fracture repair	Mortality, myocardial infarction, cardiac failure, rapid atrial fibrillation, hypotension, impaired renal function, pseudo-obstruction, chest infection, wound infection, urinary tract infection, cellulitis ,pancreatitis, pulmonary embolus, deep hemorrhage requiring transfusion, hematemesis. (Note: authors reported each of these complications separately for each treatment group. They did not report the total number of patients who had complications, they reported only the total number of events.)
Sinclair et al.(47)	1997	Hip fracture repair	NR

HDU High dependency unit
ICU Intensive care unit
NR Not reported

Table D-7. Results for Key Question 2 - Mortality

Study	N	% operative deaths (n/N)			% post-operative deaths (in-hospital or within 30 days) (n/N)			% total deaths during trial (n/N)		
		EDM group	Control group	p-value	EDM group	Control group	p-value	EDM group	Control group	p-value
Trials comparing EDM + CVP + conventional protocol to CVP + conventional protocol										
Noblett et al.(42)	103	0 (0/51)	0 (0/52)	1.0	0 (0/51)	1.9 (1/52)	1.0	0 (0/51)	1.9 (1/52)	1.0
Wakeling et al.(43)	128	0 (0/64)	0 (0/64)	1.0	0 (0/64)	0 (0/64)	1.0	0 (0/64)	1.6 (1/64)	1.0
Conway et al.(44)	57	0 (0/29)	0 (0/28)	1.0	0 (0/29)	3.6 (1/28)	0.49	0 (0/29)	3.6 (1/28)	0.49
Gan et al.(11)	100	0 (0/50)	0 (0/50)	1.0	0 (0/50)	0 (0/50)	1.0	0 (0/50)	0 (0/50)	1.0
Mythen and Webb(45)	60	0 (0/30)	0 (0/30)	1.0	0 (0/30)	3.3 (1/30)	1.0	0 (0/30)	3.3 (1/30)	1.0
Trials comparing EDM + conventional protocol to CVP + conventional protocol										
Venn et al.(46)	90	0 (0/30)	0 (0/31)	1.0	10 (3/30)	19.4 (6/31)	0.30	10 (3/30)	19.4 (6/31)	0.30
Trials comparing EDM + conventional protocol to conventional protocol										
Venn et al.(46)	90	0 (0/30)	0 (0/29)	1.0	10 (3/30)	6.9 (2/29)	1.0	10 (3/30)	6.9 (2/29)	1.0
Sinclair et al.(47)	40	0 (0/20)	0 (0/20)	1.0	0 (0/20)	5.0 (1/20)	1.0	5.0 (1/20)	10 (2/20)	1.0

Table D-8. Results for Key Question 2 – Major Complications

Study	N	% major (life-threatening) complications during trial (n/N)		Peto odds ratio (95% CI)[a]	p-value[a]
		EDM group	Control group		
Trials comparing EDM + CVP + conventional protocol to CVP + conventional protocol					
Noblett et al.(42)	103	0 (0/51)	11.5 (6/52)	0.13 (0.02 to 0.64)	0.01
Conway et al.(44)	57	0 (0/29)	17.9 (5/28)	0.11 (0.02 to 0.69)	0.02
Mythen and Webb(45)	60	0 (0/30)	20 (6/30)	0.11 (0.02 to 0.60)	0.01

[a] Calculated by ECRI

Table D-9. Results for Key Question 2 – Total Complications

Study	N	% total complications during trial (n/N)		Odds ratio (95% CI)[a]	p-value[a]
		EDM group	Control group		
Trials comparing EDM + CVP + conventional protocol to CVP + conventional protocol					
Noblett et al.(42)	103	25.5 (13/51)	42.3 (22/52)	0.47 (0.20 to 1.07)	0.07
Wakeling et al.(43)	128	37.5 (24/64)	59.4 (38/64)	0.41 (0.20 to 0.84)	0.01
Conway et al.(44)	57	17.2 (5/29)	32.1 (9/28)	0.44 (0.13 to 1.54)	0.23
Gan et al.(11)	100	42 (21/50)[b]	76 (38/50)[b]	0.23 (0.10 to 0.54)	0.001
Trials comparing EDM + conventional protocol to CVP + conventional protocol					
Venn et al.(46)	90	33.3 (10/30) 46.7 (14/30)[b]	45.2 (14/31) 51.6 (16/31)[b]	0.61 (0.21 to 1.72) 0.82 (0.30 to 2.25)	0.35 0.70
Trials comparing EDM + conventional protocol to conventional protocol					
Venn et al.(46)	90	33.3 (10/30) 46.7 (14/30)[b]	55.2 (16/29) 79.3 (23/29)[b]	0.41 (0.14 to 1.16) 0.23 (0.07 to 0.72)	0.09 0.01
Sinclair et al.(47)	40	NR	NR		

[a] Calculated by ECRI

[b] Total number of complications, not number of patients with complications

NR Not Reported

Table D-10. Results for Key Question 2 – Length of Hospital Stay

Study	N	Length of hospital stay (days)		
		EDM group	Control group	p-value
Trials comparing EDM + CVP + conventional protocol to CVP + conventional protocol				
Noblett et al.(42)	103	Median: 7 (IQR: 3 to 35)	Median: 9 (IQR: 4 to 45)	0.005
Wakeling et al.(43)	128	Median: 10 (IQR: 5.75)	Median: 11.5 (IQR: 4.75)	0.03
Conway et al.(44)	57	Mean: 18.7 (±20.2) Median: 12 (Range: 7 to 103)	Mean: 12.7 (±6.0) Median: 11 (Range: 7 to 30)	NR
Gan et al.(11)	100	Mean: 5 (±3) Median: 6	Mean: 7 (±3) Median: 7	0.03 (for medians)
Mythen and Webb(45)	60	Mean: 6.4 (Range: 5 to 9)	Mean: 10.1 (Range: 5 to 48)	0.01
Trials comparing EDM + conventional protocol to CVP + conventional protocol				
Venn et al.(46)	90	Mean: 13.5 (95% CI: 10.9 to 17.5)	Mean: 13.3 (95% CI: 10.3 to 19.2)	0.96
Trials comparing EDM + conventional protocol to conventional protocol				
Venn et al.(46)	90	Mean: 13.5 (95% CI: 10.9 to 17.5)	Mean: 17.5 (95% CI: 13.9 to 24.4)	0.31
Sinclair et al.(47)	40	Median: 11 (Range: 3 to 23)	Median: 20 (Range: 5 to 220)	<0.05

Table D-11. Meta-analysis – Length of Hospital Stay (EDM + Conventional Protocol vs. Conventional Protocol)

Study	N =	Effect Size (days)	Lower 95% CI	Upper 95% CI	p-value	I^2
Venn et al.(46)	90	-4.00	-11.57	3.57	0.30	NA
Sinclair et al.(47)	40	-9.00	-15.83	-2.17	0.01	NA
Random-effects summary effect size	130	-6.76	-11.83	-1.68	0.009	NA

Table D-12. Results of Sensitivity Analyses for Major Complications (EDM + CVP + Conventional Protocol vs. CVP + Conventional Protocol)

Sensitivity analysis	Summary Peto log odds ratio (95% CI)	p value
Qualitative robustness		
Removal of Noblett et al.(42) (largest study)	-2.19 (-3.42 to -0.96)	0.0005
Removal of Conway et al.(44) (smallest study)	-2.13 (-3.30 to -0.96)	0.0004
Removal of Mythen and Webb(45)	-2.13 (-3.35 to -0.91)	0.0006
Removal of Conway et al., Mythen and Webb (two smallest studies)	-2.08 (-3.72 to -0.44)	0.01
Removal of Noblett et al., Mythen and Webb (two largest studies)	-2.19 (-3.86 to -0.51)	0.01
Assumption of no effect in studies that did not separately report major complications (Wakeling et al., Gan et al.)	-1.11 (-2.17 to -0.05)	0.04
Original random-effects meta-analysis	-2.15 (-3.14 to -1.17)	0.00002
Original random-effects meta-analysis using a different metric (Cohen's h)	-0.80 (-1.07 to -0.54) Note: ES is Cohen's h, not lnOR	<0.0000001

Table D-13. Results of Sensitivity Analyses for Total Complications (EDM + CVP + Conventional Protocol vs. CVP + Conventional Protocol)

Sensitivity analysis	Summary log odds ratio (95% CI)	p value
Quantitative and Qualitative robustness		
Removal of Noblett et al.(42) (largest study)	-1.08 (-1.58 to -0.58)	0.00003
Removal of Wakeling et al.(43)	-1.06 (-1.60 to -0.52)	0.0001
Removal of Gan et al.(11)	-0.83 (-1.33 to -0.34)	0.001
Removal of Conway et al.(44) (smallest study)	-1.02 (-1.48 to -0.56)	0.00001
Removal of Conway et al., Gan et al. (two smallest studies)	-0.84 (-1.38 to -0.30)	0.002
Removal of Noblett et al., Wakeling et al. (two largest studies)	-1.27 (-1.97 to -0.56)	0.0005
Assumption of no effect in study that did not report total complications (Mythen and Webb)	-0.84 (-1.27 to -0.40)	0.0002
Original random-effects meta-analysis	-1.00 (-1.43 to -0.57)	0.000006
Original random-effects meta-analysis using a different metric (Cohen's h)	-0.47 (-0.67 to -0.28) Note: ES is Cohen's h, not lnOR	0.000003

Table D-14. Results of Sensitivity Analyses for Length of Hospital Stay (EDM + CVP + Conventional Protocol vs. CVP + Conventional Protocol)

Sensitivity analysis	Summary days (95% CI)	p value
Qualitative robustness		
Removal of Noblett et al.(42) (largest study)	-1.32 (-2.18 to -0.45)	0.003
Removal of Wakeling et al.(43)	-1.32 (-2.35 to -0.29)	0.01
Removal of Gan et al.(11)	-1.75 (-2.89 to -0.60)	0.003
Removal of Mythen and Webb(45)	-1.28 (-2.11 to -0.44)	0.003
Removal of Conway et al.(44) (smallest study)	-1.40 (-2.23 to -0.58)	0.0008
Removal of Noblett et al, Wakeling et al. (two largest studies)	-1.20 (-2.31 to -0.08)	0.04
Removal of Conway et al., Mythen and Webb (two smallest studies)	-1.29 (-2.14 to -0.45)	0.003
Original random-effects meta-analysis	-1.39 (-2.21 to -0.57)	0.0009
Random-effects meta-analysis using only means when available	-1.80 (-2.84 to -0.76)	0.0007
Random effects meta-analysis using more conservative assumptions about ranges and interquartile ranges	-1.29 (-2.30 to -0.27)	0.01
Random effects meta-analysis using more conservative assumptions about ranges and interquartile ranges and assuming medians are equal in the study by Mythen and Webb (which only reported means and SDs)	-1.03 (-2.05 to -0.01)	0.047
Random effects meta-analysis using a different metric (Hedges' g)	-0.31 (-0.49 to -0.12) Note: ES is Hedges' g, not days	0.001

Appendix E. Evidence Tables for Key Question 3

Table E-1. Patient Enrollment Criteria for Studies Addressing Key Question 3

Reference	Year	Inclusion Criteria	Exclusion Criteria
McKendry et al.(48)	2004	Patients undergoing cardiopulmonary bypass surgery who had provided informed consent	Patients with age <18 years, those undergoing off-pump surgery, those who declined consent, or those with relative contraindications to use of the esophageal Doppler probe (esophageal varices or other oropharyngeal and esophageal disease).

NR Not Reported

Table E-2. Characteristics of Patients in Studies Addressing Key Question 3

Author/ year	Year	Cardiac output monitoring procedure	N	Age (mean ±SD)	% female	ASA grade	Goldman cardiac risk index	POSSUM score	Hemoglobin level at baseline (g/dl)	% elective surgery
McKendry et al.(48)	2004	EDM (CardioQ)	89	65.6 (10.3)	37	NR	NR	NR	NR	88
		Conventional clinical protocol	85	66.7 (10.2)	33	NR	NR	NR	NR	84

NR Not Reported

Table E-3. Fluid Management Protocols in Studies Addressing Key Question 2

Reference	Year	Type of Surgery	Fluid management protocol for EDM group	Fluid management protocol for control group
McKendry et al.(48)	2004	Cardiopulmonary bypass surgery	Patients received continuous esophageal Doppler monitoring for first 4 h in intensive care following surgery. Patients received 200 ml fluid challenges (blood or colloid as appropriate), repeated until the stroke volume index increased to ≥35 ml/m^2. Nitrates and inotropes were given as required.	Patients received conventional management as determined by intensive care and surgical teams (based primarily on monitoring arterial and CVP, but also cardiac output if considered clinically indicated). Doppler recordings were made within 10 minutes of admission to the intensive care unit and at 4 h post-operatively.

Table E-4. Study Quality Evaluation

ECRI study quality scale - questions	Study
	McKendry et al.
1. Were patients randomly assigned to groups?	Yes
2. Did the study emply stochastic randomization?	Yes
3. Were any methods used to make the groups comparable- randomization, matching, etc.?	Yes
4. Were patients assigned to groups based on factors other than patient or physician preference?	Yes
5. Were the characteristics of the patients in different groups comparable?	NR
6. Did the patients in the different study groups have similar levels of performance on outcomes at baseline?	Yes
7. Was the study prospectively planned?	Yes
8. Did 85% or more of the patients complete the study?	Yes
9. Was there a less than 16% difference in completion rates in the study's groups?	Yes
10. Were all of the study's groups concurrently treated?	Yes
11. Was compliance with treatment greater than or equal to 85% in both of the groups?	Yes
12. Were both groups treated at the same centers?	Yes
13. Were subjects blinded to treatment?	Yes
14. Did the authors test and confirm that blinding of patients was maintained?	NR
15. Was the treating physician blinded to group assignment?	No
16. Were the outcome assessors blinded to group assignment?	No
17. Was there concealment of allocation?	Yes
18. Was the outcome of interest objective and was it objectively measured?[a]	Yes
19. Were the same methods used to measure outcomes in all of the study's groups?[a]	Yes

ECRI study quality scale - questions	Study
	McKendry et al.
20. Was the instrument used to measure the outcome standard?[a]	Yes
21. Was the same treatment given to all of the patients enrolled in the experimental group?	Yes
22. Was the same treatment given to all of the patients enrolled in the control group?	Yes
23. Were the follow-up times in all of the study's relevant groups approximately equal?	Yes
24. Was the funding for this study derived from a source that does not have a financial interest in its results?	No
25. Were the author's conclusions supported by the data in the results section?	Yes
Quality score	8.5
Quality rating	High

Table E-5. Specific Complications Reported in Included Studies

Reference	Year	Type of Surgery	Reported complications
Noblett et al.(42)	2006	Cardiopulmonary bypass surgery	Death, atrial fibrillation requiring treatment, pneumothorax, cerebral vascular accident, chest infection, sternal wound infection, gastrointestinal bleed or perforated duodenal ulcer, acute renal failure, pleural effusion, infected leg wound, aortic regurgitation

Table E-6. Results of Study Addressing Key Question 3

Study	N	% total deaths (n/N)			% total complications (n/N)			Length of hospital stay (days)		
		EDM group	Control group	p-value	EDM group	Control group	p-value	EDM group	Control group	p-value
McKendry et al.(48)	174	4.5 (4/89)	2.4 (2/85)	0.68	19.1 (17/89)	30.6 (26/85)	0.08	Mean: 11.4 Median: 7	Mean: 13.9 Median: 9	NR 0.02

Appendix F. Evidence Tables for Key Question 4

Table F-1. Studies Reporting That Esophageal Doppler Monitoring Did Not Cause Any Complications

Reference	Year	Number of patients	Esophageal Doppler probe model (manufacturer)
Cipolla et al.(49)	2006	6	Hemosonic 100 (Arrow International)
Noblett et al.(42)	2006	108	CardioQ (Deltex Medical)
Collins et al.(50)	2005	58	Hemosonic 100 (Arrow International)
Koliopanos et al.(51)	2005	55	ODM II[a] (Abbott Laboratories)
Sawai et al.(52)	2005	30	Hemosonic 100 (Arrow International)
Sharma et al.(53)	2005	35	TECO (Medicina Ltd)
Bein et al.(54)	2004	10	Hemosonic 100 (Arrow International)
Feldman et al.(55)	2004	13	Model not specified (Deltex Medical)
Seoudi et al.(58)	2003	15	Model EP90[a] (Deltex Medical)
Conway et al.(44)	2002	57	TECO 2 (Medicina Ltd)
Su et al.(59)	2002	24	Hemosonic 100 (Arrow International)
Venn et al.(46)	2002	30	CardioQ (Deltex Medical)
Odenstedt et al.(60)	2001	14	Dynemo 3000[b] (Somatec, Inc.)
Madan et al.(61)	1999	14	Model EP-90Q5[a] (Deltex Medical)
Elliott et al.(62)	1998	19	ODM 2[a] (Abbott Laboratories)
Lefrant et al.(18)	1998	64	ODM I[a] (Deltex Medical)
Valtier et al.(63)	1998	46	ODM I [a] (Deltex Medical)
Krishnamurthy et al.(71)	1997	16	ODM II[a] (Abbott Laboratories)
Sinclair et al.(47)	1997	40	ODM 2[a] (Abbott Laboratories)

[a] Earlier models of CardioQ
[b] Earlier model of Hemosonic