Hemodynamic interventions and -monitoring in critically ill patients

Geert Koster

https://www.publicatie-online.nl/publicaties/geert-koster/.

Printed by	Ridderprint | www.ridderprint.nl Alblasserdam, the Netherlands
Cover design & layout	Publiss | www.publiss.nl

rijksuniversiteit groningen

Hemodynamic interventions and -monitoring in critically ill patients

Proefschrift

ter verkrijging van de graad van doctor aan de
Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. C. Wijmenga
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

woensdag 1 december 2021 om 11.00 uur

door

Geert Koster

geboren op 24 augustus 1977
te Leiden

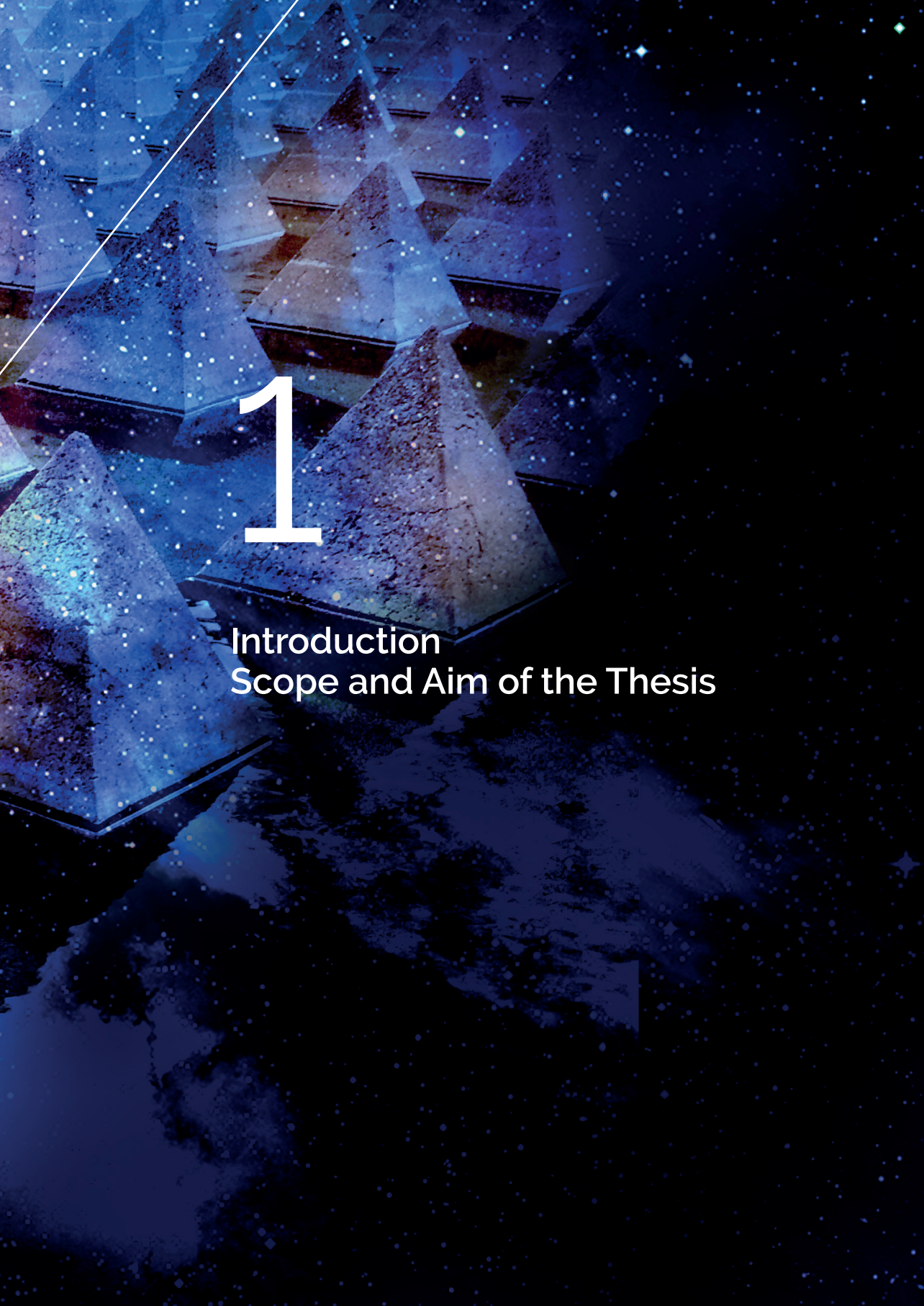
Promotor
Prof. dr. J.C.C. van der Horst

Copromoter
Dr. F. Keus

Beoordelingscommissie
Prof. dr. J.G. van der Hoeven
Prof. dr. J.C. ter Maaten
Prof. dr. M.A. Mariani

Paranimfen
Roel Bergstra
Waldo de Bildt

Table of contents



1

Introduction
Scope and Aim of the Thesis

General introduction

Every day we care for the sickest of our patients at the Intensive Care Unit (ICU). We examine the patient, and we request or perform diagnostic tests. Based on all the retrieved information, we arrive at one or multiple diagnoses, and we start or halt interventions, ultimately to benefit our patients.

Technical advancements have influenced the clinicians work importantly over the past decade; today, doctors have, for instance, electronic health data systems collecting many patients' (vital parameters) data automatically and have increasing opportunities for point-of-care diagnostics, ranging from blood gas analyses to handheld ultrasonography devices. The most accurate diagnostic information is preferred for the guidance of treatment strategies. It is assumed that the clinician is well informed and prepared. To remain up to date, clinicians need to know the literature and the skills for critical appraisal. To guide decisions on both diagnostics and interventions, clinicians at the ICU have to be trained and remain trained.

Interventions in medicine are ideally based on a clear pathophysiological rationale. In a critically ill patient, the pathophysiological basis of some form of cardiac dysfunction ought to be the incentive for prescribing inotropes. The degree of cardiac dysfunction on which to initiate therapy is usually not the point of discussion. For instance, many patients walk around with a left ventricular ejection fraction of just 10% in outpatients' settings. The clinical presentation triggers the prescription of inotropes, including a 'classic' characteristic of cold, clammy skin, mottling, reduced capillary refill and reduced consciousness reflecting a state of shock or so-called confusion of the body. This should be reflected by a derangement of hemodynamic variables, i.e. (low) cardiac output, (low or high) central/mixed venous saturation and (elevated) lactate. The lack of universally accepted definitions on what is considered (too) low, (too) high or normal introduces variety in definitions and in clinical decisions to start therapy and variety in treatment algorithms making comparisons between studies difficult.

Inotrope interventions have become the standard of care with supposedly beneficial effects on outcomes of patients. Though without the knowledge of true beneficial or harmful effects, we might walk on thin ice. Extensive research has been performed evaluating the effects of inotropes. Appraisal of all literature on the effect of inotropes is complicated by the large clinical heterogeneity resulting from variations in patient selection (different underlying pathology, indications, and definitions), variations in inotropes (timing and dosing), variations in outcomes (definitions) and settings. More importantly, the risks of bias and the risks of random error are often insufficiently addressed. Also, the outcomes are usually not selected according to their importance to patients. Frequently,

randomised trials that evaluated inotrope interventions have repeatedly applied strict selection criteria and focused on surrogates as their primary outcome (most likely for chances of significant findings), hampering their applicability in clinical practice.

Safe use of inotropes requires hemodynamic monitoring. Reductions of patient harm motivated the development of less invasive and non-invasive techniques and have led to a variety of devices presumed to measure some specific hemodynamic variable. Unfortunately, in patients with shock, these devices appear unreliable.[1,2] So, in the specific critically ill population in which inotropes are frequently used and typically claimed to be indicated, techniques to monitor the required hemodynamic variables still have to be validated. While the well-known invasive techniques (pulmonary artery catheter (PAC) and transpulmonary thermodilution method (TPTD)) have their disadvantages, these techniques have indeed been proven to be reliable in most critically ill patients. Due to the invasiveness, complexity and limitations of the PAC and TPTD [3,4], ultrasonography as a non-invasive tool has gained enormous popularity, not the least due to miniaturisation.

In contrast to the invasive techniques, ultrasonography of the heart may reveal the underlying pathophysiological problem (diagnosis) and monitor the effectiveness of interventions. In literature, there are clear pros and cons to its use in daily practice.[5-8] The basis of these pros and cons lies in the variety in the operator's expertise (and the absence of a 'golden' standard). The operator dependency and the fact that ultrasonography is not a continuous hemodynamic monitor should be considered.

Thesis outline

This thesis focuses on two themes: evaluating the evidence on the use of inotropes in critically ill patients and the evidence for monitoring hemodynamic variables using ultrasonography in critically ill patients.

Evidence on the use of inotropes in critically ill patients with cardiac dysfunction

Part 1 focuses on using inotropes in critically ill patients with cardiac dysfunction and/or shock. Inotropes are considered the first step in reversing the shock state in patients with cardiac dysfunction and hypotension, excluding septic cardiomyopathy.[9] Due to low quality - and sparse data, the choice for the inotrope and the exact starting triggers for initiating inotropes are mainly left to the treating physicians. **Dopamine** was one of the first inotropes used for this indication.[10] Although guidelines mention dopamine for use in heart failure[9], its use in clinical practice is declining along with reports on potential harms.[11] In **chapter 2,** we

evaluated the benefits and harms of dopamine.

In the eighties, a new class of inotropes emerged: the phosphodiesterase inhibitors. **Milrinone** is the most frequently evaluated drug in its class with the best side effect profile. This class distinguishes itself from other classes by reducing pulmonary – and systemic vascular resistance and its inotropic properties. Phosphodiesterase inhibitors are beneficial in cases of pulmonary hypertension or right ventricular failure. Although several guidelines suggest milrinone, its safety profile remains controversial.[9,12] Previous meta-analyses on the effects of milrinone had methodological flaws hampering their interpretation. **Chapter 3** described the systematic review and meta-analyses on milrinone's effects in critically ill patients with cardiac dysfunction.

The most recent inotropes (in the nineties) are the calcium sensitisers, which are claimed to be the ideal inotrope: increasing cardiac output without increasing myocardial oxygen demand and with just a slight decrease in peripheral resistance; all without relevant side-effects. **Levosimendan**, the commercially available drug in this class, is recommended in the guideline and can also reverse the effects of beta-blocker therapy if this is considered to contribute to the shock.[9] In **chapter 4,** we presented the results of a systematic review and meta-analyses on the effects of levosimendan.

Current evidence on critical care ultrasonography

Part 2 of this thesis focuses on the evidence for the use of critical care ultrasonography for hemodynamic monitoring and diagnosis in critically ill patients in the ICU. Clinical examination is still common to practice for every clinician and is considered the first diagnostic tool. Patients with shock may present with a variety of symptoms and clinical signs (see above). However, recent literature has shown that clinical examination accuracy for diagnosing shock (i.e., a low cardiac index) is poor.[13] So, additional hemodynamic variables are necessary for accurate diagnosis of shock.[14,15] Guidelines give no advice on which device or tool should be used to monitor any given intervention for reversal of shock. In **chapter 5**, we overviewed existing literature on the use of critical care ultrasonography in shock. Ultrasonography in critically ill patients is increasingly applied by various specialties, ranging from experts (i.e., certified ultrasonography technicians or cardiologists) to novices (i.e., medical students). In the past, the use of ultrasonography was limited to the radiologist or cardiologist. However, time, staffing and money restraints prohibited their patient care involvement 24 hours a day, seven days a week. Recent viewpoints support the use of ultrasonography by non-experts, provided adequate supervision, and quality control are accounted for.[16] Surveys show that in daily practice, both novices and experts use ultrasonography (to

varying degrees) to guide the management of critically ill patients in various settings, although several factors hamper the availability of non-experts. **Chapter 6** studied the feasibility of ultrasonography by medical students to accurately measure an ultrasonography-derived cardiac output.

The use and availability of new diagnostic tools, such as ultrasonography, questions the use of established diagnostic tests such as lung auscultation. Many different lung pathological processes can be observed in critically ill patients, challenging to discriminate only by clinical examination. Additional diagnostic tests are frequently used (i.e., chest X-ray or computer tomography of the thorax). The use of lung ultrasound for discriminating the various lung pathologies has frequently been studied and showed promising results regarding accuracy.[17,18] In time, lung ultrasound may replace lung auscultation in the critically ill patient in the ICU. The data of a sub-study of the Simple Intensive Care Studies-I (SICS-I) on the diagnostic accuracy of clinical examination versus lung ultrasound for detecting pulmonary oedema is presented in **chapter 7.**

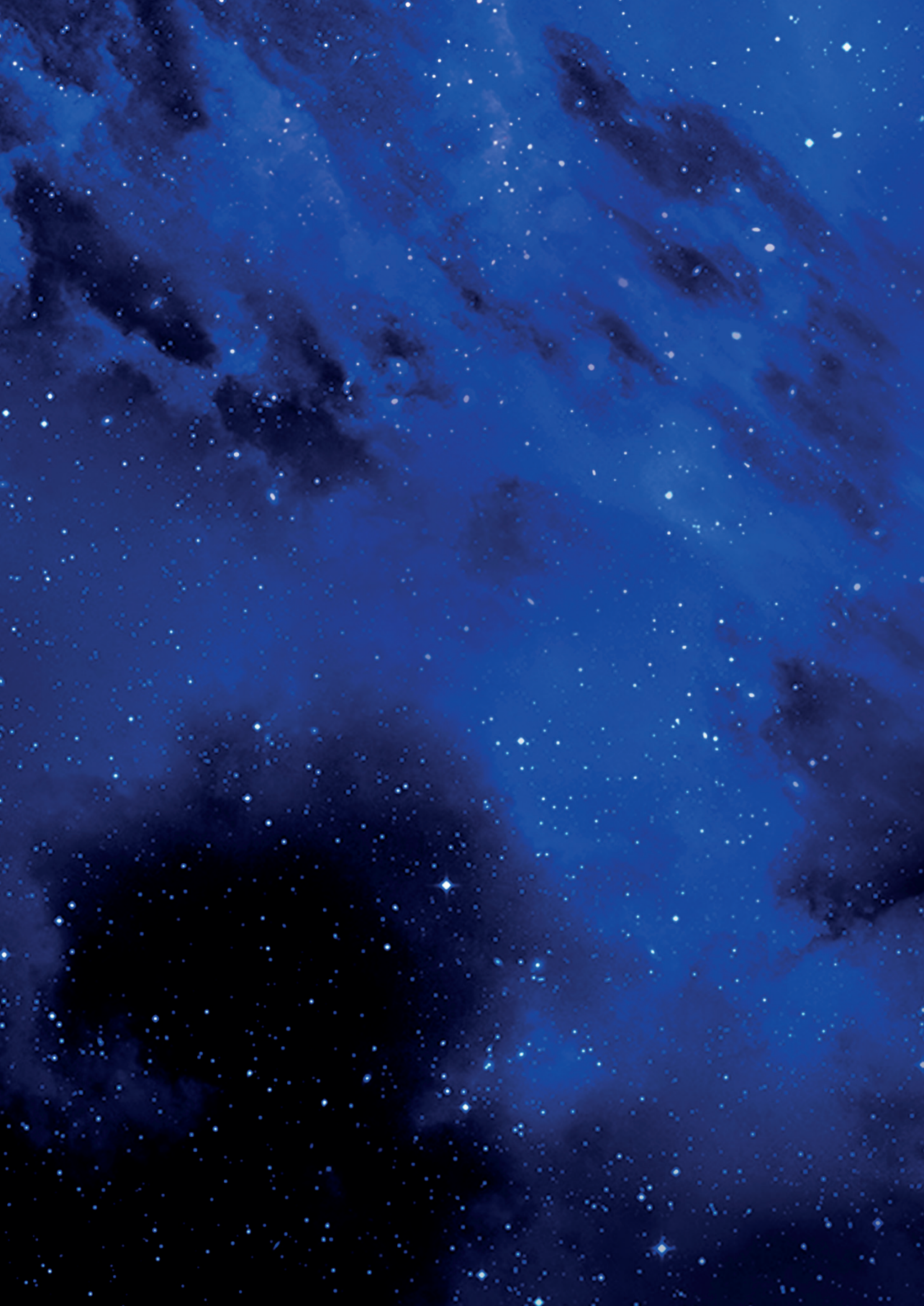
Finally, conclusions on the analyses' results on inotropes and ultrasonography in critically ill patients were integrated and discussed along with future perspectives in **chapter 8**. A summary is given in **chapter 9**.

References

1. Kaufmann T, Clement RP, Hiemstra B, Vos JJ, Scheeren TWL, Keus F, van der Horst ICC; SICS Study Group. Disagreement in cardiac output measurements between fourth-generation FloTrac and critical care ultrasonography in patients with circulatory shock: a prospective observational study. J Intensive Care 2019 Apr 11;7:21. doi: 10.1186/s40560-019-0373-5.
2. Teboul, JL., Saugel, B., Cecconi, M., De Backer D, Hofer CK, Monnet X, Perel A, Pinsky MR, Reuter DA, Rhodes A, Squara P, Vincent JL, Scheeren TW. Less invasive hemodynamic monitoring in critically ill patients. Intensive Care Med 2016:42;1350. doi:10.1007/s00134-016-4375-7
3. De Backer D, Vincent JL. The pulmonary artery catheter: is it still alive? Curr Opin Crit Care 2018 Jun;24(3):204-208. doi: 10.1097/MCC.0000000000000502.
4. De Backer D, Bakker J, Cecconi M, Hajjar L, Liu DW, Lobo S, Monnet X, Morelli A, Myatra SN, Perel A, Pinsky MR, Saugel B, Teboul JL, Vieillard-Baron A, Vincent JL. Alternatives to the Swan-Ganz catheter. Intensive Care Med 2018 Jun;44(6):730-741. doi: 10.1007/s00134-018-5187-8. Epub 2018 May 3.
5. Vieillard-Baron A, Millington SJ, Sanfilippo F, ChewJ M, Diaz-Gomez J, McLean A, Pinsky MR, Pulido J, Mayo P, Fletcher N. Intensive Care Med 2019;45:770-788. doi:10.1007/s00134-019-05604-2
6. Cecconi M, De Backer D, Antonelli M, Beale R, Bakker J, Hofer C, Jaeschke R, Mebazaa A, Pinsky MR, Teboul JL, Vincent JL, Rhodes A. Consensus on circulatory shock and hemodynamic monitoring. Task force of the European Society of Intensive Care Medicine. Intensive Care Med 2014;40:1785–1815
7. Cholley B. Echocardiography in the intensive care unit: beyond "eyeballing". A plea for the broader use of the aortic velocity-time integral measurement. Intensive Care Med 2019 Jun;45(6):898-901. doi: 10.1007/s00134-019-05618-w.
8. Wetterslev M, Moller-Sorensen H, Johansen RR, Perner A. Systematic review of cardiac output measurements by echocardiography vs. thermodilution: the techniques are not interchangeable. Intensive Care Med 2016;42:1223–123
9. Ponikowski P, Voors AA, Anker SD, Bueno H, Cleland JGF, Coats AJS, Falk V, González-Juanatey JR, Harjola VP, Jankowska EA, Jessup M, Linde C, Nihoyannopoulos P, Parissis JT, Pieske B, Riley JP, Rosano GMC, Ruilope LM, Ruschitzka F, Rutten FH, van der Meer P; ESC Scientific Document Group., ESC Scientific Document Group. 2016 ESC Guidelines for the diagnosis and treatment of acute and chronic heart failure: The Task Force for the diagnosis and treatment of acute and chronic heart failure of the European Society of Cardiology (ESC); Developed with the special contribution of the Heart Failure Association (HFA) of the ESC. Eur Heart J. 2016 Jul 14;37(27):2129-2200. doi: 10.1093/eurheartj/ehw128
10. MacCannell KL, McNay JL, Meyer MB, Goldberg LI. Dopamine in the treatment of hypotension and shock. N Engl J Med. 1966 Dec 22;275(25):1389-98. doi: 10.1056/NEJM196612222752501
11. De Backer D, Biston P, Devriendt J, Madl C, Chochrad D, Aldecoa C, Brasseur A, Defrance P, Gottignies P, Vincent JL; SOAP II Investigators. Comparison of dopamine and norepinephrine in the treatment of shock. N Engl J Med. 2010 Mar 4;362(9):779-89. doi: 10.1056/NEJMoa0907118.
12. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE Jr, Drazner MH, Fonarow GC, Geraci SA, Horwich T, Januzzi JL, Johnson MR, Kasper EK, Levy WC, Masoudi FA, McBride PE, McMurray JJ, Mitchell JE, Peterson PN, Riegel B, Sam F, Stevenson LW, Tang WH, Tsai EJ, Wilkoff BL. 2013 ACCF/AHA Guideline for the Management of Heart Failure: Executive Summary. J Am Coll

Cardiol 2013;62:1495–1539. doi: 10.1016/j.jacc.2013.05.020

13. Hiemstra B, Koster G, Wiersema R, Hummel YM, van der Harst P, Snieder H, Eck RJ, Kaufmann T, Scheeren TWL, Perner A, Wetterslev J, de Smet AMGA, Keus F, van der Horst ICC; SICS Study Group. The diagnostic accuracy of clinical examination for estimating cardiac index in critically ill patients: the Simple Intensive Care Studies-I. Intensive Care Med. 2019 Feb;45(2):190-200. doi: 10.1007/s00134-019-05527-y
14. Vincent JL, De Backer D. Circulatory shock. N Engl J Med. 2013 Oct 31;369(18):1726-34. doi: 10.1056/NEJMra1208943
15. Cecconi M, De Backer D, Antonelli M, Beale R, Bakker J, Hofer C, Jaeschke R, Mebazaa A, Pinsky MR, Teboul JL, Vincent JL, Rhodes A. Consensus on circulatory shock and hemodynamic monitoring. Task force of the European Society of Intensive Care Medicine. Intensive Care Med. 2014 Dec;40(12):1795-815. doi: 10.1007/s00134-014-3525-z
16. Neskovic AN, Edvardsen T, Galderisi M, Garbi M, Gullace G, Jurcut R, Dalen H, Hagendorff A, Lancellotti P; European Association of Cardiovascular Imaging Document Reviewers:, Popescu BA, Sicari R, Stefanidis A.Focus cardiac ultrasound: the European Association of Cardiovascular Imaging viewpoint. Eur Heart J 2014;15:956-960. doi:10.1093/ehjci/jeu081
17. Mojoli F, Bouhemad B, Mongodi S, Lichtenstein D. Lung Ultrasound for Critically Ill Patients. J Respir Crit Care Med. 2019 Mar 15;199(6):701-714. doi: 10.1164/rccm.201802-0236CI.
18. Volpicelli G, Elbarbary M, Blaivas M, Lichtenstein DA, Mathis G, Kirkpatrick AW, Melniker L, Gargani L, Noble VE, Via G, Dean A, Tsung JW, Soldati G, Copetti R, Bouhemad B, Reissig A, Agricola E, Rouby JJ, Arbelot C, Liteplo A, Sargsyan A, Silva F, Hoppmann R, Breitkreutz R, Seibel A, Neri L, Storti E, Petrovic T; International Liaison Committee on Lung Ultrasound (ILC-LUS) for International Consensus Conference on Lung Ultrasound (ICC-LUS). International evidence-based recommendations for point-of-care lung ultrasound. Intensive Care Med. 2012 Apr;38(4):577-91. doi: 10.1007/s00134-012-2513-4.



PART 1

Hemodynamic interventions

2

Dopamine for cardiac dysfunction in critically ill adults: a systematic review with meta-analysis and Trial Sequential Analysis

Hiemstra B, Koster G, Wetterslev J, Gluud C, Jakobsen JC, Scheeren TW, Keus F, van der Horst ICC.
Acta Anaesthesiologica Scandinavica 2019 Apr;63(4):424-437.
doi: 10.1111/aas.13294

Abstract

Background

Dopamine has been used in patients with cardiac dysfunction for more than five decades. Yet, no systematic review has assessed the effects of dopamine in critically ill patients with cardiac dysfunction.

Methods

This systematic review was conducted following The Cochrane Handbook for Systematic Reviews of Interventions. We searched for trials including patients with observed cardiac dysfunction published until 19 April 2018. Risk of bias was evaluated and Trial Sequential Analyses were conducted. The primary outcome was all-cause mortality at longest follow-up. Secondary outcomes were serious adverse events, myocardial infarction, arrhythmias, and renal replacement therapy. We used GRADE to assess the certainty of the evidence.

Results

We identified 17 trials randomising 1218 participants. All trials were at high risk of bias and only one trial used placebo. Dopamine compared with any control treatment was not significantly associated with relative risk of mortality (60/457 [13%] vs 90/581 [15%]; RR 0.91; 95% confidence interval 0.68 - 1.21) or any other patient-centered outcomes. Trial Sequential Analyses of all outcomes showed that there was insufficient information to confirm or reject our anticipated intervention effects. There were also no statistically significant associations for any of the outcomes in subgroup analyses by type of comparator (inactive compared to potentially active), dopamine dose (low compared to moderate dose), or setting (cardiac surgery compared to heart failure).

Conclusion

Evidence for dopamine in critically ill patients with cardiac dysfunction is sparse, of low quality, and inconclusive. The use of dopamine for cardiac dysfunction can neither be recommended nor refuted.

Introduction

Dopamine is a natural catecholamine which has various cardiovascular effects throughout a dose-dependent activation of dopaminergic, α- and β-adrenergic receptors.[1] Low-dose dopamine (<4 μg·kg^{-1}·min^{-1}) is hypothesised to primarily provide mesenteric and renal arteriole vasodilation, moderate-dose dopamine (4-10 μg·kg^{-1}·min^{-1}) is hypothesised to have particularly positive inotropic and chronotropic effects, and high-dose dopamine (>10 μg·kg^{-1}·min^{-1}) is considered a vasopressor due to the increase of systemic vascular resistance.[1, 2] These doses are arbitrary as there is a wide interindividual variability of dopamine receptor sensitivity.[2]

Guidelines for treatment of heart failure mention dopamine among other drugs to treat acute heart failure.[3, 4] Several randomised clinical trials (RCTs) have failed to show clinical benefits associated with use of dopamine in patients with acute heart failure[5-7] and circulatory shock.[8] Previous meta-analyses advocate cautious use of high-dose dopamine.[9] Despite the decline in its use, dopamine is still the used inotrope in 25% of acute heart failure patients and in 14% of the patients undergoing cardiac surgery.[10, 11]

The debate about the benefits and harms of dopamine in critically ill patients with cardiac dysfunction remains.[11, 12] Our objective was to conduct a systematic review with meta-analyses and Trial Sequential Analyses (TSA) of RCTs comparing the benefits and harms of dopamine compared to placebo, no intervention, or any potentially active comparator in critically ill patients with cardiac dysfunction.

Methods

This systematic review was conducted following our published protocol (CRD42016042867),[13] the recommendations of The Cochrane Handbook for Systematic Reviews of Interventions,[14] The Cochrane Hepato-Biliary Group Module,[15] and was reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[16]

Eligibility criteria

We considered all RCTs eligible for inclusion irrespective of language, blinding, publication status, sample size, or control intervention(s) for assessment of benefits and harms. Quasi-randomised and observational studies were included for assessment of potential harms and results were analysed separately.

Only RCTs with critically ill adult patients with cardiac dysfunction were included in our main analysis. Critical illness encompassed any clinical setting wherein patients with objectively measured cardiac dysfunction seemed to require intravenous dopamine without restrictions on dose or duration of administration. Cardiac dysfunction was defined as a left ventricular ejection fraction (LVEF) below 45% and/or a low cardiac output syndrome. Low cardiac output syndrome was defined as a pre-existing or developing state of cardiac insufficiency with underlying left- or right-ventricular systolic dysfunction seemed to require inotrope support to maintain a systolic blood pressure >90 mmHg and a cardiac index >2.2 $L \cdot min^{-1} \cdot m^{-2}$.[17] RCTs including both patients with and without cardiac dysfunction were included in the review only if the majority (more than 50%) of the included patients had cardiac dysfunction. During the selection process, we had to exclude a substantial number of trials because not all trials objectively measured cardiac dysfunction for each patient. We realised that our eligibility criteria may not reflect all the situations in which doctors decide to administer dopamine. To increase the external validity of our systematic review, we conducted a post hoc analysis including trials in which a substantial proportion of patients (more than 25%) were assumed to have cardiac dysfunction.

Outcomes

The primary outcome was all-cause mortality. The secondary outcomes were serious adverse events (SAEs), myocardial infarction, arrhythmias (including supra- and ventricular tachycardia and fibrillation), and renal failure requiring renal replacement therapy. SAEs were defined according to the International Conference on Harmonisation of Good Clinical Practice definitions, excluding mortality to avoid double counts.[18] Myocardial infarction, arrhythmias, and renal replacement therapy were defined according to the criteria used in the individual trials. We included data at longest follow-up.

Search methods

We used a sensitive search strategy that was likely to include all clinical settings wherein cardiac dysfunction was prevalent: eg shock, heart failure, cardiac surgery (Appendix S1). We searched the Cochrane Central Register of Controlled Trials (CENTRAL), PubMed, Web of Science, CINAHL, and Embase until 19 April 2018. We also searched the World Health Association's (WHO's) trial platform, ClinicalTrials.gov, and FDA and EMA homepages for ongoing trials. Last, we searched the references of the selected trials and previous meta-analyses to identify further relevant trials.

Trial selection, data extraction, and bias risk assessments

Two authors independently identified trials for inclusion and extracted study, patient and intervention characteristics, evaluated outcomes, and risks of bias according to the domains of bias in The Cochrane Handbook for Systematic Reviews of Interventions.[14] Trials with one or more of the risks of bias domains classified at high or unclear risk were considered trials at high risk of bias.[14] The authors of the individual trials were contacted in case of any unclear or missing information.

All data on the outcomes of all trials were assessed for the risks of systematic errors ('bias'), the risks of other design errors, and the risks of random errors. The three-dimensional Manhattan error matrix plot was used to facilitate the overview of available evidence at a glance.[19] We used a funnel plot to explore small trial bias.[14]

Statistical methods

Results were presented as relative risks (RR), odds ratios (OR), and Peto's OR with 95% confidence interval (CI) when applicable. We used both a fixed-effect model and a random-effects model for our meta-analyses and presented both models in case of discrepancy. Considering the anticipated clinical diversity, we emphasised the results from the random-effects model as it provides the most conservative estimate of effect and/or CI. Heterogeneity was explored by inspection of forest plots and the chi-squared test with significance set at P-value of 0.10, and the quantity of heterogeneity was measured by I^2.[20]

We used TSA on all outcomes to control for the risks of random errors ("the play of chance") and adjust the thresholds for statistical significance when few data are present or when tested repeatedly, comparable to interim analyses in a single RCT. TSA calculates a diversity-adjusted required information size (RIS) which compares well to a sample size calculation for an RCT, and widens the thresholds for statistical significance before the RIS is accrued. The RIS was calculated based on an anticipated relative risk reduction (RRR) of 10% and appropriately adjusted for heterogeneity in terms of diversity (D2) according to an overall type-I error of 5% and a power of 90% considering early and repetitive testing.[21] P-values less than TSA-adjusted significance levels were considered statistically significant.[21] We explain the interpretation of a TSA-graph in Figure S1. The concepts of TSA are explained in detail in the TSA Manual (http://www.ctu.dk/tsa) as well as in a recent overview.[21] We used the software package Review Manager 5.3.5 for the meta-analyses and the TSA program v.0.9.5.10 beta (http://www.ctu.dk/tsa) for the TSA.

Sensitivity and subgroup analyses

All outcomes were dichotomous. We constructed best-worst and worst-best case scenarios as sensitivity analyses for participants lost to follow-up. Following our protocol, we conducted subgroup analyses to explore clinical heterogeneity according to: (a) risk of bias in trials; (b) control intervention (inactive compared to a potentially active control); (c) trials assessing a low dose (<4 $\mu g \cdot kg^{-1} \cdot min^{-1}$) compared to a moderate (4-10 $\mu g \cdot kg^{-1} \cdot min^{-1}$) or high dose (>10 $\mu g \cdot kg^{-1} \cdot min^{-1}$); (d) clinical setting (patients having cardiac surgery compared to patients not having cardiac surgery).

GRADE assessments

We used the Grading of Recommendations Assessment, Development and Evaluations (GRADE) approach to rate and assess the quality of the body of evidence for each outcome and constructed a "Summary of findings" table.[22]

Results

Study selection

After screening the literature search, titles and abstracts, 341 articles out of 10 858 hits remained (Figure 1). After assessment of full-texts, 86 studies were included in our systematic review. Additional data was obtained from three studies.[5,6,23] The main meta-analysis included 17 RCTs with in total 1218 patients.[5-7,24-37] Two observational studies were assessed for harmful outcomes.[23,38]

Characteristics of included trials

The characteristics of the 17 trials included in our meta-analyses are summarised in Table 1. In- and exclusion criteria of each trial are presented in Table S1. Nine trials had a two-arm design, seven trials consisted of three treatment arms, and one administered four different treatments. One trial was placebo-controlled,[7] four trials used no intervention in the control group,[6,25,30,32] and 14 trials used a potentially active control intervention: eight trials administered an inotropic drug and six a diuretic drug. The administration duration of the study drugs varied from only during the perioperative period up to a maximum of 5 days. Seven of the 17 trials included solely patients who all had objectively verified cardiac dysfunction defined by an LVEF below 45% or a low cardiac output syndrome.[25,27,29,34,36,37] In a sensitivity analyses we only included these seven trials; findings were comparable to the analysis of 17 trials (e-Table 2, Appendix S2).

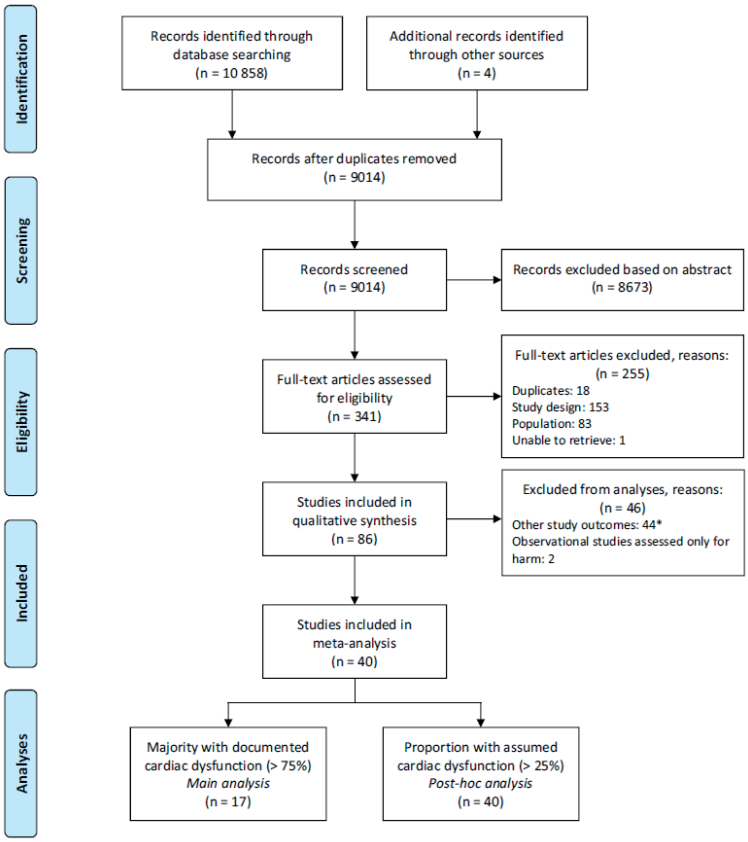


FIGURE 1. PRISMA flow diagram. *All authors from the studies published since 1990 were contacted for additional data in case of missing outcomes of interest

Risk of bias

All 17 trials were at overall high risk of bias (Figure 2). Fourteen trials were at high risk of other bias, because nine trials (53%) did not provide a statement on conflicts of interest, two trials (12%) allowed cross-over to another inotrope, and three trials (18%) were at risk of vested interests.

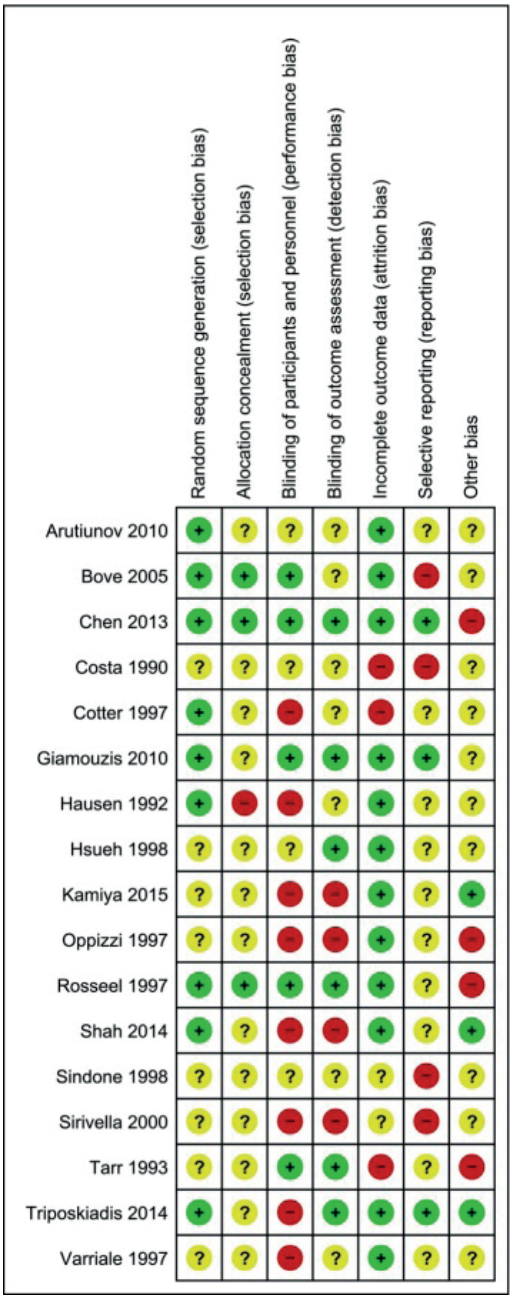


FIGURE 2. Risk of bias assessment. Red, high risk; yellow, unclear risk; green, low risk

TABLE 1. Characteristics of the included trials

Trial, year	N	Dopamine dose	Comparator(s)	Cardiac function	Outcomes
Acute heart failure					
Kamiya[24]	24	*Low dose*: 1.9 ± 0.8 μg•kg^{-1}•min^{-1}	Furosemide 17.1 ± 7.2 μg•kg^{-1}•min^{-1}	LVEF per group: Dopamine: 38% ± 16% Comparator: 43% ± 20%	Mortality (in-hospital) Serious adverse events Arrhythmias
Chen[7]	360	*Low dose*: 2.0 μg•kg^{-1}•min^{-1}	Placebo	LVEF: 33% (IQR 22-50%) Proportion LVEF <50%: 74%	Mortality (60 days) Serious adverse events Arrhythmias
Varriale[25]	20	*Low dose*: 2.0 μg•kg^{-1}•min^{-1}	Control	Mean LVEF: 28.3% ± 9.1% Depressed LV-function was an inclusion criterion	Mortality (in-hospital) Arrhythmias
Shah[26]	90	*Low dose*: 2.5 μg•kg^{-1}•min^{-1}	**(1)** Control **(2)** Furosemide 2dd 50 mg	Mean LVEF: 33%	Mortality (30 days) Serious adverse events
Arutiunov[27]	41	*Low dose*: 3.1 ± 0.2 μg•kg^{-1}•min^{-1}	Levosimendan (unknown dose) + ivabradine 2dd 5 mg	Mean LVEF: 22% LVEF < 35% was an inclusion criterion	Mortality (30 days) Myocardial infarction
Hsueh[28]	20	*Moderate dose*: 4.0 μg•kg^{-1}•min^{-1}	Dobutamine 4.0 μg•kg^{-1}•min^{-1}	LVEF: ± 33% ± 10 LVEF < 45% was an inclusion criterion	Mortality (72 hours) Arrhythmias
Cotter[29]	20	*Moderate dose*: **(1)** 4.0 + furosemide 2dd 40 mg **(2)** 4.0 + furosemide 5 mg•kg^{-1}	Furosemide 10 mg•kg^{-1}•24h^{-1}	LVEF > 40% was an exclusion criterion	Mortality (in-hospital) Arrhythmias
Giamouzis[5]	60	*Moderate dose*: 5.0 μg•kg^{-1}•min^{-1}	Furosemide 20 mg•h^{-1}	LVEF: 36% ± 12% Proportion LVEF < 40%: 70%	Mortality (60 days) Serious adverse events
Triposkiadis[6]	161	*Moderate dose*: 5.0 μg•kg^{-1}•min^{-1}	**(1)** Control **(2)** Furosemide 20 mg•h^{-1}	LVEF: 31% (25% - 45%) Proportion LVEF <40%: 58%	Mortality (1 year) Serious adverse events Arrhythmias Renal replacement therapy
Sindone[30]	67	Not specified (abstract only)	**(1)** Control **(2)** Dobutamine **(3)** Milrinone	CI 1.9 ± 0.7 L•min^{-1}•m^{-2}	Mortality (1 year)
Cardiac surgery					

Trial, year	N	Dopamine dose	Comparator(s)	Cardiac function	Outcomes
Sirivella[31]	100	*Low dose*: **(1 + 2)** 2-3 μg•kg^{-1}•min^{-1} + mannitol + furosemide 0.6-0.8 mg•kg^{-1}* (other inotropes were given)	Furosemide 1.4-3 mg•kg^{-1} + bumetadine 0.014 mg•kg^{-1} (other inotropes were given)	LVEF: 35% Mean CO: 2.4 ± 0.2 L•min^{-1}	Renal replacement therapy
Costa[32]	36	*Low dose*: **(1)** 2.5 μg•kg^{-1}•min^{-1} **(2)** 2.5 μg•kg^{-1}•min^{-1} + nitroprusside	Control	Renal dysfunction was attributable to severe HF in all but three patients	Renal replacement therapy
Bove[33]	80	*Low dose*: 2.5 μg•kg^{-1}•min^{-1} (65% received other inotropes)	Fenaldopam 0.5 μg•kg^{-1}•min^{-1} (68% received other inotropes)	LVEF per group: Dopamine: 43% ± 16% Comparator: 44% ± 17%	Mortality (in-hospital) Renal replacement therapy
Rosseel[34]	70	*Low dose*: 3.1 ± 1.6 μg•kg^{-1}•min^{-1}	Dopexamine 1.2 ± 0.6 μg•kg^{-1}•min^{-1}	Low cardiac output syndrome was an inclusion criterium	Mortality (in-hospital) Serious adverse events
Hausen[35]	41	*Moderate dose*: 5-7 μg•kg^{-1}•min^{-1} + glyceroltrinitrate (57% received adrenaline)	**(1)** Enoximone 5-20 μg•kg^{-1}•min^{-1} (62% received adrenaline) **(2)** Piroximone 3-6 μg•kg^{-1}•min^{-1} (43% received adrenaline)	A preoperative cardiac index < 2.5 L•min^{-1}•m^{-2} was an inclusion criterion	Mortality (6 ± 3 months) Myocardial infarction Arrhythmias
Oppizzi[36]	26	*Moderate dose*: 5-10 μg•kg^{-1}•min^{-1} (15% crossed over)	Enoximone bolus 0.5 mg•kg^{-1}, followed by 5-10 μg•kg^{-1}•min^{-1} (5% crossed over)	LVEF < 35% was an inclusion criterion	Mortality (in-hospital) Serious adverse events Myocardial infarction Arrhythmias
Tarr[37]	75	*Moderate dose*: 5-10 μg•kg^{-1}•min^{-1} (36% received other inotropes)	**(1)** Enoximone 5-10 μg•kg^{-1}•min^{-1} (0% received other inotropes) **(2)** Dobutamine 7-14 μg•kg^{-1}•min^{-1} (12% received other inotropes)	Cardiac index per group: Dopamine: 1.73 ± 0.08 L•min^{-1}•m^{-2} Comparators: 1.83 ± 0.11 L•min^{-1}•m^{-2}	Mortality (in-hospital)

Legend. AHF, acute heart failure; LVEF, left-ventricular ejection fraction. Trials are sorted by setting and administered dose. We selected studies that provided data on cardiac function and accepted definitions of diagnoses according to criteria used in each individual RCT.

[a]The timing of administering the experimental intervention differed between the treatment arms.

Outcomes

Table 2 summarises the meta-analysed intervention effect estimates. Due to absence of trials at overall low risk of bias and also due to absence of trials administering high-dose dopamine, we were unable to conduct these predefined subgroup analyses. None of the comparisons or outcomes could be analysed with the TSA using our prespecified parameters. As a sensitivity analyses, we conducted a TSA with a type I error of 5%, type II error of 10%, and an RRR of 20% on our primary outcome mortality to evaluate the direction of the cumulative Z-curve.

TABLE 2. Risk and odds ratios of all outcomes with subgroups analyses

	Trials*	Patients	Events	RR or OR	95% CI	Test for Interaction
Mortality	**15**	**1038**	**150**	**0.91**	**0.68 to 1.21**	**P = 1.00**
(1) Placebo or control	5	452	84	0.90	0.61 to 1.33	
(1) Potentially active control	12	586	66	0.92	0.59 to 1.43	
(2) Low dose dopamine	7	568	68	0.84	0.54 to 1.30	
(2) Moderate dose dopamine	7	403	74	0.98	0.65 to 1.47	
(3) Acute heart failure	10	746	132	0.90	0.67 to 1.23	
(3) Cardiac surgery	5	292	18	0.93	0.35 to 2.48	
Serious adverse events	**6**	**582**	**113**	**1.20**	**0.91 to 1.57**	**P = 0.92**
(1) Placebo or control	2	324	41	1.48	0.82 to 2.67	
(1) Potentially active control	5	258	72	1.34	0.75 to 2.40	
(2) Low dose dopamine	3	335	80	1.16	0.78 to 1.71	
(2) Moderate dose dopamine	3	267	33	1.70	0.86 to 3.39	
(3) Acute heart failure	4	486	59	1.54	0.94 to 2.53	
(3) Cardiac surgery	2	96	54	1.45	0.43 to 4.90	
Myocardial infarction	**5**	**339**	**16**	**1.63**	**0.56 to 4.71**	**P = 0.99**
(1) Placebo or control	1	83	2	2.00	0.12 to 33.2	
(1) Potentially active control	5	256	14	1.57	0.50 to 4.95	
(2) Low dose dopamine	2	111	8	1.68	0.15 to 18.8	
(2) Moderate dose dopamine	3	228	8	1.99	0.47 to 8.36	
(3) Acute heart failure	2	202	7	2.91	0.55 to 15.3	
(3) Cardiac surgery	3	137	9	1.09	0.27 to 4.33	
Ventricular tachyarrhythmias	**8**	**538**	**24**	**1.46**	**0.52 to 4.10**	**P = 0.97**
(1) Placebo or control	3	329	12	3.23	0.36 to 28.6	
(1) Potentially active control	6	209	12	0.94	0.28 to 3.15	
(2) Low dose dopamine	3	270	10	2.12	0.08 to 55.3	
(2) Moderate dose dopamine	5	268	14	1.09	0.35 to 3.43	

	Trials*	Patients	Events	RR or OR	95% CI	Test for Interaction
(3) Acute heart failure	6	471	21	1.29	0.38 to 4.39	
(3) Cardiac surgery	2	67	3	2.18	0.17 to 27.6	
Renal replacement therapy	**4**	**371**	**51**	**0.44**	**0.07 to 2.75**	**P = 0.94**
(1) Placebo or control	2	113	1	0.64	0.03 to 15.3	
(1) Potentially active control	3	258	50	0.42	0.05 to 3.67	
(2) Low dose dopamine	3	210	48	0.26	0.02 to 3.43	
(2) Moderate dose dopamine	1	161	3	1.16	0.15 to 9.15	
(3) Acute heart failure	1	161	3	1.16	0.15 to 9.15	
(3) Cardiac surgery	3	210	48	0.26	0.02 to 3.43	
Atrial tachyarrhythmias	**2**	**181**	**3**	**1.16**	**0.14 to 9.65**	**P = 0.99**
(1) Placebo or control	2	103	1	0.64	0.03 to 16.2	
(1) Potentially active control	1	78	2	1.81	0.11 to 30.2	
(2) Low dose dopamine	1	20	0	-	-	
(2) Moderate dose dopamine	1	161	3	1.16	0.14 to 9.65	
(3) Acute heart failure	2	181	3	1.16	0.14 to 9.65	
(3) Cardiac surgery	0	0	0	-	-	

Legend. RR, relative risk; OR, odds ratio; CI, confidence interval.

*Some trials compared dopamine with both a control intervention and a potentially active control (i.e. three-arm design), which is why the combined number of trials in subgroup analysis 1 differ from the total amount.

Comparison 1: all critically ill patients with cardiac dysfunction

All-cause mortality

All-cause mortality was reported in 15 of the 17 trials with a total of 1038 included patients. One trial reported mortality only during their 72-hour study period, seven trials reported in-hospital mortality, four trials 30- to 60-day mortality, and three trials mortality after 6-12 months of follow-up (Table 1). Dopamine did not statistically significantly affect mortality when compared with any control intervention (60/457 [13%] vs 90/581 [15%]; RR 0.91; 95% CI 0.68-1.21; I^2 0%), or when compared with an inactive control or with a potentially active control (Figure 3).

TSA on all trials showed that 19% of the RIS data was accrued and that about another 4292 patients need to become randomised in RCTs before the RIS will be reached (Figure 4; RR 0.91; TSA-adjusted CI 0.50-1.67).

SAEs

The occurrence of SAEs was reported in six trials with 582 included patients. Dopamine was not statistically significantly associated with SAEs when compared with any control intervention (62/268 (23%) vs 51/314 (16%); RR 1.20; 95% CI 0.91 to 1.57; I^2 2%; Figure 5). In a sensitivity analysis, we included mortality in our SAEs and found no statistically significant associations (122/457 (27%) vs 141/581 (24%); RR 1.06; 95% CI 0.89 to 1.27; I^2 0%). TSA on all trials showed that only 12% of the data was accrued and that about 4405 additional patients need to become randomised in RCTs before the RIS will be reached (RR 1.20; TSA-adjusted CI 0.41 to 3.41; Figure S2).

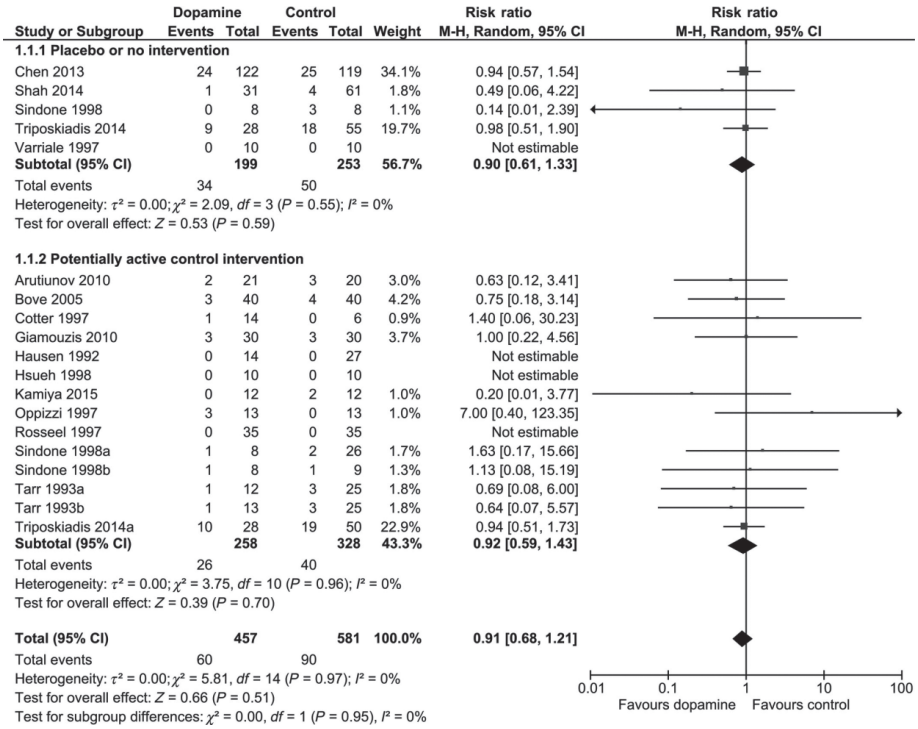


FIGURE 3. Forest plot of mortality in all trials stratified by intervention. Forest plot of all-cause mortality in trials stratified by intervention. Size of squares for risk ratio (RR) reflects the weight of the trial in the meta-analysis. Horizontal bars are 95% confidence intervals (CI)

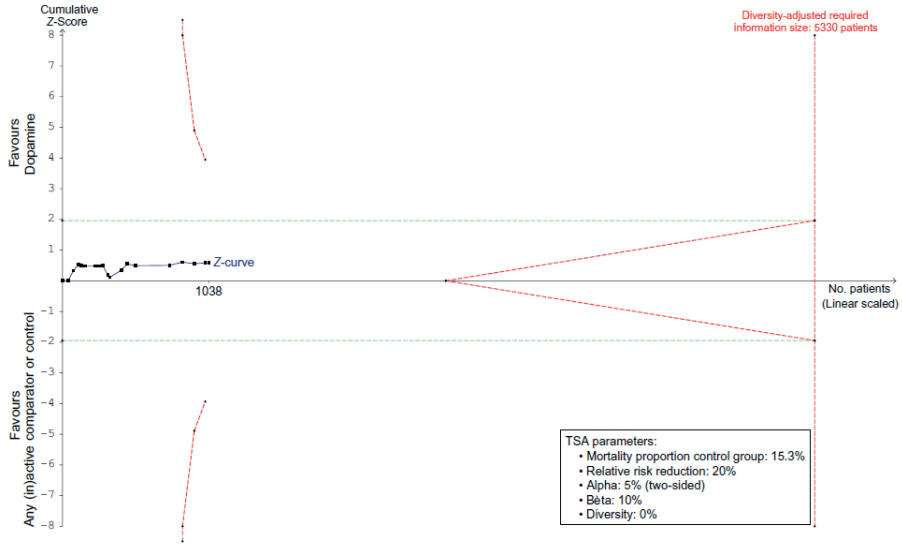


FIGURE 4. Trial Sequential Analysis for all-cause mortality. The Trial Sequential Analysis is based on 15 trials, which is the meta-analysed effect of dopamine vs any (in)active comparator intervention. The blue cumulative z-curve was constructed using a random-effects model. The horizontal green dotted lines represent the conventional naïve boundaries for benefit (positive) or harm (negative). The red dotted lines represent the trial sequential boundaries for benefit (positive), harm (negative), or futility (middle triangular area)

Other outcomes

There were no significant differences in favour of any intervention on the other outcomes (Table 2). None of the outcomes could be analysed with TSA using our prespecified parameters because less than 5% of RIS was accrued.

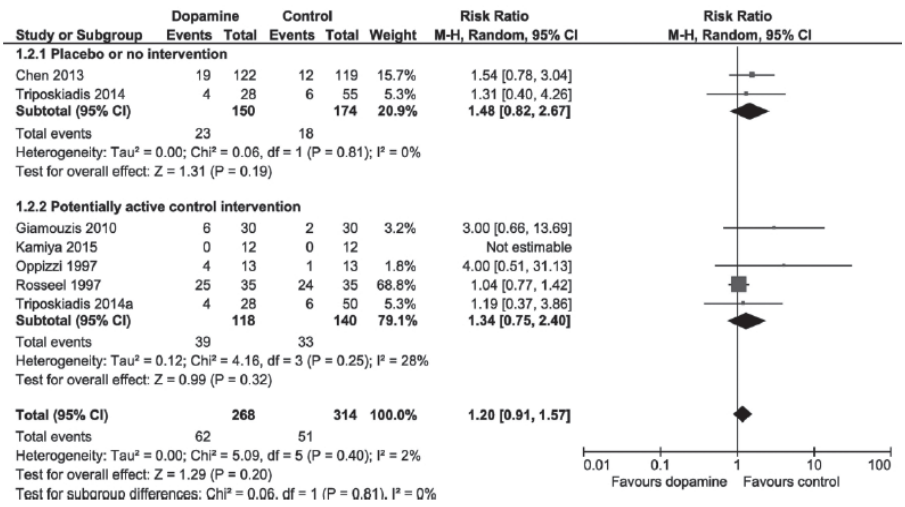


FIGURE 5. Forest plot of serious adverse events in all trials stratified by intervention. Size of squares for risk ratio (RR) reflects the weight of the trial in the meta-analysis. Horizontal bars are 95% confidence intervals (CI).

Comparison 2: trials subdivided by dopamine dose (low compared to moderate)

All-cause mortality

Seven trials administered low-dose dopamine (i.e. < 4 $\mu g \cdot kg^{-1} \cdot min^{-1}$) and seven trials a moderate dose (4 to 10 $\mu g \cdot kg^{-1} \cdot min^{-1}$). Trials that studied low-dose dopamine in patients with heart-failure targeted to increase diuresis by improving renal perfusion, whereas low-dose dopamine during cardiac surgery was used to preserve renal function. Moderate dose-dopamine was administered in both patients with heart-failure and cardiac surgery patients to increase renal perfusion and ameliorate cardiac function. One trial that reported mortality did not report on the dopamine dose.[30] No statistically significant associations between different doses of dopamine and mortality were found (Table 2).

SAEs

The occurrence of SAEs was recorded in three trials that administered low-dose dopamine and in four trials administering moderate-dose dopamine. No significant differences were found for either low or moderate dose dopamine (Table 2).

Other outcomes

In the low-dose dopamine group there was significant heterogeneity (I^2 90%, P=0.002) due to one trial reporting use of renal replacement therapy in 36 of the 40 patients (90%) in the control group versus 2 of the 42 patients (5%) in the dopamine group. No significant differences were observed for any dose on any of the outcomes (Table 2).

Comparison 3: trials subdivided by setting (heart failure compared to cardiac surgery)

All-cause mortality

Ten trials were conducted in patients admitted with acute heart failure and seven trials in patients undergoing cardiac surgery. Heart failure was often based on clinical symptoms classified by the New York Heart Association (NYHA) and a depressed LVEF (Table S1). The type of cardiac surgery varied between the trials: two trials included patients having cardiac artery bypass grafting,[34,36] two trials included patients having mitral valve surgery,[35,37] and three trials included patients having various cardiac surgeries.[31-33] Subgroup analyses by clinical setting did not show any statistically significant associations on mortality (Table 2).

SAEs

SAEs were reported in four trials that included patients with acute heart failure and in two trials that included patients undergoing cardiac surgery. There were no statistically significant associations on occurrence of SAEs in both settings (Table 2).

Other outcomes

There was no significant difference in favour of any intervention on the proportion of myocardial infarction, renal replacement therapy, and ventricular or atrial tachyarrhythmias (Table 2).

Post-hoc meta-analyses with broader inclusion criteria of cardiac dysfunction

These post-hoc meta-analyses included trials in which a substantial proportion of patients (> 25%) were assumed to have cardiac dysfunction. This broader inclusion criterion added ten trials with patients suffering from shock (n = 1679) or septic shock (n =444), who received high-dose dopamine for treatment of hypotension. This meta-analysis included 40 trials with 4182 patients and full details can be found in Supplements 2.

Dopamine seemed associated with increased mortality, increased SAEs, and increased tachyarrhythmias when compared with a potentially active control intervention (Table S2). The excess mortality was largely attributable to the trials which administered high-dose dopamine and accounted for 87% of weight in the pooled effect (Figure S3). All but one of these trials compared dopamine with noradrenaline and two trials allowed other cardioactive co-interventions with dobutamine or open-label noradrenaline. TSA including all trials reporting on mortality showed that it is highly unlikely to show a beneficial effect of dopamine with further trials, as the cumulative Z-curve would have to cross the futility area (Figure S4).

Observational studies

One quasi-randomised study and one observational study were assessed for harms.[23,38] One study compared dopamine to levosimendan and recorded SAEs and arrhythmias[38]; the other evaluated dopamine to an intra-aortic balloon pump and reported myocardial infarction and renal replacement therapy proportions.[23] Dopamine did not significantly affect any of these outcomes (Table S3).

TABLE 3. GRADEpro summary of finding table of the outcomes of interest

Quality assessment							№ of patients		Effect		Quality	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Dopamine	Any (in) active comparator	Relative (95% CI)	Absolute (95% CI)		
Mortality at maximum follow-up												
15	RCTs	serious[a]	not serious	serious[b]	serious[c]	none	60/457 (13.1%)	90/581 (15.5%)	**RR 0.91** (0.68 to 1.21)	**14 fewer per 1.000** (from 33 more to 50 fewer)	⨁◯◯◯ VERY LOW	CRITICAL
Serious adverse events												
6	RCTs	serious[a]	not serious	serious[b]	serious[c]	none	62/268 (23.1%)	51/314 (16.2%)	**RR 1.20** (0.91 to 1.57)	**32 more per 1.000** (from 15 fewer to 93 more)	⨁◯◯◯ VERY LOW	CRITICAL
Myocardial infarction												
5	RCTs	serious[a]	not serious	serious[b]	very serious[d]	none	9/139 (6.5%)	7/200 (3.5%)	**OR 1.63** (0.56 to 4.71)	**21 more per 1.000** (from 15 fewer to 111 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Ventricular tachyarrhythmias												
8	RCTs	serious[a]	not serious	serious[b]	very serious[d]	none	14/255 (5.5%)	10/313 (3.2%)	**OR 1.46** (0.52 to 4.10)	**14 more per 1.000** (from 15 fewer to 87 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Renal replacement therapy												
4	RCTs	serious[a]	serious[e]	serious[b]	serious[c]	none	9/174 (5.2%)	42/197 (21.3%)	**RR 0.44** (0.07 to 2.75)	**119 fewer per 1.000** (from 198 fewer to 373 more)	⨁◯◯◯ VERY LOW	IMPORTANT
Atrial tachyarrhythmias												
2	RCTs	serious[a]	not serious	serious[b]	very serious[d]	none	1/66 (1.5%)	2/115 (1.7%)	**OR 1.16** (0.14 to 9.65)	**3 more per 1.000** (from 15 fewer to 128 more)	⨁◯◯◯ VERY LOW	NOT IMPORTANT

Explanations: a. There were no trials at overall low risk of bias; b. There was considerable difference in population types (i.e. heart failure, cardiac surgery) and both dosing and length of administration of the study drugs; c. The confidence intervals include both appreciable harm and benefit and less than 5% of the required information size was accrued; d. Odds ratios are based on very few events (< 25); e. There was considerable statistical heterogeneity (I^2 77%, P=0.004), which was caused by one study at high risk of bias. Abbreviations: RCTs, randomised clinical trials; CI, confidence interval; RR, risk ratio; OR, odds ratio.

Quality of evidence

Based on GRADE, the certainty of the evidence on all outcomes was judged as 'very low' and was mainly attributable to serious risks of bias, serious indirectness, and very serious imprecision (Table 3). The Manhattan error matrix plots showed that there are lacunas in the evidence of dopamine regarding both systematic errors and random errors (Figure S5). The funnel plots showed no clear arguments for small trial bias including publication bias (Figure S6).

Discussion

Our main meta-analysis consisting of 17 trials with 1218 patients did not provide high-quality evidence to support or refute the use of dopamine. All trials were at overall high risk of bias, only one trial compared dopamine with placebo, and TSA showed that further thousands of patients need to be randomised before firm conclusions can be drawn. The use of dopamine as preferred inotrope in up to 25% of heart failure patients lacks evidence from RCTs.

The largest trial on dopamine thus far observed that high-dose dopamine, as compared with noradrenaline, is associated with increased 28-day mortality in the subgroup of patients with cardiogenic shock.[8] We could not include these patients in our main meta-analysis because cardiac function was not measured in each patient and the randomisation procedure was not stratified for the cardiogenic shock subgroup. The increased mortality was supported by a meta-analysis including trials randomising patients with cardiogenic shock receiving high-dose dopamine.[39] We were unable to include these trials because the meta-analysis did not elaborate on cardiac function of each trial population and the full-text manuscripts were inaccessible to us (i.e. the Wanfang and Weipu Database). Based on these studies, high-dose dopamine for treatment of cardiogenic shock seems associated with increased harm.

Dopamine for treatment of cardiac dysfunction also seems harmful according to observational data.[11] Nevertheless, the quality of current evidence on the possible benefits or harms of dopamine, milrinone, levosimendan, and probably all other inotropes is considered very low. [40,41] There is currently no high-quality evidence on which inotrope should preferentially be administered to patients with cardiac dysfunction.

Previous systematic reviews on dopamine in critically ill adult patients differ in design; all studied dopamine in patients with cardiogenic,[39,42] hypotensive,[9] or septic shock.[43-47] Some identified a potentially harmful effect of dopamine on mortality and occurrence of arrhythmias,[39,43,44,46] while others were inconclusive.[9,42,45,47] These systematic reviews used different inclusion criteria and most studied high-

dose dopamine.[9,39,43-47] The main analysis of our systematic review included fewer patients (n = 1218) compared to eight of the other reviews (n = 510 [39], n = 70 [42], n = 1400 [9], n = 2043 [44], n = 1408 [43,47], n = 3819 [45], n = 1718 [46]) due to our more stringent inclusion criteria on cardiac dysfunction. We selected patients with objectively measured cardiac dysfunction because these patients would presumably benefit the most from an inotropic drug based on a pathophysiological reasoning. Critically ill patients with a normal cardiac function probably benefit less from the inotropic effects of dopamine and are more likely to only suffer potential harms.

Limitations and strengths

Potential biases may have arisen during the review process. Our systematic review mainly included small trials (i.e. less than 100 patients per trial) that used haemodynamic variables as their primary outcome. Therefore, our effect estimates may contain covariate imbalances and the included trials were individually underpowered for our outcomes.[48] Such problems with imbalance and power are, however, best mitigated through the conduct of meta-analyses.

It can be debated whether our inclusion criteria fully reflect daily clinical practise. We were interested in patients with cardiac dysfunction based on cardiac index and LVEF measurements, which are operator dependent and may have considerable interobserver variability.[49,50] Though, these are currently the advocated measures to quantify left-ventricular function and often used as trigger to start inotropic treatments.[51]

Although statistical heterogeneity was often absent, our meta-analyses had considerable clinical heterogeneity because 1) not all trials included patients who all have objectively verified cardiac dysfunction and 2) dopamine was administered in different doses to patients in different clinical settings, based on different assumed pathophysiological mechanisms. In fact, very few of the included trials had objective haemodynamic targets to direct infusion of dopamine and other inotropes. We probably cannot move forward understanding the role of inotropes before we understand the pathophysiology of shock on organ level.

More insight is needed into the pathophysiology of shock on organ level with bridging to haemodynamic goals to achieve optimal organ function support in critical ill patients. To detect possible sources of clinical heterogeneity, we first conducted subgroup analyses on dopamine dose, clinical setting, and a sensitivity analysis of trials exclusively including patients with cardiac dysfunction. Second, we conducted post-hoc meta-analyses with a broader inclusion criterion for cardiac dysfunction.

Conclusion

Evidence for dopamine in critically ill adults with cardiac dysfunction is sparse and of low quality due to high risks of systematic errors and random errors. The use of dopamine in patients with cardiac dysfunction can neither be recommended nor refuted.

References

1. Bangash MN, Kong ML, Pearse RM. Use of inotropes and vasopressor agents in critically ill patients. *Br J Pharmacol*. 2012;165(7):2015-2033.
2. Jentzer JC, Coons JC, Link CB, Schmidhofer M. Pharmacotherapy update on the use of vasopressors and inotropes in the intensive care unit. *J Cardiovasc Pharmacol Ther*. 2015;20(3):249-260.
3. Yancy CW, Jessup M, Bozkurt B, et al. 2013 ACCF/AHA guideline for the management of heart failure: A report of the american college of cardiology foundation/american heart association task force on practice guidelines. *Circulation*. 2013;128(16):240.
4. Ponikowski P, Voors AA, Anker SD, et al. 2016 ESC guidelines for the diagnosis and treatment of acute and chronic heart failure: The task force for the diagnosis and treatment of acute and chronic heart failure of the european society of cardiology (ESC)developed with the special contribution of the heart failure association (HFA) of the ESC. *Eur Heart J*. 2016;37(27):2129-2200.
5. Giamouzis G, Butler J, Starling RC, et al. Impact of dopamine infusion on renal function in hospitalized heart failure patients: Results of the dopamine in acute decompensated heart failure (DAD-HF) trial. *J Card Fail*. 2010;16(12):922-930.
6. Triposkiadis FK, Butler J, Karayannis G, et al. Efficacy and safety of high dose versus low dose furosemide with or without dopamine infusion: The dopamine in acute decompensated heart failure II (DAD-HF II) trial. *Int J Cardiol*. 2014;172(1):115-121.
7. Chen HH, Anstrom KJ, Givertz MM, et al. Low-dose dopamine or low-dose nesiritide in acute heart failure with renal dysfunction: The ROSE acute heart failure randomized trial. *JAMA*. 2013;310(23):2533-2543.
8. De Backer D, Biston P, Devriendt J, et al. Comparison of dopamine and norepinephrine in the treatment of shock. *N Engl J Med*. 2010;362(9):779-789.
9. Gamper G, Havel C, Arrich J, et al. Vasopressors for hypotensive shock. *Cochrane Database Syst Rev*. 2016;2:CD003709.
10. Thongprayoon C, Cheungpasitporn W, Harrison AM, et al. Temporal trends in the utilization of vasopressors in intensive care units: An epidemiologic study. *BMC Pharmacol Toxicol*. 2016;17(1):z. doi: 10.1186/s40360-016-0063-z [doi].
11. Mebazaa A, Motiejunaite J, Gayat E, et al. Long-term safety of intravenous cardiovascular agents in acute heart failure: Results from the european society of cardiology heart failure long-term registry. *Eur J Heart Fail*. 2018;20(2):332-341. doi: 10.1002/ejhf.991 [doi].
12. da Costa BR, Juni P. Systematic reviews and meta-analyses of randomized trials: Principles and pitfalls. *Eur Heart J*. 2014;35(47):3336-3345. doi: ehu424 [pii].
13. van der Horst I, Keus F, Hiemstra B, Koster G, Wetterslev J, Gluud C. Dopamine for cardiac dysfunction in critically ill adult patients: A systematic review with meta-analysis and trial sequential analysis. *PROSPERO 2016: CRD42016042867. Available from http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42016042867*.
14. Higgins JP, Green S, eds. *Cochrane handbook for systematic reviews of interventions.* 5.1.0 (updated March 2011) ed. The Cochrane Collaboration, 2011. Available from www.cochrane-handbook.org.
15. Gluud C, Nikolova D, Klingenberg SL. Cochrane hepato-biliary group. About cochrane (cochrane review groups (CRGs)). . 2015(Issue 4):Art. No.: LIVER.
16. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *J Clin Epidemiol*. 2009;62(10):1006-1012.

17. Rao V, Ivanov J, Weisel RD, Ikonomidis JS, Christakis GT, David TE. Predictors of low cardiac output syndrome after coronary artery bypass. *J Thorac Cardiovasc Surg*. 1996;112(1):38-51.
18. International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use (ICH) adopts consolidated guideline on good clinical practice in the conduct of clinical trials on medicinal products for human use. *Int Dig Health Legis*. 1997;48(2):231-234.
19. Keus F, Wetterslev J, Gluud C, van Laarhoven CJ. Evidence at a glance: Error matrix approach for overviewing available evidence. *BMC Med Res Methodol*. 2010;10:90.
20. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Stat Med*. 2002;21(11):1539-1558.
21. Wetterslev J, Jakobsen JC, Gluud C. Trial sequential analysis in systematic reviews with meta-analysis. *BMC Med Res Methodol*. 2017;17(1):7.
22. Guyatt GH, Oxman AD, Kunz R, et al. What is "quality of evidence" and why is it important to clinicians? *BMJ*. 2008;336(7651):995-998.
23. Mao CT, Chen TH, Chen DY, Tsai ML, Lin YS, Wang SH. Comparison of intra-aortic balloon pump and dopamine for patients with acute myocardial infarction and unstable haemodynamics undergoing percutaneous coronary intervention. *EuroIntervention*. 2015.
24. Kamiya M, Sato N, Nozaki A, et al. Renal effects of added low-dose dopamine in acute heart failure patients with diuretic resistance to natriuretic peptide. *J Cardiovasc Pharmacol*. 2015;65(3):282-288. doi: 10.1097/FJC.0000000000000193 [doi].
25. Varriale P, Mossavi A. The benefit of low-dose dopamine during vigorous diuresis for congestive heart failure associated with renal insufficiency: Does it protect renal function? *Clin Cardiol*. 1997;20(7):627-630.
26. Shah RA, Subban V, Lakshmanan A, et al. A prospective, randomized study to evaluate the efficacy of various diuretic strategies in acute decompensated heart failure. *Indian Heart J*. 2014;66(3):309-316. doi: 10.1016/j.ihj.2014.03.006 [doi].
27. Arutiunov GP, Arutiunov AG, Volkova AL. Study evaluating the impact of a combination of inotropic support and heart rate control on prognosis and stabilization rate in patients with decompensated chronic heart failure (LEGION). *Ter Arkh*. 2010;82(3):47-52.
28. Hsueh CW, Lee WL, Chen CK, et al. Dopamine and dobutamine have different effects on heart rate variability in patients with congestive heart failure. *Zhonghua Yi Xue Za Zhi (Taipei)*. 1998;61(4):199-209.
29. Cotter G, Weissgarten J, Metzkor E, et al. Increased toxicity of high-dose furosemide versus low-dose dopamine in the treatment of refractory congestive heart failure. *Clin Pharmacol Ther*. 1997;62(2):187-193. doi: S0009-9236(97)90067-9 [pii].
30. Sindone, A, MacDonald, P, K, A. Haemodynamic, neurohumoral and symptomatic effects of dobutamine, dopamine and milrinone in severe heart failure. *Aust N Z J Med*. 1998;28:113. http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/712/CN-00320712/frame.html.
31. Sirivella S, Gielchinsky I, Parsonnet V. Mannitol, furosemide, and dopamine infusion in postoperative renal failure complicating cardiac surgery. *Ann Thorac Surg*. 2000;69(2):501-506. doi: S0003-4975(99)01298-9 [pii].
32. Costa P, Ottino GM, Matani A, et al. Low-dose dopamine during cardiopulmonary bypass in patients with renal dysfunction. *J Cardiothorac Anesth*. 1990;4(4):469-473.

33. Bove T, Landoni G, Calabro MG, et al. Renoprotective action of fenoldopam in high-risk patients undergoing cardiac surgery: A prospective, double-blind, randomized clinical trial. *Circulation*. 2005;111(24):3230-3235. doi: 111/24/3230 [pii].

34. Rosseel PM, Santman FW, Bouter H, Dott CS. Postcardiac surgery low cardiac output syndrome: Dopexamine or dopamine? *Intensive Care Med*. 1997;23(9):962-968.

35. Hausen B, Heublein B, Vogelpohl J, von der Leyen H, Haverich A. Comparison of enoximone and piroximone in patients after mitral valve operation: A prospective and controlled clinical study. *J Cardiovasc Pharmacol*. 1992;19(3):299-307.

36. Oppizzi M, Montorsi E, Tosoni A, et al. The effectiveness of enoximone in patients with serious left ventricular dysfunction submitted for aorto-coronary bypass. *Minerva Anestesiol*. 1997;63(1-2):17-27.

37. Tarr TJ, Moore NA, Frazer RS, Shearer ES, Desmond MJ. Haemodynamic effects and comparison of enoximone, dobutamine and dopamine following mitral valve surgery. *Eur J Anaesthesiol Suppl*. 1993;8:15-24.

38. Dzhaiani NA, Kositsyna IV, Gnidkina NA, Tereshchenko SN. Efficacy of levosimendan vs dopamine in patients with resistant cardiac failure. *Ter Arkh*. 2011;83(6):53-59.

39. Rui Q, Jiang Y, Chen M, Zhang N, Yang H, Zhou Y. Dopamine versus norepinephrine in the treatment of cardiogenic shock: A PRISMA-compliant meta-analysis. *Medicine (Baltimore)*. 2017;96(43):e8402. doi: 10.1097/MD.0000000000008402 [doi].

40. Koster G, Bekema HJ, Wetterslev J, Gluud C, Keus F, van der Horst, I C. Milrinone for cardiac dysfunction in critically ill adult patients: A systematic review of randomised clinical trials with meta-analysis and trial sequential analysis. *Intensive Care Med*. 2016;42(9):1322-1335.

41. Koster G, Wetterslev J, Gluud C, et al. Effects of levosimendan for low cardiac output syndrome in critically ill patients: Systematic review with meta-analysis and trial sequential analysis. *Intensive Care Med*. 2015;41(2):203-221.

42. Schumann J, Henrich EC, Strobl H, et al. Inotropic agents and vasodilator strategies for the treatment of cardiogenic shock or low cardiac output syndrome. *Cochrane Database Syst Rev*. 2018;1:CD009669. doi: 10.1002/14651858.CD009669.pub3 [doi].

43. De Backer D, Aldecoa C, Njimi H, Vincent JL. Dopamine versus norepinephrine in the treatment of septic shock: A meta-analysis. *Crit Care Med*. 2012;40(3):725-730.

44. Vasu TS, Cavallazzi R, Hirani A, Kaplan G, Leiby B, Marik PE. Norepinephrine or dopamine for septic shock: Systematic review of randomized clinical trials. *J Intensive Care Med*. 2012;27(3):172-178.

45. Zhou F, Mao Z, Zeng X, et al. Vasopressors in septic shock: A systematic review and network meta-analysis. *Ther Clin Risk Manag*. 2015;11:1047-1059.

46. Avni T, Lador A, Lev S, Leibovici L, Paul M, Grossman A. Vasopressors for the treatment of septic shock: Systematic review and meta-analysis. *PLoS One*. 2015;10(8):e0129305.

47. Oba Y, Lone NA. Mortality benefit of vasopressor and inotropic agents in septic shock: A bayesian network meta-analysis of randomized controlled trials. *J Crit Care*. 2014;29(5):706-710.

48. Turner RM, Bird SM, Higgins JP. The impact of study size on meta-analyses: Examination of underpowered studies in cochrane reviews. *PLoS One*. 2013;8(3):e59202. doi: 10.1371/journal.pone.0059202 [doi].

49. Cole GD, Dhutia NM, Shun-Shin MJ, et al. Defining the real-world reproducibility of visual grading of left ventricular function and visual estimation of left ventricular ejection fraction: Impact of image quality, experience and accreditation. *Int J Cardiovasc Imaging*. 2015;31(7):1303-1314.

50. Pellikka PA, She L, Holly TA, et al. Variability in ejection fraction measured by echocardiography, gated single-photon emission computed tomography, and cardiac magnetic resonance in patients with coronary artery disease and left ventricular dysfunction. *JAMA Network Open*. 2018;1(4):e181456. doi: 10.1001/jamanetworkopen.2018.1456.

51. Lang RM, Badano LP, Mor-Avi V, et al. Recommendations for cardiac chamber quantification by echocardiography in adults: An update from the American society of echocardiography and the European association of cardiovascular imaging. *Eur Heart J Cardiovasc Imaging*. 2015;16(3):233-270.

3

Milrinone for cardiac dysfunction in critically ill adult patients: a systematic review of randomised clinical trials with meta-analysis and trial sequential analysis

Koster G, Bekema HJ, Wetterslev J, Gluud C, Keus F, van der Horst ICC.
Intensive Care Med. 2016 Sep;42(9):1322-35. doi: 10.1007/s00134-016-4449-6

Abstract

Background

Milrinone is an inotrope widely used for treatment of cardiac failure. Because previous meta-analyses had methodological flaws, we decided to conduct a systematic review of the effect of milrinone in critically ill adult patients with cardiac dysfunction.

Methods

This systematic review was performed according to The Cochrane Handbook for Systematic Reviews of Interventions. Searches were conducted until November 2015. Patients with cardiac dysfunction were included. The primary outcome was serious adverse events (SAE) including mortality at maximum follow-up. The risk of bias was evaluated and trial sequential analyses were conducted. The quality of evidence was assessed by the Grading of Recommendations Assessment, Development and Evaluation criteria.

Results

A total of 31 randomised clinical trials fulfilled the inclusion criteria, of which 16 provided data for our analyses. All trials were at high risk of bias, and none reported the primary composite outcome SAE. Fourteen trials with 1611 randomised patients reported mortality data at maximum follow-up (RR 0.96; 95% confidence interval 0.76-1.21). Milrinone did not significantly affect other patient-centred outcomes. All analyses displayed statistical and/or clinical heterogeneity of patients, interventions, comparators, outcomes, and/or settings and all featured missing data.

Conclusion

The current evidence on the use of milrinone in critically ill adult patients with cardiac dysfunction suffers from considerable risks of both bias and random error and demonstrates no benefits. The use of milrinone for the treatment of critically ill patients with cardiac dysfunction can be neither recommended nor refuted. Future randomised clinical trials need to be sufficiently large and designed to have low risk of bias.

Introduction

Milrinone (Corotrope®/Primacor®) is a type III phosphodiesterase inhibitor primarily used for inotropic support in the treatment of cardiac dysfunction. Although milrinone is implemented in several guidelines, its efficacy and safety profile remain controversial [1, 2].

Three meta-analyses have evaluated milrinone in critically ill patients [3–5]. One meta-analysis included adult cardiac surgery patients and observed that milrinone was associated with a significant increase in mortality while an update of the review found no significant effects [4, 5]. One other meta-analysis evaluated milrinone for the treatment of acute heart failure after acute myocardial infarction and suggested that milrinone might be safe and effective in these patients [3]. Unfortunately, only four trials with a limited number of 303 patients were included. None of these meta-analyses met all key methodological criteria for being a systematic review [6]. None of them were based upon a previously published protocol [3–5]. They lacked or had insufficient assessment of the risk of bias, and bias risks were insufficiently incorporated in the analyses and conclusions. They also lacked sufficient evaluation of the risks of random errors [7–9]. Just one domain having unclear risk of bias or high risk of bias is potentially sufficient to bias the findings. Furthermore, none of the previous meta-analyses assessed the outcomes according to the patients' perspective following the Grading of Recommendations Assessment, Development and Evaluation (GRADE) [10]. GRADE assesses the quality of evidence by evaluating risk of bias, heterogeneity, indirectness, imprecision and publication bias [10].

Our objective was to perform a systematic review with meta-analyses and trial sequential analysis (TSA) of randomised clinical trials (RCTs) according to The Cochrane Handbook for Systematic Reviews of Interventions and The Cochrane Hepato-Biliary Group Module comparing the benefits and harms of milrinone in critically ill adult patients with cardiac dysfunction [6, 7].

Methods

This systematic review was conducted according to our published protocol following the recommendations of The Cochrane Handbook for Systematic Reviews of Interventions and The Cochrane Hepato-Biliary Group Module and reported according to the PRISMA statement [6, 7, 11]. The protocol for this systematic review was registered at PROSPERO (no. CRD42014009061) [12].

Eligibility criteria

We considered all randomised clinical trials for inclusion, irrespective of language, blinding, publication status or sample size for assessment of benefits and harms. Quasi-randomised studies and observational studies with more than 500 patients were not included regarding assessment of benefits, but were considered for inclusion regarding assessment of harms and were planned to be analysed separately from the randomised trials [6].

Only trials with adult patients having cardiac dysfunction were considered. Cardiac dysfunction was defined as left ventricular ejection fraction (LVEF) below 40% and/or low cardiac output. Low cardiac output syndrome was defined as a pre-existing or developing state of cardiac insufficiency with underlying left or right ventricular systolic dysfunction requiring inotrope support [13]. We accepted the definitions of the diagnoses according to the criteria used in each individual randomised trial. Milrinone was considered the experimental intervention. There were no restrictions on dose, continuous or intermittent administration, or duration of treatment. However, trials with oral and/or inhaled milrinone were excluded as such routes of administration were judged inappropriate for critically ill patients.

All trials were included independent the type of control intervention, i.e., no intervention, placebo, dobutamine, levosimendan, or any other inotrope or vasopressor. While this may introduce heterogeneity, subgroup comparisons were preplanned according to inactive (placebo or no intervention) and potentially active control interventions (e.g., other inotropes or vasopressors).

All outcomes were graded according to the patients' perspective following GRADE [9]. The primary outcome was serious adverse events (SAE). SAE is a composite outcome summarising all serious events necessitating an intervention, operation, prolonged hospital stay or mortality according to ICH-GCP definitions [14]. This outcome was chosen for balancing the potential benefits and harms. The secondary outcomes were all-cause mortality, myocardial infarction, arrhythmia (including supra- and ventricular tachycardia and ventricular fibrillation) and duration of mechanical ventilation. Time-specific analyses of mortality were conducted according to availability of data (e.g. 30, 90 and/or 180 days). Length of stay (both intensive care unit and total hospital stay) is a potentially highly biased surrogate outcome for recovery and was therefore not considered.

Search strategy

We searched the Cochrane Central Register of Controlled Trials (CENTRAL) in The Cochrane Library, PubMed/MEDLINE, EMBASE, Web of Science and CINAHL until November 2015 (see supplements). We searched the references of the identified

trials and systematic reviews to identify any further relevant trials, i.e. backward snowballing. We also searched the WHO's trial platform and ClinicalTrials.gov for ongoing trials and contacted the FDA and EMA.

Study selection and data extraction

Two authors independently identified the trials for inclusion. Excluded studies were listed with reasons for exclusion. The following data was extracted: year of publication, country in which the trial was conducted, year of conduct of the trial, single-centre or multicentre trial, inclusion and exclusion criteria, all outcomes, details on interventions and characteristics of the trials, e.g. baseline imbalance, early stopping and other than intention-to-treat analysis. The authors of the individual trials were contacted in case of any unclear or missing information.

Bias risk assessment

Two authors independently assessed the risks of bias of the trials following instructions in The Cochrane Handbook for Systematic Reviews of Interventions [6]. The following risk of bias domains were extracted from each trial: random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessors, incomplete outcome data, selective outcome reporting and other bias including bias due to vested interest and/or academic bias [15–20]. Trials were classified as low risk of bias if all the domains were assessed as low risk. Trials were considered to have high risk of bias if one or more of these bias risk domains were scored as unclear or high risk of bias.

Error matrix approach

Data on the outcomes of all trials were assessed for the risks of bias (measured by the level of evidence), the risks of random error (measured by standard error) and design errors (measured by GRADING the outcomes) [21]. The three-dimensional Manhattan error matrix was used to facilitate the overview of available evidence at a glance [21].

Statistical analysis

We performed the meta-analyses according to The Cochrane Handbook for Systematic Reviews of Interventions [6] and The Cochrane Hepato-Biliary Group Module [7] and used the software package Review Manager 5.30 [22]. For TSA, the TSA program v.0.9beta (http://www.ctu.dk/tsa) was used [23].

Results were presented as relative risks (RR) with 95% confidence interval (CI) if there were two or more trials for an outcome. For rare events (<5% in the control

group) we calculated odds ratios (OR) and for very rare events (<2% in the control group) we calculated Peto's OR with 95% CI [24]. We also reported risk differences (RD) if conclusions were different from risk ratio. P values less than TSA-adjusted significance levels were considered statistically significant.

We calculated both a fixed-effect [25] and a random-effects [26] model for meta-analysis and presented both models in case of discrepancy. Considering the anticipated clinical heterogeneity we emphasised the random-effects model except if one or two trials dominated the available evidence [27]. Heterogeneity was explored by the Chi-squared test with significance set at a P value of 0.10, and the quantity was measured by I^2 [6, 28].

Analyses were performed on intention-to-treat [6]. In case of statistically significant RR, we calculated the number needed to treat (NNT) or number needed to harm (NNH) with 95% CI.

Predefined subgroup analyses were conducted according to (1) the bias risk of trials (low risk of bias compared to trials with unclear and high risk of bias; hypothesis: trials with unclear or high risk of bias are associated with more favourable beneficial effects); (2) the control intervention (inactive compared to potentially active; hypothesis: milrinone appears more favourable when compared to an inactive control intervention than potentially active control intervention); (3) clinical setting (patients having cardiac surgery compared to patients not having cardiac surgery; hypothesis: milrinone shows benefit in patients having cardiac surgery and not in other patients).

Funnel plots were used to explore small trial bias when data of more than ten randomised trials were available [6, 29, 30].

Trial sequential analysis (TSA)

We conducted TSA to control the statistically significance levels when data are reanalysed repetitively or are too sparse to draw firm conclusions, and accordingly, appropriately widen the confidence intervals [8, 9, 31–33]. TSA depends on the quantification of the required information size (the meta-analysis sample size). We calculated the diversity (D2)-adjusted required information size (DARIS) for a random-effects meta-analysis [34]. Trial sequential monitoring boundaries cannot be calculated when less than 5% of the DARIS has been accrued. We conducted TSA with the intention to maintain an overall 5% risk of a type I error and a power of 90%. We used the unweighted control event proportion in the control group and we anticipated an intervention effect of a 10% relative risk reduction (RRR). Sensitivity analyses were conducted using an RRR of 20% as well as the lower confidence limit

of the RRR of the intervention effect suggested by the meta-analysis of the trials with low risk of bias [27]. We intended to provide the CI adjusted for sparse data and repetitive testing, which we describe as the TSA-adjusted CI.

GRADE approach

We used the GRADE system to assess the quality of the body of evidence associated with each of the major outcomes in our review using GRADE software (ims.cochrane.org/revman/other-resources/gradepro) [10]. The quality measure of a body of evidence considers within-study risk of bias, indirectness, heterogeneity, imprecision and risk of publication bias.

Results

The search strategy identified 9336 hits (Fig. 1). Three additional publications were identified by backward snowballing: two could be included [35, 36] and one was irretrievable [37]. After removal of duplicates and screening, 244 hits remained. Of the 244 hits, 213 were excluded after full text evaluation. The remaining 31 publications were included in this systematic review. All authors of the 31 publications were contacted for missing data; only three authors responded [38–40], but no additional data was obtained. Of the 31 publications, 15 evaluated only surrogate outcomes (such as haemodynamic variables). Accordingly, only 16 randomised trials provided data for analyses [35, 36, 38, 40–52].

No ongoing trials, quasi-randomised studies or observational studies were identified.

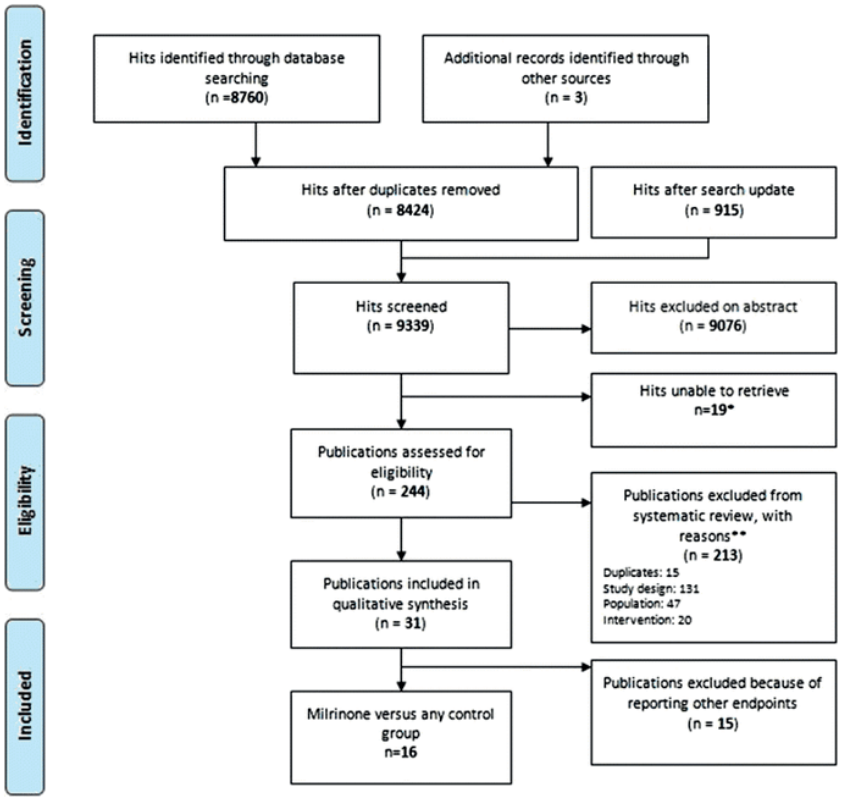


FIGURE 1. PRISMA flow diagram.
Legend. Asterisk not available at Dutch libraries or the universities linked through the University of Groningen. Double asterisk study design: no RCT or prospective non-randomised <500 patients. Population: no cardiac dysfunction or low cardiac output syndrome. Intervention: not milrinone

Characteristics of the included trials

The characteristics of the 16 randomised trials that provided data for analyses are listed (Table 1). Two trials used a three-arm parallel group design; all others had a two-arm parallel group design. There were five multicentre trials.

Eight trials evaluated patients after cardiac surgery [38, 41, 43, 45–48, 51], four trials evaluated patients with chronic heart failure [40, 42, 44, 50], three trials evaluated patients with acute heart failure after acute myocardial infarction [35, 36, 49] and one trial evaluated patients with severe sepsis [52].

Milrinone was administered in different doses. Nine trials used a 50 μg/kg bolus and one trial a 30 μg/kg bolus. Continuous infusion rates ranged from

0.25 to 1.0 μg/kg/min. Eight trials used an inactive comparator and eight trials used a potentially active comparator, including catecholamines, dobutamine, levosimendan, nifedipine or nesirrtide. Many trials applied milrinone as an add-on intervention to standard care including other inotropes.

TABLE 1. Baseline characteristics of included trials

Trial	Patients	Clinical setting	Milrinone bolus and infusion rate	Comparator	Outcomes
Cardiac surgery setting					
Doolan (1997)	30	CHF, ONCAB and valve surgery	50 μg/kg; 0.5 μg/kg/min for minimum 4h	placebo (67% received open label milrinone)	Primary: none Secondary: arrhythmia Other: none Surrogate: failure of weaning from CPB, hemodynamic - and biochemical parameter
Feneck (2001)	120	AHF, cardiac surgery	50 μg/kg; 0.5 μg/kg/min to 0.75 μg/kg/min	dobutamine 10μg/kg/min up to 20μg/kg/min if clinically indicated	Primary: none Secondary: MI, arrhythmias, hypotension Other: adverse events, severe low output state Surrogate: hemodynamic parameters
Möllhoff (2002)	30	CHF, cardiac surgery	no bolus; 0.375 μg/kg/min for at least 24h (dobutamine or epinephrine added when needed; no data)	nifedipine 0.2μg/kg/min for at least 24h (dobutamine or epinephrine added when needed; no data)	Primary: mortality (1 year) Secondary: arrhythmias Other: hospital stay, angina pectoris at 1yr, NYHA at 1yr Surrogate: hemodynamic – and biochemical parameters
Al Shawaf (2006)	30	CHF, cardiac (valve) surgery	50 μg/kg; 0.3 to 0.5 μg/kg/min for 24h	levosimendan bolus 12μg/kg followed by 0.1 to 0.2μg/kg/min for 24h	Primary: mortality (48 hour) Secondary: MI, arrhythmias, mechanical ventilation Other: renal replacement therapy Surrogate: hemodynamic parameters
Brackbill (2007)	40	CHF, cardiac surgery	50μg/kg; 0.375 μg/kg/min for approximately 24h (20% received epinephrine; 35% dopamine)	nesiretide bolus 2μg/kg with infusion 0.01μg/kg/min for approx. 24h (30% received epinephrine; 40% dopamine)	Primary: mortality (30 days) Secondary: none Other: renal failure, length of ICU and hospital stay Surrogate: hemodynamic parameters

Trial	Patients	Clinical setting	Milrinone bolus and infusion rate	Comparator	Outcomes
De Hert (2007)	30	CHF, cardiac surgery	no bolus; 0.5 µg/kg/min (+ dobutamine 5µg/kg/min)	levosimendan 0.1µg/kg/min maximum 24h (+ dobutamine 5µg/kg/min)	Primary: mortality (30 days) Secondary: MI, arrhythmias, mechanical ventilation Other: length of ICU and hospital stay Surrogate: biochemical – and hemodynamic parameters
Jebeli (2010)	70	CHF, ONCAB	50µg/kg; 0.5µg/kg/min for 24h (>90% received dopamine)	placebo (>90% received dopamine)	Primary: mortality (until ICU discharge) Secondary: MI, arrhythmia, mechanical ventilation Other: myocardial ischemia, non-myocardial ischemic events, renal failure, duration of CPB, inotropic support and ICU stay Surrogate: biochemical parameters, need for IABP
Hadadzadeh (2013)	80	CHF, OPCAB	50 µg/kg; 0.5 µg/kg/min for 24h (47.5% received other 'inotropic support')	placebo (52.5% received other 'inotropic support')	Primary: mortality (until ICU discharge) Secondary: MI, arrhythmia, mechanical ventilation Other: myocardial ischemia, ICU stay, inotropic support duration, renal failure, non-myocardial ischemic events Surrogate: need for IABP, biochemical parameters
Non-cardiac surgery setting					
Biddle (1987)	79	CHF	50 or 75 µg/kg; 0.5 µg/kg/min up to 0.75 or 1.0 µg/kg/min	dobutamine 2.5 µg/kg/min, increased to max. 15.0 µg/kg/min	Primary: mortality (48 hour) Secondary: arrhythmias Other: adverse reactions Surrogate: hemodynamic parameters
Karlsberg (1996)	30	AHF post MI	50 µg/kg; First 3h uptitration from 0.5, 0.62 to 0.75 µg/kg/min and then 0.25 to 0.75 µg/kg/min	dobutamine 2.5 µg/kg/min titrated to a max. of 15µg/kg/min	Primary: mortality (24 hour) Secondary: arrhythmias Other: adverse events Surrogate: echocardiographic – and hemodynamic parameters

Trial	Patients	Clinical setting	Milrinone bolus and infusion rate	Comparator	Outcomes
Siostrzonek (2000)	20	CHF	no bolus; 0.5 µg/kg/min >24h (+ dobutamine)	conventional (+ dobutamine)	Primary: mortality (24 hour) Secondary: none Other: length of ICU stay Surrogate: days until complete weaning from catecholamine, hemodynamic - and biochemical parameters
Cuffe (2002)	949	AHF, CHF	no bolus; 0.5 µg/kg/min, titrated to 0.375-0.75 µg/kg/min and maintained for 48-72h (11.5% received dobutamine)	placebo (9.3% received dobutamine)	Primary: mortality (60 days) Secondary: SAE, MI, arrhythmias, hypotension Other: length of stay Surrogate: none
Aranda (2003)	36	CHF, awaiting Htx	no bolus; 0.25µg/kg/min, titrated 0.125µg/kg/min to max 0.75µg/kg/min (32% received dopamine)	dobutamine 2.5 µg/kg/min titrated 2.5 µg/kg/min to a max. 10 µg/kg/min (18% received dopamine)	Primary: mortality (until hospital discharge) Secondary: arrhythmias Other: Htx, inotrope (bridge to Htx), LVAD (bridge to Htx), IABP (bridge to Htx), length of stay Surrogate: switched to alternate drug, hemodynamic parameters, cost effectiveness
Yang (2007)	120	AHF post MI	no bolus; 0.5 µg /kg/min for 5 hr once a day for 7 days	control	Primary: mortality (7 days) Secondary: none Other: none Surrogate: echocardiographic -, hemodynamic - and biochemical parameters
Pang (2011)	50	AHF post MI	50 µg/kg; 0.5 µg/kg/min (+ dobutamine)	placebo (+ dobutamine)	Primary: mortality (5 days) Secondary: none Other: none Surrogate: gastrointestinal reaction, hemodynamic parameters
Wang (2015) *	60	Sepsis	30 µg/kg; 0.375 to 0.5 µg/kg/min	control	Primary: mortality (28 days) Secondary: none Other: adverse events, ICU/hospital stay Surrogate: hemodynamic - and biochemical parameters

Legend. AHF = acute heart failure, CHF = chronic heart failure, CPB = cardiac pulmonary bypass, Htx = heart transplant, IABP = intra-arterial balloon pump, ICU = intensive care unit, LVAD = left ventricular assist device, MI = myocardial infarction, ONCAB = on-pump coronary artery bypass graft; OPCAB = off-pump coronary artery bypass graft, SAE = serious adverse events. The maximal length of follow-up for the outcome mortality is reported between the quotes in the 'Outcomes' column.

* Only the comparison between milrinone and control was included

Bias risk assessment

Three trials (19%) had low risk of bias regarding sequence generation, two trials (13%) had low risk of bias regarding allocation concealment, five trials (31%) had low risk of bias regarding blinding of participants, six trials (38%) had low risk of bias regarding blinding of outcome assessors, five trials (31%) had low risk of bias regarding incomplete outcome data, four trials (25%) were without selective outcome reporting and two trials (13%) were assessed as low risk of bias concerning industry and/or academic bias (Fig. 2). Accordingly, all trials were assessed as high risk of bias.

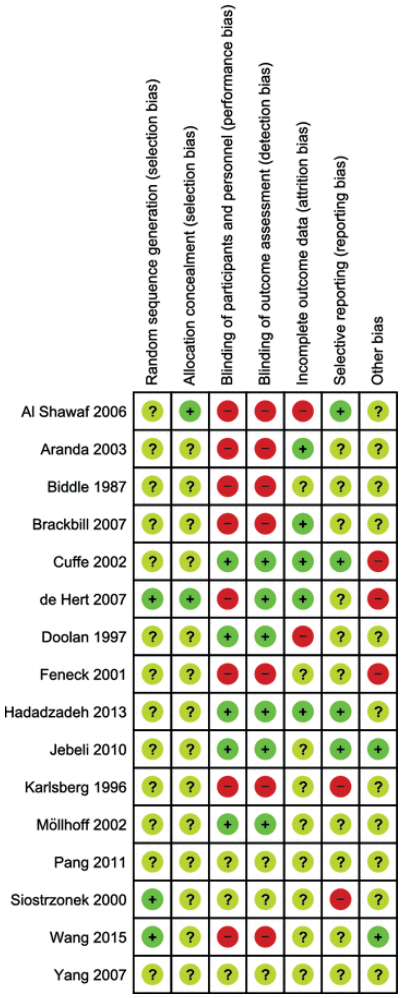


FIGURE 2. Risk of bias assessment. Review of authors' judgements about each risk of bias domain for each included study. Red high risk, green low risk, yellow unclear

Outcomes

The pooled intervention effect estimates with the 95% CI of the outcomes are specified according to control intervention and setting (Table 2 and supplements).

TABLE 2. Conventional risk ratios with 95% confidence intervals (CI) for the evaluated outcome measures including all patients stratified by intervention

	Number of trials	Number of patients	Conventional meta-analysis	
			RR with 95% CI	Test of interaction
Mortality				
Inactive control	5	1267	0.99 (0.77 to 1.27)	
Potentially active control	9	344	0.76 (0.31 to 1.89)	
Any control	14	1611	0.96 (0.76 to 1.21)	P=0.59
MI				
Inactive control	3	1060	0.54 (0.11 to 2.69)	
Potentially active control	2	60	1.09 (0.36 to 3.29)	
Any control	5	1120	0.73 (0.25 to 2.09)	P=0.48
VT/VF				
Inactive control	4	1087	0.80 (0.40 to 1.61)	
Potentially active control	3	139	1.11 (0.58 to 2.15)	
Any control	7	1226	0.96 (0.65 to 1.41)	P=0.50
SVT				
Inactive control	2	988	1.43 (0.80 to 2.54)	
Potentially active control	2	150	0.60 (0.13 to 2.70)	
Any control	4	1138	0.89 (0.43 to 1.87)	P=0.29
MV duration*				
Inactive control	2	150	-2.85 (-5.00 to -0.69)	
Potentially active control	2	60	12.66 (-3.48 to 28.80)	
Any control	4	210	1.03 (-4.87 to 6.93)	P=0.06

Legend**.** All pooled estimates are reported using risk ratio and calculated using a random-effects model unless stated otherwise. MV duration is reported in mean difference with 95% CI.
TSA adjusted risk ratios with the predefined α=0.05 (two sided), β=0.10 (power 90%), and an anticipated relative risk increase of 10% could not be calculated in any outcome with <5% of the DARIS accrued. DARIS = diversity adjusted required information size; MI = myocardial infarction; SVT: supraventricular tachyarrhythmia; VT/VF = ventricular tachyarrhythmia; MV = mechanical ventilation.

In the absence of trials that reported the primary composite outcome SAE including mortality, we have chosen to report all-cause mortality at maximum follow-up as the most important outcome. There were insufficient data for time-specific analyses of mortality. Meta-regression was not performed because of insufficient data.

Subgroup analyses according to risk of bias were not performed as no trial was assessed as having low risk of bias.

All analyses were conducted with stratification by control intervention, unless stated otherwise.

Comparison 1: all critically ill patients with cardiac dysfunction

All-cause mortality

Fourteen trials with 1611 randomised patients reported mortality. Pooled data showed that mortality at maximal follow-up was 11% in both groups (RR 0.96; 95% CI 0.76–1.21; I^2 0%; Fig. 3).

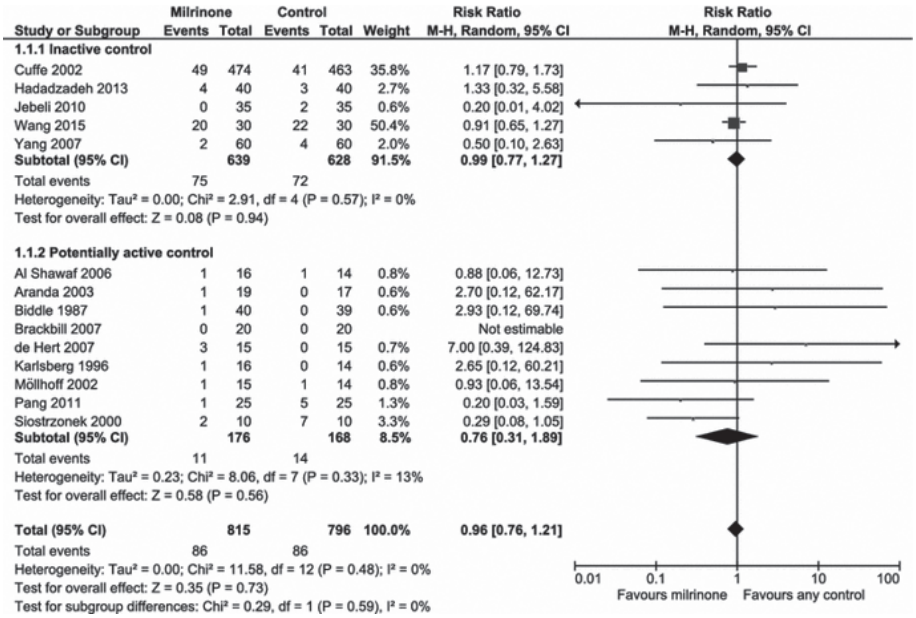


FIGURE 3. Forest plot of all-cause mortality in trials stratified by intervention.
Legend. Size of squares for risk ratio (RR) reflects the weight of the trial in the pooled analyses. Horizontal bars 95% confidence intervals (CI)

Subgroup analyses on type of control intervention and clinical setting showed differences in mortality event proportions in the control groups (inactive control group 0–70%; potentially active control group 9–73%; cardiac surgery setting control group 0–7%; non-cardiac surgery setting control group 0–73%), but tests of interaction showed no statistically significant differences between the groups (P = 0.59 and P = 0.83, respectively; Table 2 and supplements). No comparison could be analysed with TSA using the prespecified type I error of 5% and type II error of 10% because less than 5% of DARIS was accrued.

As a sensitivity analysis, we conducted TSA with an RRR of 20% and power of 80% which showed that 20% of the data was accrued and thousands of additional randomised patients are needed before futility or the required information size will be reached (RR 0.96; TSA adjusted CI 0.60–1.53; see supplements).

Myocardial infarction

Five trials with 1120 patients reported myocardial infarction (MI). MI at maximal follow-up occurred in 3% in the inactive control group versus 15% in the potentially active control group. Two small trials [41, 45] had a potentially active control group. There were no statistically significant differences in MI between milrinone and any control group (RR 0.73; 95% CI 0.25–2.09; I^2 61%, P = 0.48; Table 2). No comparison could be analysed with TSA using the prespecified type I error of 5% and type II error of 10% because less than 5% of DARIS was accrued.
Subgroup analyses based on clinical setting revealed discrepancy between fixed- and random-effects models driven by different weighting of one trial with 94% relative risk reduction (random-effects model RR 0.53; 95% CI 0.24–1.17, and fixed-effect model RR 0.45; 95% CI 0.25–0.81; I^2 34%; see supplements) [48].

Other outcomes

Ventricular tachyarrhythmias [i.e. ventricular tachycardia (VT)/ventricular fibrillation (VF)] were reported in seven randomised trials (1226 patients) with equal event rate percentages (7%) in both groups (RR 0.96; 95% CI 0.65–1.41; I^2 0%). No comparison could be analysed with TSA using the prespecified type I error of 5% and type II error of 10% because less than 5% of DARIS was accrued.

The pooled results and the subgroup analyses showed no associations between milrinone and ventricular tachyarrhythmia (see supplements).

Supraventricular tachyarrhythmia's (SVT) were reported in four trials (1138 patients). There was a statistically significant heterogeneity between the trials in both subgroup analyses (both I^2 55%; P = 0.08). SVT varied from 5 to 18% in the different subgroups (inactive versus potentially active and cardiac surgery versus

non-cardiac surgery). Analyses of the pooled data (RR 0.89; 95% CI 0.43–1.87) and the subgroups showed no significant associations (see supplements). No comparison could be analysed with TSA using the prespecified type I error of 5% and type II error of 10% since less than 5% of DARIS was accrued.

Mechanical ventilation duration was reported in four trials (210 patients) in a cardiac surgery setting; duration ranged from 11 to 34 h in the control group and 10–65 h in the milrinone group. There was statistically significant heterogeneity (I^2 80%; P = 0.002). No significant differences were found. Test of interaction was not significant (P = 0.06).

Comparison 2: patients with cardiac dysfunction after cardiac surgery

All-cause mortality

Six trials with 279 randomised patients reported mortality data. Mortality at maximal follow-up was 4% in both groups (RR 1.04; 95% CI 0.30–3.63; I^2 0%). No comparison could be analysed with TSA using the prespecified type I and type II error because less than 5% of DARIS was accrued.

Two trials used an inactive comparator and four trials used a potentially active comparator. No significant associations between milrinone and mortality were found (see supplements).

Myocardial infarction

MI was reported in four trials including 210 patients. There was significant statistical heterogeneity between the trials (I^2 58%; P = 0.09). There was discrepancy between the fixed- and the random-effects models driven by different weighting of one trial [48] (fixed-effect model RR 0.42; 95% CI 0.21–0.86; random-effects model RR 0.47; 95% CI 0.13–1.72; see supplements). No comparison could be analysed with TSA using the prespecified type I and type II error because less than 5% of DARIS was accrued.

Other outcomes

Five trials with 240 patients documented ventricular tachyarrhythmia and no significant associations were found (see supplements).

SVTs were reported in three trials with 230 randomised patients and no significant associations were found between milrinone and SVTs (see supplements).

Comparison 3: patients with cardiac dysfunction not having cardiac surgery

All-cause mortality

Eight trials with 1332 randomised patients reported mortality. Mortality at maximal follow-up was 11% in the milrinone group versus 12% in the control group (RR 0.91; 95% CI 0.64–1.28; I^2 0%). No comparison could be analysed with TSA using the prespecified type I error of 5% and type II error of 10% because less than 5% of DARIS was accrued.

Three trials used an inactive control and five trials used a potentially active control. Subgroup analyses on type of control intervention showed no significant difference (test of interaction P = 0.34). No significant associations between milrinone and mortality were found (see supplements).

Other outcomes

Ventricular tachyarrhythmia's (VT/VF) were reported in two trials (986 patients). No significant associations between milrinone and VT/VF were found (RR 1.19 95% CI 0.68–2.06).

There was insufficient data on other secondary outcomes.

Error matrix approach

The Manhattan error matrix plots of milrinone showed that there is a similar amount of evidence regarding the benefits and harms of milrinone. All trials had high risks of systematic errors (bias) and the large majority of the trials also had high risks of random errors (see supplements).

Small trial bias

Funnel plots showed no clear arguments for small trial bias including publication bias (see supplements).

GRADE approach

The quality of the evidence was assessed as very low for all outcomes based on risk of bias limitations, indirectness, inconsistency, imprecision and other considerations. Table 3 shows the GRADEpro summary of findings table with stratification by control intervention.

TABLE 3. GRADE pro summary of findings table of the outcomes of interest stratified by control intervention

Quality assessment						
№ of studies	**Study design**	**Risk of bias**	**Inconsistency**	**Indirectness**	**Imprecision**	**Other considerations**
Mortality at maximal follow-up						
14	randomised trials	serious [1]	not serious [2§]	serious [2,3,4]	very serious [5]	none [6]
Myocardial infarction						
5	randomised trials	serious [1]	serious [7]	serious [3,4]	very serious [5]	none
Ventricular arrhythmia						
7	randomised trials	serious [1]	not serious [8]	serious [3,4]	very serious [5]	none
Supraventricular arrhythmia						
4	randomised trials	serious [1]	serious [7]	serious [3,4]	very serious [5]	none
Mechanical ventilation duration						
4	randomised trials	serious [1]	serious [7]	very serious [3,4]	serious [5]	none

Legend. CI: Confidence interval; RR: Risk ratio; MD: Mean difference

1. No trial with low risk of bias in all domains; many bias assessment items are not reported in the trials
2. There was considerable clinical heterogeneity in setting (reflected by >5% difference in control event rate between populations) and comparator. All in all not enough to downgrade the evidence for inconsistency concerning the outcome mortality
3. There was considerable difference in dosing, titration and duration of milrinone; furthermore trials evaluated milrinone in a minority of trials against the most valuable inactive comparator (placebo); and even then different conventional inotropes were used between the trials.
4. Most trials used surrogate outcomes instead of patient-important outcomes
5. Many trials with few patients and few events; less than 5% of DARIS accrued
6. Funnel plots showed no clear asymmetry
7. There was considerable clinical and statistical heterogeneity
8. Despite different clinical settings and interventions the CI were overlapping and statistical heterogeneity was low

№ of patients		Effect		Quality	Importance
milrinone	any control	Relative (95% CI)	Absolute (95% CI)		
86/815 (10.6%)	86/796 (10.8%)	**RR 0.96** (0.76 to 1.21)	**4 fewer per 1000** (from 23 more to 26 fewer)	⨁◯◯◯ VERY LOW	
	6.9%		**3 fewer per 1000** (from 14 more to 17 fewer)		
19/568 (3.3%)	26/552 (4.7%)	**RR 0.73** (0.25 to 2.09)	**13 fewer per 1000** (from 35 fewer to 51 more)	⨁◯◯◯ VERY LOW	
	22.5%		**61 fewer per 1000** (from 169 fewer to 245 more)		
41/622 (6.6%)	42/604 (7.0%)	**RR 0.96** (0.65 to 1.41)	**3 fewer per 1000** (from 24 fewer to 29 more)	⨁◯◯◯ VERY LOW	
	5.1%		**2 fewer per 1000** (from 18 fewer to 21 more)		
38/577 (6.6%)	36/561 (6.4%)	**RR 0.89** (0.43 to 1.87)	**7 fewer per 1000** (from 37 fewer to 56 more)	⨁◯◯◯ VERY LOW	
	12.9%		**14 fewer per 1000** (from 74 fewer to 112 more)		
		-	MD **1.03 higher** (4.87 lower to 6.93 higher)	⨁◯◯◯ VERY LOW	

Discussion

Our systematic review evaluating the effects of milrinone for critically ill adult patients with cardiac dysfunction found few data on outcomes critical for decision making. Thirty-one randomised clinical trials fulfilled our inclusion criteria. All included trials had high risk of bias, most as a result of not reporting bias protection, and nearly all trials had large risks of random errors. Fifteen trials only reported surrogate outcomes. No trial reported the primary outcome, SAE (including mortality). All-cause mortality was reported in 14 trials with 1611 patients. No significant effect on any patient-centred outcome was found.

A general issue is that systematic reviews depend on the strengths of the included randomised trials. Trials with unclear or high risks of bias are associated with overestimation of benefits and underestimation of harms [15, 16, 18, 19]. The unknown true intervention effect may be beneficial, neutral or harmful. Previous meta-analyses on milrinone differ in design from our systematic review and they come to different conclusions [3–5]. One study focussed only on patients with myocardial infarction [3] and two only on cardiac surgery patients, in which the latter was an update [4, 5]. The meta-analysis on patients with myocardial infarction observed no significant effect on mortality, but stated that milrinone increased left ventricular ejection fraction and cardiac output [3]. The first meta-analysis evaluating patients having cardiac surgery suggested an increase in mortality using milrinone, which disappeared in the updated meta-analysis [4, 5]. Our prepublished protocol, a sensitive search strategy and thorough evaluations of the risks of systematic errors and random errors may explain differences with these previous publications [3–5]. First, previous meta-analyses ignored exploring associations of bias risk with intervention effect estimates. Final conclusions ought to be derived from trials with low risk of bias, of which there were none [6]. Second, despite including more patients (n = 1611) as compared to previous meta-analyses (n = 303 [3], n = 518 [4], n = 1037 [5]) the number of included patients is still far too small to draw any firm conclusions. We think that any significance needs the perspective of sample size considerations, in individual trials and also in meta-analyses [27, 31, 53–55]. Third, previous meta-analyses combined patients with normal cardiac functions [4, 5] and children [5] with patients with cardiac dysfunction into one pooled estimate. We included trials that randomised adult patients who had cardiac dysfunction. It is unlikely that patients benefit from milrinone when their cardiac function is unaffected, i.e. when the pathophysiological basis for cardiac stimulation is lacking.

Co-interventions with medications with an efficacy profile similar to milrinone might also have obscured results. Trials that evaluated milrinone versus placebo could also be considered add-on trials since co-interventions were allowed.

The largest trial that evaluated milrinone versus placebo allowed at least co-interventions with dobutamine in their randomised patients; other inotropes were not reported [44]. The results of this trial suggest that milrinone may be harmful in patients with heart failure (LVEF <40%) compared with standard treatment (ACE inhibitor and diuretics). Furthermore, the sickest patients were excluded in this trial [44]. For daily practice it is of utmost interest to know which vasopressor, inotrope, vasodilator or any combination is indicated for which patient and at what target [56, 57]. We found that for milrinone and levosimendan for critically ill adult patients with cardiac dysfunction evidence from trials with low risk of bias and low risk of random error is lacking to support its use [58]. Other interventions are currently being evaluated in systematic reviews which might feed future evidence-based guidance for clinicians or substantiate new trials.

Limitations

During the process of the systematic review we were non-adherent to our prepublished protocol for several reasons. We rephrased the title and terminology for an improved description of the cardiac state of the patients at interest (i.e. cardiac dysfunction instead of cardiac support or myocardial dysfunction). We divided subgroup comparisons into inactive versus (potentially) active control interventions. Since no data was found on the predefined subgroup comparison milrinone versus vasopressors we were unable to report this comparison. There were also no data on the composite outcome SAE (mortality included) and, therefore, all-cause mortality became the most important outcome. The outcome hypotension was regarded as a surrogate outcome and therefore omitted.

We frequently found significant statistical heterogeneity, but even when absent, there was still considerable clinical heterogeneity in patients, interventions, comparators, outcomes or settings. Pooling the data was frequently considered disputable, even in the absence of statistically significant tests of interactions. One example is the pooled intervention effect estimate of mortality (comparison 1), which has low statistical heterogeneity (I^2 0%; P = 0.50) and no subgroup differences, but the clinical heterogeneity is obvious, also reflected by control event rates for all-cause mortality varying from 3.6 to 12.0%. The large variety in types of control interventions further increases clinical heterogeneity.

Milrinone dose and duration varied among the included trials. Also, there were differences in definitions of outcomes. Further, 15 trials evaluated surrogate outcomes, such as haemodynamic and biochemical parameters. Finally, most trials had short follow-up; only one trial evaluated 1-year follow-up [51], so that mortality analyses reflect rather short follow-up.

Conclusions

The quantity and quality of evidence for benefit or harm of milrinone in critically ill adult patients with cardiac dysfunction are very low because of high risks of systematic and random errors. Future randomised clinical trials need to be large and well designed by following SPIRIT guidelines and reported according to CONSORT guidelines. The widespread use of milrinone in critical care cannot be advocated or refuted on the basis of the current evidence.

References

1. McMurray JJV, Adamopoulos S, Anker SD et al. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: the Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of th e European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. Eur Heart J 2012;33:1787–1847. doi:10.1093/eurheartj/ehs104
2. Yancy CW, Jessup M, Bozkurt B et al. 2013 ACCF/AHA guideline for the management of heart failure: executive summary. J Am Coll Cardiol 2013;62:1495–1539. doi:10.1016/j.jacc.2013.05.020
3. Tang X, Liu P, Li R et al. Milrinone for the treatment of acute heart failure after acute myocardial infarction: a systematic review and meta-analysis. Basic Clin Pharmacol Toxicol 2015;117:186–194. doi:10.1111/bcpt.12385
4. Zangrillo A, Biondi-Zoccai G, Ponschab M et al. Milrinone and mortality in adult cardiac surgery: a meta-analysis. J Cardiothorac Vasc Anesth 2012;26:70–77. doi:10.1053/j.jvca.2011.06.022
5. Majure DT, Greco T, Greco M et al. Meta-analysis of randomized trials of effect of milrinone on mortality in cardiac surgery: an update. J Cardiothorac Vasc Anesth 2013;27:220–229. doi:10.1053/j.jvca.2012.08.005
6. The Cochrane Collaboration. In: Higgins JPT, Green S (eds) Cochrane handbook for systematic reviews of interventions version 5.1.0. 2011. http://www.handbook.cochrane.org. Accessed 14 June 2016
7. Gluud C, Nikolova D, Klingenber SL et al. Cochrane Hepato-Biliary Group. About The Cochrane Collaboration (Cochrane Review Groups (CRGs)) 2015, issue 4. Art. No.:LIVER
8. Brok J, Thorlund K, Wetterslev J, Gluud C. Apparently conclusive meta-analyses may be inconclusive—trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. Int J Epidemiol 2009;38:287–298. doi:10.1093/ije/dyn188
9. Thorlund K, Devereaux PJ, Wetterslev J et al. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? Int J Epidemiol 2009;38:276–286. doi:10.1093/ije/dyn179
10. Guyatt GH, Oxman AD, Kunz R et al. What is "quality of evidence" and why is it important to clinicians? BMJ 2008;336:995–998. doi:10.1136/bmj.39490.551019.BE
11. Moher D, Liberati A, Tetzlaff J, Altman DG, For the PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009 Jul 21;339:b2535. doi:10.1136/bmj.b2535
12. Bekema HJ, Koster G, Wetterslev J et al. Milrinone for cardiac support: a systematic review of randomised clinical trials with meta-analyses and trial sequential analyses. In: PROSPERO 2014:CRD42014009061. http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014009061. Accessed 30 May 2016
13. Rao V, Ivanov J, Weisel RD et al. Predictors of low cardiac output syndrome after coronary artery bypass. J Thorac Cardiovasc Surg 1996;112:38–51
14. ICH Harmonised Tripartite Guideline. Clinical safety data management: definitions and standards for expedited reporting. Fed Regist 1995;60:11284–11287
15. Wood L, Egger M, Gluud LL et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 2008;336:601–605. doi:10.1136/bmj.39465.451748.AD

16. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273:408–412
17. Moher D, Jadad AR, Tugwell P. Assessing the quality of randomized controlled trials. Current issues and future directions. Int J Technol Assess Health Care 1996;12:195–208
18. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in metaanalyses. Ann Intern Med 2001;135:982–989
19. Savović J, Jones HE, Altman DG et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med 2012;157:429–438
20. Lundh A, Gøtzsche PC. Recommendations by Cochrane review groups for assessment of the risk of bias in studies. BMC Med Res Methodol 2008;8:22. doi:10.1186/1471-2288-8-22
21. Keus F, Wetterslev J, Gluud C, van Laarhoven CJHM. Evidence at a glance: error matrix approach for overviewing available evidence. BMC Med Res Methodol 2010;10:90. doi:10.1186/1471-2288-10-90
22. The Cochrane Collaboration. Review Manager (RevMan) [Computer Program] Version 5.3 Copenhagen. 2014.The Nordic Cochrane Centre
23. Thorlund K, Engstrøm J, Wetterslev J, Brok J, Imberger G, Gluud C. User manual for trial sequential analysis (TSA). 2011 http://ctu.dk/tsa/files/tsa_manual.pdf. Accessed 1 Jan 2016
24. Zhang J, Yu KF. What's the relative risk? A method of correcting the odds ratio in cohort studies of common outcomes. JAMA 1998;280:1690–1691
25. Demets DL. Methods for combining randomized clinical trials: strengths and limitations. Stat Med 1987;6:341–350
26. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials 1986;7:177–188
27. Jakobsen JC, Wetterslev J, Winkel P et al. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. BMC Med Res Methodol 2014;14:120. doi:10.1186/1471-2288-14-120
28. Higgins JPT, Thompson SG. Quantifying heterogeneity in a metaanalysis. Stat Med 2002;21:1539–1558
29. Egger M, Davey Smith G, Schneider M, Minder C. Bias in metaanalysis detected by a simple, graphical test. BMJ 1997;315:629–634
30. Macaskill P, Walter SD, Irwig L. A comparison of methods to detect publication bias in meta-analysis. Stat Med 2001;20:641–654
31. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative meta-analysis. J Clin Epidemiol 2008;61:64–75
32. Wetterslev J, Engstrøm J, Gluud C, Thorlund K. Trial sequential analysis: methods and software for cumulative meta-analyses. In: Chandl J, Clarke M, Higgins J (eds) Cochrane methods, Cochrane DB Syst Rev 2012;suppl 1, pp 1–56
33. Pogue JM, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. Control Clin Trials 1997;18:580–593 (discussion 661–6)
34. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model metaanalyses. BMC Med Res Methodol 2009;9:86. doi:10.1186/1471-2288-9-86

35. Pang Z, Zhao W. Milrinone in treatment of patients with cardiac shock following acute myocardial infarction. J Clin Med Pr 2011;15:82–83

36. Yang G, Han Y, Tong M, Deng J. Milrinone in treatment of patients with acute myocardial infarction and acute left heart failure. Chin Heart J 2007;19:454–456

37. Tang X, Li R, Jing Q et al. Observation on the clinical efficacy of milrinone in the treatment of heart failure patients with acute myocardial infarction. J Clin Med Pr 2012;16:163–165

38. Feneck RO, Sherry KM, Withington PS, Oduro-Dominah A. Comparison of the hemodynamic effects of milrinone with dobutamine in patients after cardiac surgery. J Cardiothorac Vasc Anesth 2001;15:306–315. doi:10.1053/jcan.2001.23274

39. Heringlake M, Wernerus M, Grünefeld J et al. The metabolic and renal effects of adrenaline and milrinone in patients with myocardial dysfunction after coronary artery bypass grafting. Crit Care 2007;11:R51. doi:10.1186/cc5904

40. Aranda JM, Schofield RS, Pauly DF et al. Comparison of dobutamine versus milrinone therapy in hospitalized patients awaiting cardiac transplantation: a prospective, randomized trial. Am Heart J 2003;145:324–329. doi:10.1067/mhj.2003.50

41. Al-Shawaf E, Ayed A, Vislocky I et al. Levosimendan or milrinone in the type 2 diabetic patient with low ejection fraction undergoing elective coronary artery surgery. J Cardiothorac Vasc Anesth 2006;20:353–357. doi:10.1053/j.jvca.2006.02.012

42. Biddle TL, Benotti JR, Creager Ma et al. Comparison of intravenous milrinone and dobutamine for congestive heart failure secondary to either ischemic or dilated cardiomyopathy. Am J Cardiol 1987;59:1345–1350. doi:10.1016/0002-9149(87)90917-9

43. Brackbill ML, Stam MD, Schuller-Williams RV, Dhavle AA. Perioperative nesiritide versus milrinone in high-risk coronary artery bypass graft patients. Ann Pharmacother 2007;41:427–432. doi:10.1345/aph.1H500

44. Cuffe MS, Califf RM, Adams KF et al. Short-term intravenous milrinone for acute exacerbation of chronic heart failure. JAMA 2002;287:1541–1547

45. De Hert SG, Lorsomradee S, Cromheecke S, Van der Linden PJ. The effects of levosimendan in cardiac surgery patients with poor left ventricular function. Anesth Analg 2007;104:766–773. doi:10.1213/01. ane.0000256863.92050.d3

46. Doolan LA, Jones EF, Kalman J et al. A placebo-controlled trial verifying the efficacy of milrinone in weaning high-risk patients from cardiopulmonary bypass. J Cardiothorac Vasc Anesth 1997;11:37–41. doi:10.1016/S1053-0770(97)90250-0

47. Hadadzadeh M, Hosseini SH, Mostafavi Pour Manshadi SM et al. Effect of Milrinone on short term outcome of patients with myocardial dysfunction undergoing off-pump coronary artery bypass graft: a randomized clinical trial. Acta Med Iran 2013;51(10):681–686

48. Jebeli M, Ghazinoor M, Mandegar MH et al. Effect of milrinone on short-term outcome of patients with myocardial dysfunction undergoing coronary artery bypass graft: a randomized controlled trial. Cardiol J 2010;17:73–78

49. Karlsberg RP, DeWood MA, DeMaria AN et al. Comparative efficacy of short-term intravenous infusions of milrinone and dobutamine in acute congestive heart failure following acute myocardial infarction. Milrinone–Dobutamine Study Group. Clin Cardiol 1996;19:21–30

50. Siostrzonek P, Koreny M, Delle-Karth G et al. Milrinone therapy in catecholamine-dependent critically ill patients with heart failure. Acta Anaesthesiol Scand 2000;44:403–409

51. Möllhoff T, Schmidt C, Van Aken H et al. Myocardial ischaemia in patients with impaired left ventricular function undergoing coronary artery bypass grafting—milrinone versus nifedipin. Eur J Anaesthesiol 2002;19:796–802

52. Wang Z, Wu Q, Nie X et al. Combination therapy with milrinone and esmolol for heart protection in patients with severe sepsis: a prospective, randomized trial. Clin Drug Investig 2015;35:707–716. doi:10.1007/s40261-015-0325-3

53. Pogue J, Yusuf S. Overcoming the limitations of current metaanalysis of randomised controlled trials. Lancet 1998;351:47–52

54. Imberger G, Gluud C, Boylan J, Wetterslev J. Systematic reviews of anesthesiologic interventions reported as statistically significant: problems with power, precision, and type 1 error protection. Anesth Analg 2015;121:1611–1622. doi:10.1213/ANE.0000000000000892

55. Roberts I, Ker K, Edwards P et al. The knowledge system underpinning healthcare is not fit for purpose and must change. BMJ 2015;350:h2463. doi:10.1136/bmj.h2463

56. Lamontagne F, Meade MO, Hébert PC et al. Canadian critical care trials group. Higher versus lower blood pressure targets for vasopressor therapy in shock: a multicentre pilot randomized controlled trial. Intensive Care Med 2016;42:542–550. doi:10.1007/s00134-016-4237-3

57. Mebazaa A, Tolppanen H, Mueller C et al. Acute heart failure and cardiogenic shock: a multidisciplinary practical guidance. Intensive Care Med 2016;42:147–163. doi:10.1007/s00134-015-4041-5

58. Koster G, Wetterslev J, Gluud C et al. Effects of levosimendan for low cardiac output syndrome in critically ill patients: systematic review with meta-analysis and trial sequential analysis. Intensive Care Med 2015;41(2):203–221. doi:10.1007/s00134-014-3604-1

4

Effects of levosimendan for low cardiac output syndrome in critically ill patients: systematic review with meta-analysis and trial sequential analysis

Koster G, Wetterslev J, Gluud C, Zijlstra JG, Scheeren TW, van der Horst IC, Keus F.
Intensive Care Med. 2015 Feb;41(2):203-21. doi: 10.1007/s00134-014-3604-1

Abstract

Background

To assess the benefits and harms of levosimendan for low cardiac output syndrome in critically ill patients.

Methods

We conducted a systematic review with meta-analyses and trial sequential analyses (TSA) of randomised clinical trials comparing levosimendan with any type of control. Two reviewers independently assessed studies for inclusion. The Cochrane Collaboration methodology was used. Random-effects risk ratios (RR) and 95% confidence intervals (CI) were derived for the principal primary outcome mortality at maximal follow-up.

Results

A total of 88 trials were included in the systematic review and 49 trials (6,688 patients) in the meta-analysis. One trial had low risk of bias and nine trials (2,490 patients) were considered lower risk of bias. Trials compared levosimendan with placebo, control interventions, and other inotropes. Pooling all trials including heterogenous populations was considered inappropriate. Pooled analysis of 30 trials including critically ill patients not having cardiac surgery showed an association between levosimendan and mortality (RR 0.83, TSA-adjusted 95% CI 0.59-0.97), while trials with lower risk of bias showed no significant difference (RR 0.83, TSA-adjusted 95% CI 0.48-1.55). Conventional meta-analysis of all 14 trials including cardiac surgery patients showed an association, while lower risk of bias trials showed no association between levosimendan and mortality (RR 0.52, 95% CI 0.37-0.73 versus RR 1.02, 95% CI 0.48-2.16).

Conclusion

The available evidence is inconclusive whether or not levosimendan may have a beneficial effect on mortality due to risks of systematic errors and random errors. Further well-designed randomised trials are needed.

Introduction

Levosimendan is an inotropic agent that may enhance myocardial contractility without increasing myocardial oxygen demand in patients with low cardiac output syndrome [1]. Clinical trials evaluating levosimendan have suggested survival benefits both in patients with acute heart failure and in patients after cardiac surgery [2, 3, 4, 5]. These favourable results were, however, not confirmed by larger clinical trials [6, 7].

Several meta-analyses, either based or not based on systematic reviews, have been performed [8, 9, 10, 11], and the most recent meta-analyses have stated that levosimendan was associated with survival benefits [9, 10].

Before one should accept levosimendan for patients with low cardiac output syndromes, one needs to secure that all evidence, i.e., the randomised clinical trials and the systematic reviews, are without design errors, systematic errors, and random errors [12]. First, the entire review process has to be outlined in advance by pre-published protocols [13]. Outcomes, subgroup analyses and an extensive sensitive search should be predefined. Risks of bias need to be assessed and incorporated into the analyses. Risks of random errors should be considered, especially in cumulative analyses of data from randomised clinical trials [14, 15, 16, 17, 18]. Evidence has to be graded according to The Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group [19]. Only then can reflections on heterogeneity, indirectness, and publication bias be considered, including variations in populations, therapies, and settings, among other reasons for heterogeneity [13].

Our objective was to assess the benefits and harms of levosimendan for low cardiac output syndrome in critically ill patients. Therefore, we performed a systematic review according to a prepublished protocol using The Cochrane Collaboration methodology with meta-analyses and trial sequential analyses of randomised clinical trials.

Methods

This systematic review was conducted following recommendations of the 'Cochrane Handbook for Systematic Reviews' [13] and reported according to the PRISMA statement (at: http://www.prisma-statement.org). The protocol is published in the PROSPERO register [20].

Eligibility criteria

Randomised clinical trials were considered for inclusion irrespective of language, publication status, and predefined outcomes. Levosimendan was considered the experimental intervention, without any restrictions regarding dose or duration of administration. Trials were included independent of the type of control intervention, i.e., placebo, no intervention, or any other inotrope or device. Only adult (age >18 years) critically ill patients with low cardiac output syndromes were included. Low cardiac output syndrome was defined as a pre-existing or developing state of cardiac insufficiency with underlying left or right ventricular systolic dysfunction requiring inotrope support [21]. Trials with oral administration of levosimendan may be appropriate for outpatient settings, but not for critical care practice and were therefore excluded. Other exclusion criteria were trials in animals, trials including patients without low cardiac output syndromes, quasi-randomised trials, observational studies, crossover trials, and trials comparing different treatment regimens of levosimendan.

Search strategy

We searched PubMed/Medline, Web of Science, EMBASE, Cumulative Index to Nursing and Allied Health Literature (CINAHL) and the Cochrane Central Register of Controlled Trials (CENTRAL) in The Cochrane Library. We also hand-searched the reference lists of included trials and other systematic reviews of levosimendan to identify additional trials. Unpublished trials were sought through trial registries (http://www.controlled-trials.com, http://www.clinicaltrials.gov, and http://www.centerwatch.com). No time restrictions were applied. The electronic literature search strategy was last updated 1 February 2014 (eTable 1).

Study selection and data extraction

Two authors independently reviewed all titles and abstracts and excluded trials that were obviously irrelevant (Fig. 1). All remaining trials were evaluated in full text or in abstract if no full text was available. Disagreements were resolved through discussion.

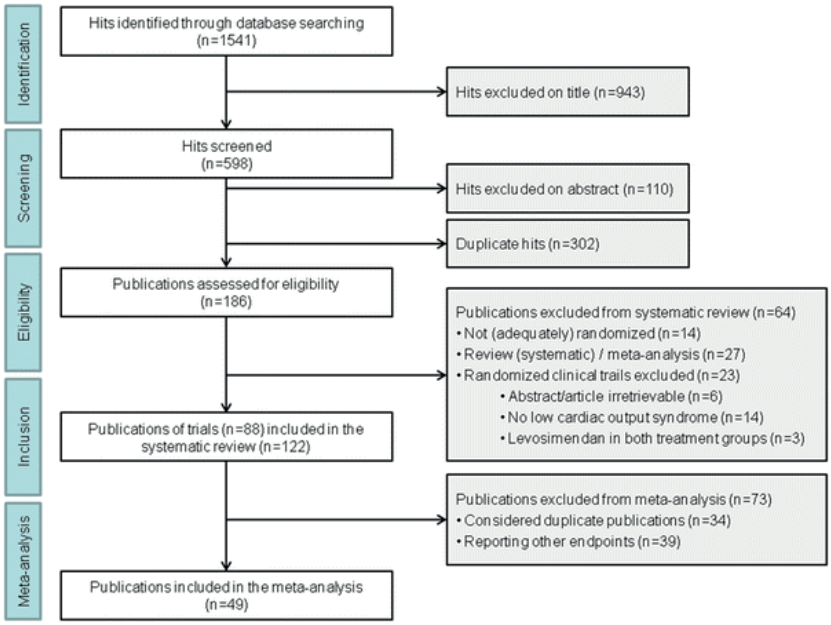


FIGURE 1. Flow diagram of identified trials

Information was extracted from the included trial reports including trial characteristics (single or multicentre, and country), characteristics of participants (disease severity), criteria for inclusion and exclusion, type of interventions, and outcomes. In cases of missing data, the corresponding authors were contacted.

The primary outcomes were all-cause mortality, serious adverse events (SAE) and myocardial infarction (MI) at maximal follow-up. Time-specific analyses of mortality were conducted according to availability of data (e.g., 30, 90, and/or 180 days). SAE was defined as the composite outcome measure summarising all serious events necessitating an intervention, operation, or prolonged hospital stay according to the International Conference on Harmonisation of Good Clinical Practice (ICH-GCP) definitions [22]; mortality was excluded in our definition of SAE to avoid double counts. MI was defined according to the individual trials.

The secondary outcomes were arrhythmias and hypotension. Data on arrhythmias were divided into ventricular tachyarrhythmias (VT) and supraventricular tachyarrhythmias (SVT). Hypotension was defined according to the individual trials. Length of stay of the index admission (both ICU and hospital stay) was not considered as it is a potentially highly biased surrogate outcome that does not consider re-admissions within a specific time frame [13]. All outcomes were graded from the patients' perspective according to GRADE (eTable 2) [19].

Risk of bias assessment

To evaluate the validity of the included trials, we assessed the risk of bias according to The Cochrane Collaboration methodology, including the domains of random sequence generation, allocation concealment, blinding of participants and personnel, blinding of outcome assessors, incomplete outcome data, selective outcome reporting, bias due to vested financial interest (funding), and academic bias. If one or more of the domains were judged as having a high or unclear risk of bias, we classified the trial as having a high risk of bias [13]. In cases of absence of trials with low risk of bias in all domains, we a priori formulated a group of trials with lower risk of bias if the trials had adequate assessments of sequence generation, allocation concealment, blinding of participants and personnel, as well as blinding of outcome assessment. The a priori defined group of trials with lower risk of bias may still suffer significant bias from other components.

Error matrix approach

All data on the outcomes of all trials were assessed for the risk of systematic errors, the risk of other design errors, and the risk of random errors. Data was presented in a three-dimensional Manhattan error matrix plot, which facilitates the overview of available evidence at a glance [12].

Statistical analysis

Review Manager 5.2.11 was used for meta-analyses [23]. For trial sequential analyses (TSA), the TSA program v.0.9 beta (http://www.ctu.dk/tsa) was used [24, 25]. For each included trial, we calculated the relative risk (RR) with 95% confidence intervals (CI) for dichotomous outcomes. We also reported risk differences (RD) if conclusions differed from RR. In cases of statistical significance, we calculated the number-needed-to-treat (NNT) or number-needed-to-harm (NNH) with 95% CI.

Heterogeneity among trials was explored by the Chi-squared test with significance set at a P value of 0.10 [26], and quantified with inconsistency factor (I^2). If the I^2 statistic was 0, we reported the results from a fixed-effect model; otherwise, we reported the results from the random-effects model anticipating abundant clinical heterogeneity (in populations, interventions, and settings) except if one or two trials dominated the available evidence. All reported results are from the random-effects model unless stated otherwise.

Predefined subgroup analyses were conducted: (1) according to the stratification of bias risk of trials (lower risk of bias compared to high risk of bias); (2) according to populations: patients having cardiac surgery or patients not having cardiac surgery; and (3) according to the control intervention (placebo or control intervention or any other inotrope).

Trial sequential analysis

We conducted TSA—an analysis that widens the confidence intervals and controls the p value when data are too sparse or reanalyzed to draw firm conclusions [25, 27, 28]. TSA is similar to interim analyses in a single trial in which sequential monitoring boundaries are used. In a similar manner, trial sequential monitoring boundaries can be applied to meta-analyses [14, 15, 16, 27]. TSA depends on the quantification of the required information size (the meta-analysis sample size). We calculated the diversity (D2) adjusted required information size (DIS) since the heterogeneity adjustment with I^2 tends to underestimate the required information size [29]. However, TSA could not be calculated when less than 5% of DIS was accrued. We conducted TSA with the intention to maintain an overall 5% risk of a type I error and a power of 90%. For the calculation of the required information size, we anticipated an intervention effect of a 10% relative risk reduction (RRR). We conducted sensitivity analyses using a RRR of 20%. The TSA were conducted using the unweighted control event proportion calculated from the actual meta-analyses. We provide the 95% CI adjusted for sparse data and repetitive testing, which we describe as the TSA-adjusted 95% CI.

Results

Our search strategy identified a total of 1,541 hits (Fig. 1). After screening, 186 publications remained of which another 64 were excluded (eTable 3 including references, and eTable 4). Finally, 122 publications described 88 randomised clinical trials. Of the 88 trials, 24 were only published as abstracts. Two publications (Russian and Spanish) were translated. All authors were contacted for missing data of whom 14 responded and 3 provided additional data. Thirty-nine trials evaluated surrogate outcomes, such as haemodynamic variables, without reporting any clinical outcome critical for decision making [e1–e39]. A total of 49 trials with 6,688 patients were included in the analyses [2, 3, 4, 5, 7, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73]. We identified 13 ongoing trials (eTable 5).

Characteristics of trials

Characteristics of the 49 trials are listed in Table 1. Of these, 44 trials used a two-arm parallel group design, 4 used a three-arm parallel group design, and 1 used a four-arm parallel group design. There were 33 single-centre trials and 16 multi-centre trials. Duration of follow-up varied from 6 h to 12 months (eTable 3).

TABLE 1. Characteristics of all included trials in the systematic review in critically ill patients with low cardiac output syndrome divided on basis of control intervention

Trial	Setting/country	Population	No. of randomised patients	Comparator
Adamopoulos [30]	Multicentre/Greece	AHF, CHF	69	Placebo or dobutamine
Asaad [31]	Single centre/Kuwait	CHF, CABG	20	Placebo
Dogan [32]	Single centre/Turkey	CHF, CABG	200	Placebo
Eriksson [33]	Multicentre/Finland	CHF, CABG	60	Placebo
Flevari [34]	Single centre/Greece	CHF	45	Placebo
Husebye [35]	Single centre/Norway	AHF post MI	61	Placebo
Iliuta [36]	Single centre/Romania	CHF, CABG	600	Placebo
Kleber [37]	Multicentre/Germany, Sweden	Right HF	28	Placebo
Lathinen [38]	Single centre/Finland	CHF, CABG and/or valve surgery	200	Placebo
Levin [4]	Multicentre/Argentina	CHF, valve surgery	77	Placebo
Levin [39]	Multicentre/Argentina	CHF	80	Placebo
Lilleberg [40]	Single centre/Finland	CHF	22	Placebo
Llorens [41]	Single centre/Spain	AHF	45	Placebo
Moiseyev [2]	Multicentre/Russia, Latvia	AHF post MI	504	Placebo
Nieminen [42]	Multicentre/Finland, Germany, Sweden, the Netherlands	CHF	116	Placebo, vehicle or dobutamine
Packer [7]	Multicentre/USA, Israel, Australia	AHF, CHF	700	Placebo
Slawsky [43]	Multicentre/USA	AHF, CHF	146	Placebo
Tritapepe [44]	Single centre/Italy	CHF, CABG	102	Placebo
Zairis [3]	Multicentre/Greece	CHF	227	Placebo or dobutamine
Berger [45]	single centre/Austria	CHF	75	Prostaglandin E1
Biteker [46]	Single centre/Turkey	CHF	24	Control
Kurt [47]	Single centre/Turkey	AHF, CHF	60	Control
Lomivorotov [48]	Single centre/Russia	CHF, CABG	90	Prophylactic IABP
Malfatto [49]	Single centre/Italy	CHF	33	Furosemide
Mavrogeni [50]	Single centre/Greece	CHF	50	Control
Zemljic [51]	Single centre/Slovenia	CHF	40	Control
Alhashemi [52]	Single centre/Saudi Arabia	Sepsis	42	Dobutamine
Al-Shawaf [53]	Single centre/Kuwait	CHF, valve surgery	30	Milrinone

Trial	Setting/country	Population	No. of randomised patients	Comparator
Alvarez [54]	Single centre/Spain	LCO after CABG and/or valve surgery	41	Dobutamine
Baysal [55]	Single centre/Turkey	CHF, CABG and/or valve surgery	128	Dobutamine
Bergh [56]	Multicentre/Sweden, Norway, Iceland	AHF, CHF	60	Dobutamine
Bonios [57]	Single centre/Greece	CHF	63	Dobutamine or dobutamine + levosimendan
DeHert [58]	Single centre/Belgium	CHF, CABG and/or valve surgery	30	Milrinone
Duygu [59]	Single centre/Turkey	AHF, CHF	62	Dobutamine
Duygu [60]	Single centre/Turkey	AHF, CHF	40	Dobutamine
Follath [61]	Multicentre/Austria, Denmark, Finland, France, Germany, Hungary, Italy, Switzerland, the Netherlands, Sweden, and the UK	CHF	203	Dobutamine
Fuhrmann [62]	Single centre/Germany	AHF	32	Enoximone
Garcia-Gonzalez [63]	Single centre/Spain	AHF post MI	22	Dobutamine
Iyisoy [64]	Single centre/Turkey	AHF, CHF	40	Dobutamine
Levin [65]	Multicentre/Argentina	CHF, CABG	137	Dobutamine
Levin [66]	Multicentre/Argentina	CHF, valve surgery	71	Dobutamine
Levin [5]	Multicentre/Argentina	CHF, CABG	253	Dobutamine
Mebazaa [67]	Multicentre/Austria, Finland, France, Germany, Israel, Latvia, Poland, Russia, and the UK	AHF, CHF	1,327	Dobutamine
Memis [68]	Single centre/Turkey	Sepsis	30	Dobutamine
Morelli [69]	Single centre/Italy	Sepsis	30	Dobutamine
Morelli [70]	Single centre/Italy	Sepsis	40	Dobutamine
Vaitsis [71]	Single centre/Greece	Sepsis	44	Dobutamine
Yilmaz [72]	Single centre/Turkey	AHF, CHF	40	Dobutamine
Yontar [73]	Single centre/Turkey	AHF, CHF	60	Dobutamine

Legend. AHF acute heart failure, CABG coronary artery bypass graft, CHF chronic heart failure, IABP intra-aortic balloon pump, MI myocardial infarction, LCO low cardiac output, No. number

Participants

All trials included critically ill patients with left ventricular dysfunction [i.e., left ventricular ejection fraction (LVEF) <45%] except one trial which evaluated right ventricular dysfunction [37]. The majority of the patients in the trials had acute and/or chronic heart failure; five trials included patients with sepsis, three included patients who underwent a percutaneous coronary intervention, one included patients with pulmonary hypertension and right ventricular failure, seven included patients who underwent valve replacement with or without Coronary Artery Bypass Graft (CABG), and eight only included patients who had CABG procedures (Table 1).

Interventions

Levosimendan was used in different doses. Most trials (n = 30) used a loading dose (6–36 μg/kg) and continued infusion for 24 h (eTable 3). A total of 19 trials (3,302 patients) compared levosimendan with placebo. One trial compared levosimendan with intra-aortic balloon pump. Six trials (282 patients) compared levosimendan with any other control intervention, including 'standard' treatment in four, and prostaglandine E1 and furosemide in one trial each. A total of 26 trials (2,655 patients) compared levosimendan with other inotropes, including enoximone (1 trial), milrinone (5 trials) and dobutamine in (20 trials) in varying doses (Table 1; eTable 3).

Bias risk assessment

Random sequence generation was assessed as low risk of bias in 17 trials (35%), allocation concealment in 14 trials (29%), blinding of participants in 18 trials (37%), blinding of outcome assessors in 15 trials (31%), incomplete outcome data in 18 trials (37%), selective outcome reporting in 31 trials (63%), and industry and/or academic bias in 5 trials (10%) (Fig. 2). Only one trial scored low risk of bias in all bias domains [41]. The a priori defined group of trials with lower risk of bias consisted of 9 trials (18%) and was therefore used for subgroup analyses [2, 31, 35, 37, 38, 41, 44, 61, 67].

Outcomes

All intervention effect estimates with the 95% CI of all outcomes are specified according to bias risk assessment and control group intervention (Table 2). Additionally, GRADEpro summary of findings tables were constructed (eTable 6, 7).

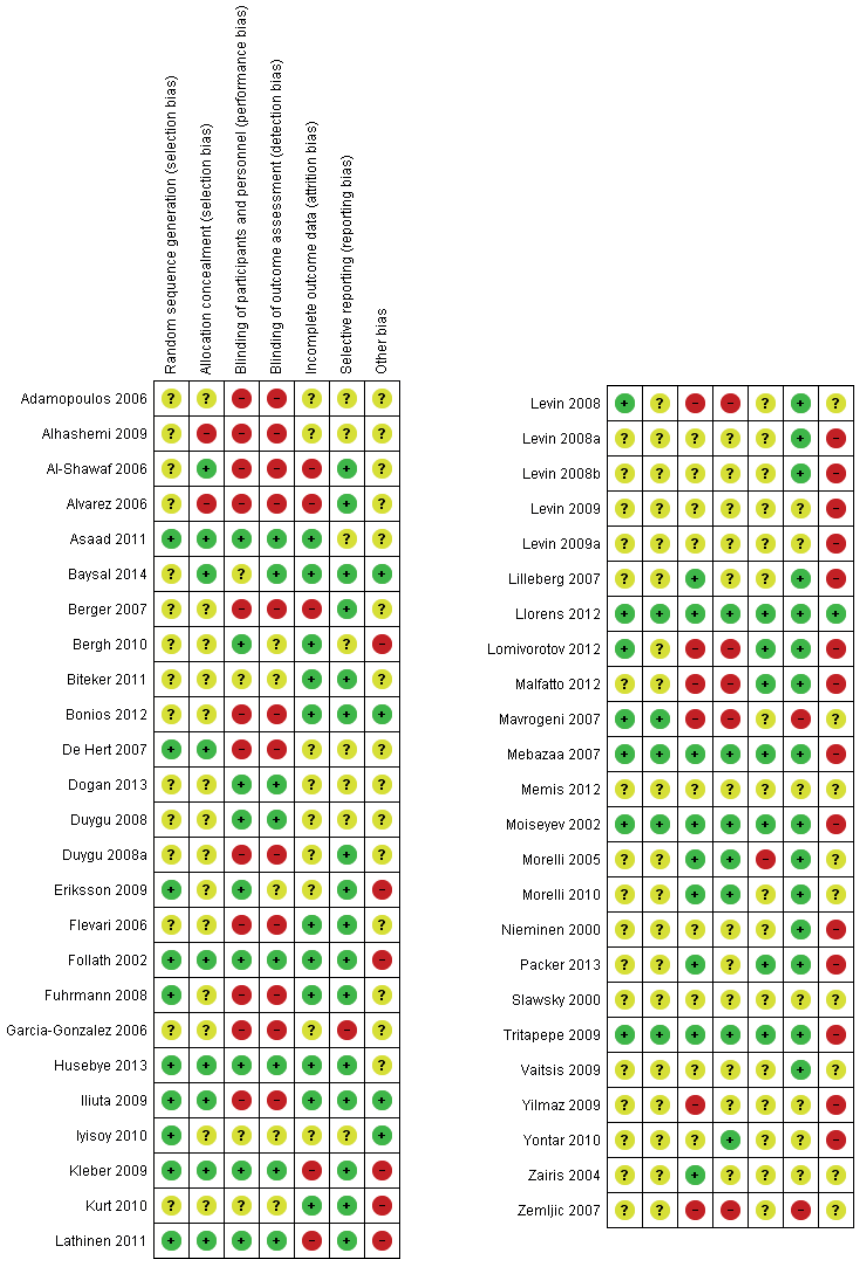


FIGURE 2. Risk of bias summary. Review of authors' judgements about each risk of bias item for each included study. Red high risk, green low risk, yellow unclear

TABLE 2. Conventional and TSA adjusted risk ratios with 95% confidence intervals for the outcome measurements of interest in patients not having cardiac surgery and in patients having cardiac surgery

	Number of trials	Total number of patients in the trials	Conventional meta-analysis RR (95% CI)	TSA-adjusted meta-analysis α = 0.05 (two-sided); power 90%; RRR 10%
Patients not having cardiac surgery				
Mortality				
Placebo				
HR	6	1,138	0.69 (95% CI 0.37–1.29)	
LR	4	638	0.70 (95% CI 0.51–0.96)*	Insufficient data <5% of DIS
All	10	1,776	0.70 (95% CI 0.49–1.00)	Insuficient data <5% of DIS
Test of interaction considering risk of bias subgroups: P = .97				
Control				
HR	6	282	0.52 (95% CI 0.29–0.92)*a	
LR	0		No trials	
All	6	282	0.52 (95% CI 0.29–0.92)*a	Insufficient data <5% of DIS
Other inotropes				
HR	15	806	0.66 (95% CI 0.53–0.82)*	
LR	2	1,530	0.86 (95% CI 0.64–1.15)	Insufficient data <5% of DIS
All	17	2,336	0.73 (95% CI 0.61–0.88)**	0.73 (95% CI 0.34–1.55)
Test of interaction considering risk of bias subgroups: P = .16				
Any control				
HR	24	2,034	0.69 (95% CI 0.56–0.85)**	
LR	6	2,168	0.83 (95% CI 0.70–0.99)*	0.86 (95% CI 0.48–1.55)
All	30	4,202	0.74 (95% CI 0.64–0.86)**	0.76 (95% CI 0.59–0.97)*
Test of interaction considering risk of bias subgroups: P = .16				
Test of interaction considering control intervention subgroups, including all trials: P = .70				
SAE				
Placebo				
HR	0		No trials	
LR	2	532	0.31 (95% CI 0.02–5.37)	Insufficient data <5% of DIS
All	2	532	0.31 (95% CI 0.02–5.37)	Insufficient data <5% of DIS
Control				
All	0		No trials	
Other inotropes				
HR	0		No trials	
LR	2	1,523	0.61 (95% CI 0.17–2.26)	Insufficient data <5% of DIS
All	2	1,523	0.61 (95% CI 0.17–2.26)	Insufficient data <5% of DIS

	Number of trials	Total number of patients in the trials	Conventional meta-analysis RR (95% CI)	TSA-adjusted meta-analysis α = 0.05 (two-sided); power 90%; RRR 10%
Any control				
HR	0		No trials	
LR	4	2,055	0.53 (95% CI 0.20–1.40)	Insufficient data <5% of DIS
All	4	2,055	0.53 (95% CI 0.20–1.40)	Insufficient data <5% of DIS
Test of interaction considering control intervention subgroups, including all trials: P = .67				
MI				
Placebo				
HR	0		No trials	
LR	2	549	1.06 (95% CI 0.40–2.80)a	Insufficient data <5% of DIS
All	2	549	1.06 (95% CI 0.40–2.80)a	Insufficient data <5% of DIS
Control				
HR	1	33	No events	
LR	0			
All	1	33	No events	
Other inotropes				
HR	0		No trials	
LR	2	265	0.06 (95% CI 0.00–1.12)	Insufficient data <5% of DIS
All	2	265	0.06 (95% CI 0.00–1.12)	Insufficient data <5% of DIS
Any control				
HR	2	95	No events	
LR	3	752	0.40 (95% CI 0.05 to 2.94)	Insufficient data <5% of DIS
All	5	847	0.40 (95% CI 0.05–2.94)	Insufficient data <5% of DIS
Test of interaction considering control intervention subgroups, including all trials: P = .07				
VT/VF				
Placebo				
HR	2	832	1.38 (95% CI 1.03–1.85)a	
LR	0		No trials	
All	2	832	1.38 (95% CI 1.03–1.85)a	Insufficient data <5% of DIS
Control				
HR	1	33	0.63 (95% CI 0.21–1.87)	
LR	0		No trials	
All	1	33	0.63 (95% CI 0.21–1.87)	Insufficient data <5% of DIS

	Number of trials	Total number of patients in the trials	Conventional meta-analysis RR (95% CI)	TSA-adjusted meta-analysis α = 0.05 (two-sided); power 90%; RRR 10%
Other inotropes				
HR	2	92	0.75 (95% CI 0.42–1.37)a	
LR	2	1,523	0.98 (95% CI 0.72–1.35)a	Insufficient data <5% of DIS
All	4	1,615	0.95 (95% CI 0.72–1.26)a	Insufficient data <5% of DIS
Test of interaction considering risk of bias subgroups: P = .44				
Any control				
HR	5	957	1.05 (95% CI 0.70–1.56)	
LR	2	1,523	0.99 (95% CI 0.72–1.35)	Insufficient data <5% of DIS
All	7	2,480	1.09 (95% CI 0.89–1.33)	1.09 (95% CI 0.49–2.45)
Test of interaction considering risk of bias subgroups: P = .82				
Test of interaction considering control intervention subgroups, including all trials: P = .11				
SVT				
Placebo				
HR	2	708	4.04 (95% CI 1.84–8.87)**a	
LR	1	504	1.40 (95% CI 0.31–6.20)	Insufficient data <5% of DIS
All	3	1,212	3.26 (95% CI 1.65–6.42)*a	Insufficient data <5% of DIS
Test of interaction considering risk of bias subgroups: P = .22				
Control				
HR	1	33	0.33 (95% CI 0.06–1.71)	
LR	0		No trials	
All	1	33	0.33 (95% CI 0.06–1.71)	Insufficient data <5% of DIS
Other inotropes				
HR	3	132	0.76 (95% CI 0.39–1.49)a	
LR	2	1,523	1.00 (95% CI 0.71–1.40)a	Insufficient data <5% of DIS
All	5	1,655	0.96 (95% CI 0.71–1.30)a	Insufficient data <5% of DIS
Test of interaction considering risk of bias subgroups: P = .48				
Any control				
HR	6	873	1.16 (95% CI 0.40–3.37)	
LR	3	2,027	1.02 (95% CI 0.73–1.41)	
All			NA	NA
Test of interaction considering control intervention subgroups, including all trials: P = .002				

	Number of trials	Total number of patients in the trials	Conventional meta-analysis RR (95% CI)	TSA-adjusted meta-analysis α = 0.05 (two-sided); power 90%; RRR 10%
Hypotension				
Placebo				
HR	2	802	1.35 (95% CI 1.14–1.61)**a	
LR	3	577	1.27 (95% CI 0.60–2.71)a	Insufficient data <5% of DIS
All	5	1,379	1.35 (95% CI 1.14–1.60)**a	Insufficient data <5% of DIS
	Test of interaction considering risk of bias subgroups: P = .88			
Control				
HR	1	75	2.77 (95% CI 0.98–7.81)	
LR	0		No trials	
All	1	75	2.77 (95% CI 0.98–7.81)	Insufficient data <5% of DIS
Other inotropes				
HR	3	215	1.79 (95% CI 0.42–7.72)	
LR	2	1,523	1.23 (95% CI 0.76–1.97)	Insufficient data <5% of DIS
All	5	1,738	1.52 (95% CI 0.80–2.90)	Insufficient data <5% of DIS
	Test of interaction considering risk of bias subgroups: P = .63			
Any control				
HR	5	1,092	1.88 (95% CI 0.97–3.63)	
LR	5	2,100	1.15 (95% CI 0.90–1.46)	1.15 (95% CI 0.43–3.08)
All	10	3,192	1.36 (95% CI 1.06–1.75)*	1.36 (95% CI 0.50–3.73)
	Test of interaction considering risk of bias subgroups: P = .17			
	Test of interaction considering control intervention subgroups, including all trials: P = .38			
Patients having cardiac surgery				
Mortality				
Placebo				
HR	4	937	0.41 (95% CI 0.19–0.88)*	
LR	3	322	1.02 (95% CI 0.48–2.16)	Insufficient data <5% of DIS
All	7	1,259	0.56 (95% CI 0.27–1.14)	Insufficient data <5% of DIS
	Test of interaction considering risk of bias subgroups: P = .10			
Control				
HR	1	60	1.00 (95% CI 0.07–15.66)	
LR			No trials	
All	1	60	1.00 (95% CI 0.07–15.66)	Insufficient data <5% of DIS

	Number of trials	Total number of patients in the trials	Conventional meta-analysis RR (95% CI)	TSA-adjusted meta-analysis α = 0.05 (two-sided); power 90%; RRR 10%
Other inotropes				
HR	6	553	0.39 (95% CI 0.23–0.66)*a	
LR			No trials	
All	6	553	0.39 (95% CI 0.23–0.66)*a	Insufficient data <5% of DIS
Any control				
HR	11	1,550	0.44 (95% CI 0.30–0.64)**	
LR	3	322	1.02 (95% CI 0.48–2.16)	Insufficient data <5% of DIS
All	14	1,872	0.52 (95% CI 0.37–0.73)**	Insufficient data <5% of DIS
Test of interaction considering risk of bias subgroups: P = .05				
Test of interaction considering control intervention subgroups, including all trials: P = .68				
SAE				
Placebo				
All			No trials	
Control				
All			No trials	
Other inotropes				
All			No trials	
Any control				
All			No trials	
MI				
Placebo				
HR	1	600	0.80 (95% CI 0.30–2.12)a	
LR	1	102	0.32 (95% CI 0.01–7.69)a	Insufficient data <5% of DIS
All	2	702	0.73 (95% CI 0.29–1.84)a	Insufficient data <5% of DIS
Test of interaction considering risk of bias subgroups: P = .59				
Control				
All			No trials	
Other inotropes				
HR	2	158	0.48 (95% CI 0.19–1.21)a	
LR			No trials	
All	2	158	0.48 (95% CI 0.19–1.21)a	Insufficient data <5% of DIS

	Number of trials	Total number of patients in the trials	Conventional meta-analysis RR (95% CI)	TSA-adjusted meta-analysis α = 0.05 (two-sided); power 90%; RRR 10%
Any control				
HR	3	758	0.61 (95% CI 0.31–1.19)a	
LR	1	102	0.32 (95% CI 0.01–7.69)a	Insufficient data <5% of DIS
All	4	860	0.59 (95% CI 0.31–1.13)a	Insufficient data <5% of DIS
Test of interaction considering risk of bias subgroups: P = .70				
Test of interaction considering control intervention subgroups, including all trials: P = .53				
VT				
Placebo				
All			No trials	
Control				
All			No trials	
Other inotropes				
HR	2	208	0.26 (95% CI 0.09–0.76)*a	
LR			No trials	
All	2	208	0.26 (95% CI 0.09–0.76)a, c	Insufficient data <5% of DIS
Any control				
HR	2	208	0.26 (95% CI 0.09–0.76)*a	
LR			No trials	
All	2	208	0.26 (95% CI 0.09–0.76)*a	Insufficient data <5% of DIS
SVT				
Placebo				
HR	2	230	0.51 (95% CI 0.20–1.31)	
LR	2	302	1.08 (95% CI 0.91–1.28)	Insufficient data <5% of DIS
All	4	532	0.78 (95% CI 0.41–1.48)	Insufficient data <5% of DIS
Test of interaction considering risk of bias subgroups: P = .13				
Control				
All			No trials	
Other inotropes				
HR	5	407	0.66 (95% CI 0.46–0.94)*	
LR			No trials	
All	5	407	0.66 (95% CI 0.46–0.94)*	Insufficient data <5% of DIS

	Number of trials	Total number of patients in the trials	Conventional meta-analysis RR (95% CI)	TSA-adjusted meta-analysis α = 0.05 (two-sided); power 90%; RRR 10%
Any control				
HR	7	638	0.58 (95% CI 0.43–0.77)**	
LR	2	302	1.08 (95% CI 0.91–1.28)	Insufficient data <5% of DIS
All			NA	NA
Test of interaction considering risk of bias subgroups: P < .001				
Test of interaction considering control intervention subgroups, including all trials: P = .18				
Hypotension				
Placebo				
HR			3.25 (95% CI 0.97–10.90)	
LR			1.63 (95% CI 1.32–2.01)**	Insufficient data <5% of DIS
All	2	277	1.80 (95% CI 1.10–2.95)	Insufficient data <5% of DIS
Test of interaction considering risk of bias subgroups: P = .27				
Control				
All			No trials	
Other inotropes				
HR	1	71	0.39 (95% CI 0.20–0.76)**	
LR			No trials	
All	1	71	0.39 (95% CI 0.20–0.76)**	Insufficient data <5% of DIS
Any control				
HR	2	148	1.06 (95% CI 0.13–8.75)	
LR	1	200	1.63 (95% CI 1.32–2.01)**	Insufficient data <5% of DIS
All			NA	NA
Test of interaction considering risk of bias subgroups: P = .69				
Test of interaction considering control intervention subgroups, including all trials: P < .01				

Legend. All pooled estimates are reported with risk ratio and calculated using a random-effects model unless stated otherwise. Insufficient data <5% DIS = adjusted confidence interval could not be calculated due to sparse data with the predefined α = 0.05 (two sided), β = 0.10 (power 90%), and an anticipated relative risk increase of 10%

All all trials included, *DIS* diversity adjusted required information size, *HR* trials with high risk of bias, *LR* trials with lower risk of bias, *MI* myocardial infarction, *SAE* serious adverse event, *SVT* supraventricular tachyarrhythmias, *VT/VF* ventricular tachyarrhythmias, *NA* not appropriate due to significant test of interaction

* P < .05, ** P < .001

[a] Fixed-effect model when I^2 was 0

Comparison 1: all critically ill patients

Mortality data were obtained from 44 trials including 6,074 patients. Figure 3 shows a forest plot with subgroups according to populations. Heterogeneity was 19%, while control event rates were 25.1% in patients not having cardiac surgery and 10.2% in patients having cardiac surgery. Due to apparent clinical heterogeneity in populations with large differences in control event rates, pooling of all data into one estimate appeared inappropriate.

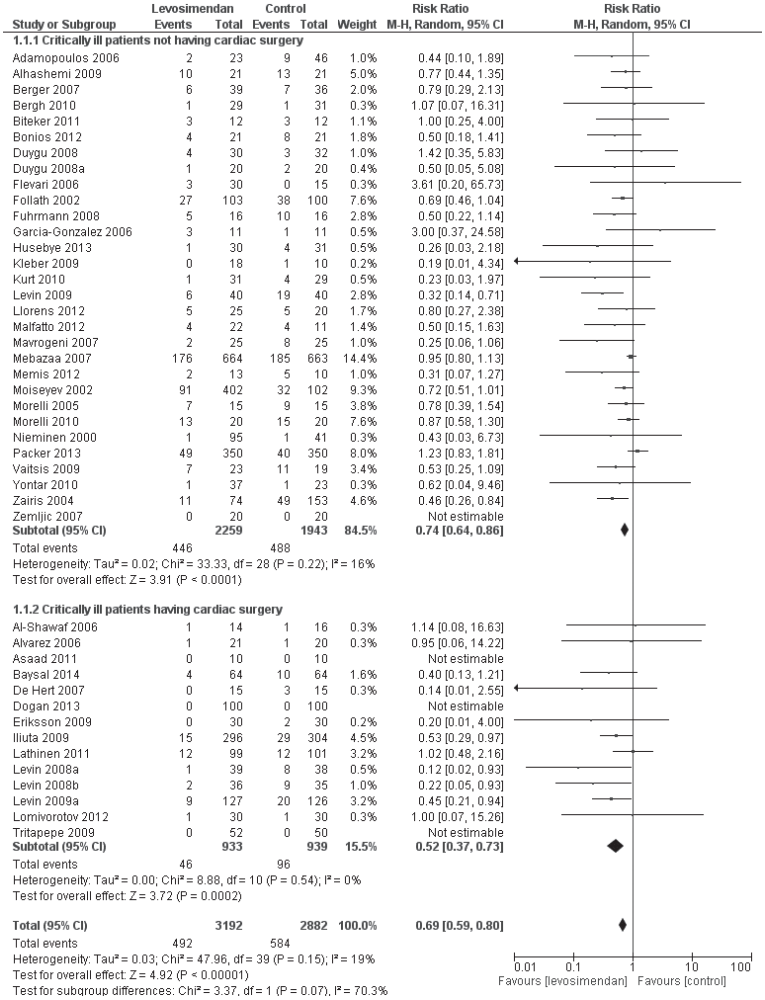


FIGURE 3. Forest plot of all-cause mortality at maximal follow-up of levosimendan compared with any control for low cardiac output syndromes in critically ill patients including all trials with subgroups according to populations. Size of squares for risk ratio (RR) reflects the weight of the trial in the pooled analyses. Horizontal bars 95% confidence intervals (CI)

Comparison 2: critically ill patients not having cardiac surgery

All-cause mortality

Mortality data were obtained from 30 trials (4,202 patients). Ten trials (1,776 patients) comparing levosimendan to placebo showed no association between levosimendan and all-cause mortality at maximal follow-up (RR 0.70, 95% CI 0.49–1.00; eFig. 1). Including only the four trials with lower risk of bias (638 patients) showed a RR of 0.70 (95% CI 0.51–0.96), but the test of interaction on subgroup difference based on bias risk was not significant (P = .97).

The subgroup analyses of six trials (282 patients) with high risk of bias comparing levosimendan to other control interventions found an association between levosimendan and mortality (fixed-effect model; RR 0.52, 95% CI 0.29–0.92).

Seventeen trials (2,336 patients) compared levosimendan to other inotropes and showed that levosimendan was associated with reduced mortality (RR 0.73, 95% CI 0.61–0.88). Subgroup analysis based on risk of bias showed that the association was only found in the 15 trials (806 patients) with high risk of bias and not in the 2 trials (1,530 patients) with lower risk of bias, although the test of interaction was not significant (P = 0.16). However, TSA showed no significant association between levosimendan and mortality (RR 0.73, TSA-adjusted 95% CI 0.34–1.55; eFig. 2).

Pooled analysis of all 30 trials showed an association between levosimendan and mortality (RR 0.74, 95% CI 0.64–0.86; NNT 19, 95% CI 13–36; eFig. 3). The tests of interaction for subgroup differences based on bias risks (P = .16) and types of control interventions (P = .70) were not statistically significant. Additionally, subgroup analysis according to funding bias showed a significant association in the trials with high risk of funding bias but not in the trials with low risk of funding bias (RR 0.71, 95% CI 0.59–0.87 vs RR 0.76, 95% CI 0.56–1.03, respectively; I^2 0%; P = 0.76; eFig. 4).

TSA including all 30 trials confirmed the association between levosimendan and mortality (RR 0.76, TSA-adjusted 95% CI 0.59–0.97; eFig. 5). However, TSA-adjusted 95% CI for the trials with lower risk of bias was 0.48–1.55; only 16% (2,169 patients) of the DIS (13,557 patients) has so far been accrued (Fig. 4). A TSA sensitivity analysis only including trials with lower risk of bias using a RRR of 20% indicates that approximately 1,100 additional randomised patients are required for a definite answer (Fig. 4).

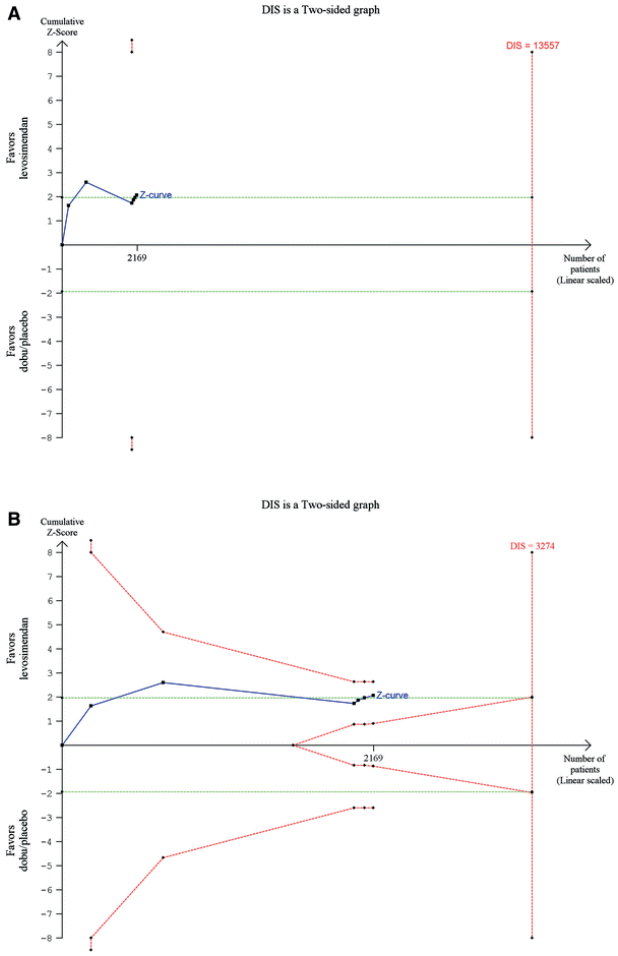


FIGURE 4a. Trial sequential analysis of mortality at maximal follow-up of levosimendan compared with any control in the six trials with lower risk of bias for low cardiac output syndromes in critically ill patients not having cardiac surgery. A diversity-adjusted required information size (DIS) of 13,557 patients was calculated using the predefined α = 0.05 (two-sided), β = 0.10 (power 90%), D2 = 25%, an anticipated relative risk reduction of 10% and an event proportion of 28.6% in the control arm. The blue cumulative z-curve was constructed using a random-effects model. The horizontal green dotted lines represent the conventional boundary's for benefit (positive) or harm (negative). The horizontal red dotted lines represent the trial sequential boundary's for benefit (positive) or harm (negative). **4b.** Trial sequential analysis of mortality at maximal follow-up of levosimendan compared with any control in the six trials with lower risk of bias for low cardiac output syndromes in critically ill patients not having cardiac surgery. A DIS of 3,274 patients was calculated using the predefined α = 0.05 (two-sided), β = 0.10 (power 90%), D2 = 25%, an anticipated relative risk reduction of 20% and an event proportion of 28.6% in the control arm. The blue cumulative z-curve was constructed using a random-effects model. The red dotted area arising at the right of the horizontal axis represents the futility area

Further, of all subgroups, only TSA of levosimendan compared with other inotropes could be calculated and showed no association (eFig. 6): the cumulative z-curve crossed the conventional boundary but not the trial sequential monitoring boundary for benefit.

Serious adverse events

SAE were obtained from four trials (2,055 patients). Pooled analysis of all trials and subgroup analyses on type of control intervention showed no association between levosimendan and SAE (eFig. 7). The test of interaction for subgroup difference based on types of control interventions was not significant (P = .67). No subgroup evaluation on bias risk was conducted as all trials had lower risk of bias. TSA could not be conducted due to too few data.

Myocardial infarction

Five trials (752 patients) provided data on MI. Pooled analysis and subgroup analyses on type of control intervention showed no association between levosimendan and MI (eFig. 8). The test of interaction for subgroup differences based on types of control interventions was not significant (P = .07). No subgroup evaluation on bias risk was conducted as the two trials with high risk of bias had zero events. TSA could not be conducted due to too few data.

Secondary outcomes

VT were reported in seven trials (2,480 patients) and no associations were found (eFig. 9).

SVT were documented in nine trials (2,900 patients). The subgroup analysis of three trials (1,212 patients) comparing levosimendan to placebo showed a significant association between levosimendan and SVT (RR 3.26, 95% CI 1.65–6.42). Other subgroup comparisons did not show an association between levosimendan and SVT (eFig. 10). The test of interaction for subgroup differences based on types of control interventions was significant (P = .002). TSA could not be conducted due to too few data.

Ten trials (3,192 patients) reported hypotension. The subgroup comparing levosimendan to placebo showed that levosimendan was associated with increased risk for hypotension (fixed-effect model; RR 1.35, 95% CI 1.14–1.60). Other subgroup comparisons showed no association between levosimendan and hypotension (eFig. 11).

TSA did not show an association between levosimendan and hypotension. The cumulative z-curve of all trials crossed the conventional boundary for harm, but not the trial sequential monitoring boundary for harm (eFig. 12).

Comparison 3: critically ill patients having cardiac surgery

All-cause mortality

Mortality data were presented in 14 trials (1,872 patients). The subgroups comparing levosimendan to placebo (7 trials, 1,259 patients) and to intra-aortic balloon pump (1 trial, 60 patients) showed no association between levosimendan and all-cause mortality at maximal follow-up (eFig. 13).The subgroup analysis of the 6 trials (553 patients) comparing levosimendan to other inotropes showed an association between levosimendan and mortality (RR 0.41, 95% CI 0.22–0.75; eFig. 14).

The pooled analysis of all 14 trials showed an association between levosimendan and mortality (RR 0.52, 95% CI 0.37–0.73; eFig. 15; NNT 17, 95% CI 12–31). The predefined risk of bias subgroup analysis showed a RR of 1.02 (95% CI 0.48–2.16) for trials with lower risk of bias.

TSA of subgroups and all trials could not be conducted due to too few data. A sensitivity analysis of TSA including all trials using a 20% RRR revealed that the diversity adjusted information size was 11,228 patients. The cumulative z-curve crossed the conventional boundary for benefit but not the trial sequential monitoring boundary for benefit (RR 0.54, TSA-adjusted 95% CI 0.13–2.22; eFig. 16).

Serious adverse events

None of the trials reported SAE other than MI, arrhythmias and hypotension.

Myocardial infarction

Four trials (860 patients) reported data on MI. Pooled analysis of all trials and subgroup analyses showed no significant association between levosimendan and MI (eFig. 17). TSA could not be performed due to too few data.

Secondary outcomes

VT were reported in two trials with high risk of bias. The pooled analysis showed an association between levosimendan and VT (fixed-effect model, RR 0.26, 95% CI 0.09–0.76). The TSA could not be conducted due to insufficient data.

Nine trials reported the incidence of SVT. A subgroup of five trials compared levosimendan with other inotropes and showed a significant association between levosimendan and SVT (RR 0.66, 95% CI 0.46–0.94; eFig. 18A). Pooled analysis and other subgroups found no significant associations (eFig. 18B). TSA could not be conducted due to too few data.

Three trials provided data on the incidence of hypotension. Two trials compared levosimendan to placebo (RR 1.80, 95% CI 1.10–2.95) and one trial compared levosimendan with dobutamine (RR 0.39, 95% CI 0.20–0.76). The test of interaction for subgroup differences based on types of control interventions was significant (P = .002; eFig. 19).

Error matrix approach

The Manhattan error matrix plots of levosimendan showed that there is more data suggesting benefit than that there is data suggesting harm. The standard errors (SE) of the data varied with only few trials having a SE below 0.20. The large majority of the trials had high risks of both systematic errors (bias) and random errors (SE) (eFig. 20 + eFig. 21).

Small trial bias

The funnel plot (eFig. 22 + eFig. 23) showed no clear arguments for small trial bias including publication bias.

Discussion

This systematic review has evaluated the effects of levosimendan in all critically ill patients. Pooling of all data into one overall intervention effect estimate appeared inappropriate due to obvious clinical heterogeneity, and therefore patients having cardiac surgery and patients not having cardiac surgery were considered separately. Conventional meta-analysis of both populations of patients having cardiac surgery and patients not having cardiac surgery showed significant associations between levosimendan and lower mortality when compared to any control. However, all but one trial had high risk of bias and TSA of trials with lower risk of bias showed that the data are too sparse to draw firm conclusions. Therefore, further randomised trials with low risk of both systematic and random errors are urgently awaited. Moreover, even superiority of levosimendan to placebo has never convincingly been established.

Each systematic review depends on the strength of the available randomised trials. There was only 1 trial with low risk of bias among a total of 88. There is convincing empirical evidence from meta-epidemiological studies that high risk of bias is associated with overestimation of beneficial effects and underestimation of harm [74, 75, 76, 77]. Therefore, we a priori defined a subgroup of trials with lower risk of bias to evaluate the effect of bias on intervention estimates. However, even trials with lower risk of bias may suffer from other bias domains, and even these intervention effects may therefore be either over- or under-estimated.

In this meta-analysis we come to other conclusions from authors of previous meta-analyses [8, 9, 10, 11]. Various reasons might explain differences. Previous meta-analyses included search strategies that yielded less randomised trials, not explained by the time span of the searches. More importantly, previous meta-analyses combined heterogeneous populations including patients having cardiac surgery and patients not having cardiac surgery into one overall intervention effect estimate. Further, previous meta-analyses did not explore the association of bias risk with overestimation of intervention effect estimates. Only evaluating the risk of bias of the trials is clearly insufficient: the bias risk assessment has to be taken into the analyses and final conclusions ought to be derived from the trials with low(er) risks of bias [13]. Finally, the risks of random errors have previously been not considered. Conclusions on significance have to be considered in the perspective of sample size considerations, in individual randomised trials, and even in meta-analyses [16, 27]. Therefore, a prepublished protocol, a sensitive search strategy, and thorough evaluation of the risks of systematic errors (bias) and random errors (the play of chance) may explain differences between this systematic review and previous publications [8, 9, 10, 11].

There were several limitations in this systematic review. A large number of trials investigated surrogate outcomes. Another limitation is that most of the included trials only had short-term follow-up (28 days). Further, there were substantial differences in the definitions of the outcomes (i.e. SAE, MI, (supra)ventricular arrhythmias, and hypotension). Also, levosimendan dosing varied between the trials. All TSA evaluations suggested insufficient data to draw firm conclusions.

Future trials need to be designed according to SPIRIT [78], and should especially focus on patient relevant outcomes, like mortality and SAE. Such trials ought to be reported according to the CONSORT guidelines [79].

Conclusion

We found no significant beneficial effects of levosimendan for low cardiac output syndromes in critically ill patients, and, therefore, daily use of levosimendan should not be propagated. Levosimendan may or may not be associated with a decrease in mortality in patients with cardiac dysfunction and further well-designed randomised trials are needed. Several trials are underway planned with some 1,500 additional patients, while probably many more patients are necessary for definitive conclusions.

References

1. Lilleberg J, Sundberg S, Nieminen MS. Dose-range study of a new calcium sensitizer, levosimendan, in patients with left ventricular dysfunction. J Cardiovasc Pharmacol 1995;26(Suppl 1):S63–S69
2. Moiseyev VS, Poder P, Andrejevs N et al. Safety and efficacy of a novel calcium sensitizer, levosimendan, in patients with left ventricular failure due to an acute myocardial infarction. A randomized, placebo-controlled, double-blind study (RUSSLAN). Eur Heart J 2002;23:1422–1432
3. Zairis M, Apostolatos C, Anastasiadis P et al. The effect of a calcium sensitizer or an inotrope or none in chronic low output decompensated heart failure: results from the calcium sensitizer or inotrope or none in low output heart failure study (CASINO). J Am Coll Cardiol 2004;43:206A–207A. doi: 10.1016/S0735-1097(04)90879-0
4. Levin R, Degrange M, Porcile R et al. Preoperative use of calcium sensitizer levosimendan reduces mortality and low cardiac output syndrome in patients with aortic stenosis and left ventricular dysfunction. Circulation 2008;118:E217
5. Levin R, Porcile R, Salvagio F et al. Levosimendan reduces mortality in postoperative low cardiac output syndrome after coronary surgery. Circulation 2009;120:S987–S988
6. Mebazaa A, Nieminen MS, Filippatos GS et al. Levosimendan vs. dobutamine: outcomes for acute heart failure patients on beta-blockers in SURVIVE. Eur J Heart Fail 2009;11:304–311. doi:10.1093/eurjhf/hfn045
7. Packer M, Colucci W, Fisher L et al. Effect of levosimendan on the short-term clinical course of patients with acutely decompensated heart failure. JACC Heart Fail 2013;1:103–111
8. Delaney A, Bradford C, McCaffrey J, Bagshaw SM, Lee R. Levosimendan for the treatment of acute severe heart failure: a metaanalysis of randomised controlled trials. Int J Cardiol 2010;138:281–289. doi: 10.1016/j.ijcard.2008.08.020
9. Landoni G, Biondi-Zoccai G, Greco M et al. Effects of levosimendan on mortality and hospitalization. A metaanalysis of randomized controlled studies. Crit Care Med 2012;40:634–646
10. Huang X, Lei S, Zhu MF, Jiang RL, Huang LQ, Xia GL, Zhi YH. Levosimendan versus dobutamine in critically ill patients: a meta-analysis of randomized controlled trials. J Zhejiang Univ Sci B 2013;14:400–415. doi: 10.1631/jzus.B1200290
11. Harrison RW, Hasselblad V, Mehta RH, Levin R, Harrington RA, Alexander JH. Effect of levosimendan on survival and adverse events after cardiac surgery: a meta-analysis. J Cardiothorac Vasc Anesth 2013;27:1224–1232. doi: 10.1053/j.jvca.2013.03.027
12. Keus F, Wetterslev J, Gluud C, van Laarhoven CJ. Evidence at a glance: error matrix approach for overviewing available evidence. BMC Med Res Methodol 2010;10:90. doi:10.1186/1471-2288-10-90
13. Higgins JPT, Green S (editors). Cochrane handbook for systematic reviews of interventions version 5.1.0 (updated March 2011). 2011. The Cochrane Collaboration
14. Brok J, Thorlund K, Wetterslev J, Gluud C. Apparently conclusive meta-analyses may be inconclusive— trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal metaanalyses. Int J Epidemiol 2009;38:287–298. doi:10.1093/ije/dyn188
15. Thorlund K, Devereaux PJ, Wetterslev J et al. Can trial sequential monitoring boundaries reduce spurious inferences from meta-analyses? Int J Epidemiol 2009;38:276–286. doi: 10.1093/ije/dyn179

16. Pogue JM,Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative metaanalysis. Control Clin Trials 1997;18:580–593; discussion 661–6
17. Pogue J, Yusuf S. Overcoming the limitations of current meta-analysis of randomised controlled trials. Lancet 1998;351:47–52. doi: 10.1016/S0140-6736(97)08461-4
18. Thorlund K, Imberger G, Walsh M et al. The number of patients and events required to limit the risk of overestimation of intervention effects in meta-analysis–a simulation study. PLoS ONE 2011;6:e25491. doi:10.1371/journal.pone.0025491
19. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ, GRADE Working Group. What is "quality of evidence" and why is it important to clinicians? BMJ 2008;336:995–998. doi:10.1136/bmj.39490.551019.BE
20. Keus F, Koster G, van der Horst ICC, Zijlstra JG, Wetterslev J, Gluud C. Levosimendan for low cardiac output syndromes: a systematic review with meta-analysis and trial sequential analysis. 2013 http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013003871
21. Rao V, Ivanov J, Weisel RD, Ikonomidis JS, Christakis GT, David TE. Predictors of low cardiac output syndrome after coronary artery bypass. J Thorac Cardiovasc Surg 1996;112:38–51. doi:10.1016/S0022-5223(96)70176-9
22. International Conference on Harmonisation Expert Working Group. International conference on harmonisation of technical requirements for registration of pharmaceuticals for human use, ed. ICH harmonised tripartite guideline; guideline for good clinical practice. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf
23. Review Manager (RevMan) [Computer Program] Version 5.2 Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012
24. Copenhagen Trial Unit. TSA— trial sequential analysis. http://ctu.dk/tsa/. 2011. Accessed 4 September 2012
25. Thorlund K, Engstrøm J, Wetterslev J, Brok J, Imberger G, Gluud C. User manual for trial sequential analysis (TSA). http://ctu.dk/tsa/files/tsa_manual.pdf. 2011. Accessed 12 May 2014
26. Higgins JP, Thompson SG. Quantifying heterogeneity in a metaanalysis. Stat Med 2002;21:1539–1558. doi: 10.1002/sim.1186
27. Wetterslev J, Thorlund K, Brok J, Gluud C. Trial sequential analysis may establish when firm evidence is reached in cumulative metaanalysis. J Clin Epidemiol 2008;61:64–75. doi:10.1016/j.jclinepi.2007.03.013
28. Wetterslev J, Engstrøm J, Gluud C, Thorlund K. Trial sequential analysis: methods and software for cumulative meta-analyses. Cochrane Methods Cochrane Database Syst Rev 2012;suppl 1:29–31
29. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model metaanalyses. BMC Med Res Methodol 2009;9:86. doi:10.1186/1471-2288-9-86
30. Adamopoulos S, Parissis JT, Iliodromitis EK et al. Effects of levosimendan versus dobutamine on inflammatory and apoptotic pathways in acutely decompensated chronic heart failure. Am J Cardiol 2006;98:102–106. doi: 10.1016/j.amjcard.2006.01.068
31. Asaad OM, Hanafy MS. Levosimendan's effect on coronary artery grafts blood flow in patients with left ventricular dysfunction, assessment by transit time flow meter. Egypt J Anaesth 2011;27:45–53
32. Dogan OF. Levosimendan use decreases atrial fibrillation in patients after coronary artery by-pass grafting: a pilot study. Heart Surg Forum 2013;16:E287– E294. doi:10.1532/HSF98.2013190

33. Eriksson HI, Jalonen JR, Heikkinen LO et al. Levosimendan facilitates weaning from cardiopulmonary bypass in patients undergoing coronary artery bypass grafting with impaired left ventricular function. Ann Thorac Surg 2009;87:448–454. doi:10.1016/j.athoracsur.2008.10.029

34. Flevari P, Parissis JT, Leftheriotis D, Panou F, Kourea K, Kremastinos DT. Effect of levosimendan on ventricular arrhythmias and prognostic autonomic indexes in patients with decompensated advanced heart failure secondary to ischemic or dilated cardiomyopathy. Am J Cardiol 2006;98:1641–1645. doi:10.1016/j.amjcard.2006.07.043

35. Husebye T, Eritsland J, Muller C et al. Levosimendan in acute heart failure following primary percutaneous coronary intervention-treated acute STelevation myocardial infarction. Results from the LEAF trial: a randomized, placebo-controlled study. Eur J Heart Fail 2013;15:565–572. doi:10.1093/eurjhf/hfs215

36. Iliuta L. The Cost-Benefit Report associated with routinely using levosimendan in patients with systolic dysfunction undergoing coronary artery bypass grafting. Eur J Heart Fail 2009;8(Suppl S2):ii392. doi:10.1093/eurjhf/hfp082

37. Kleber FX, Bollmann T, Borst MM et al. Repetitive dosing of intravenous levosimendan improves pulmonary hemodynamics in patients with pulmonary hypertension: results of a pilot study. J Clin Pharmacol 2009;49:109–115. doi:10.1177/0091270008325150

38. Lahtinen P, Pitkanen O, Polonen P, Turpeinen A, Kiviniemi V, Uusaro A. Levosimendan reduces heart failure after cardiac surgery: a prospective, randomized, placebocontrolled trial. Crit Care Med 2011;39:2263–2270

39. Levin R, Porcile R, Tanus E, Degrange M. The intermittent infusion of levosimendan reduces mortality and readmisions in patients with advanced heart failure. Circulation 2009;120: S865–S865

40. Lilleberg J, Laine A, Palkama T, Kivikko A, Pohjanjousi P, Kupari M. Duration of the haemodynamic action of a 24-h infusion of levosimendan in patients with congestive heart failure. Eur J Heart Fail 2007;9:75–82. doi:10.1016/j.ejheart.2006.04.012

41. Llorens P, Miro O, Roman F, Zapater P, Carbajosa-Dalmau J, Llanos L. Efficacy of early administration of levosimendan in emergency department in patients with acute heart failure: a randomized pilot clinical trial. Emergencias 2012;24:268–276

42. Nieminen MS, Akkila J, Hasenfuss G et al. Hemodynamic and neurohumoral effects of continuous infusion of levosimendan in patients with congestive heart failure. J Am Coll Cardiol 2000;36:1903–1912

43. Slawsky MT, Colucci WS, Gottlieb SS et al. Acute hemodynamic and clinical effects of levosimendan in patients with severe heart failure. Circulation 2000;102:2222–2227

44. Tritapepe L, De Santis V, Vitale D, Guarracino F, Pellegrini F, Pietropaoli P, Singer M. Levosimendan pretreatment improves outcomes in patients undergoing coronary artery bypass graft surgery. Br J Anaesth 2009;102:198–204

45. Berger R, Moertl D, Huelsmann M, Bojic A, Ahmadi R, Heissenberger I, Pacher R. Levosimendan and prostaglandin E1 for uptitration of betablockade in patients with refractory, advanced chronic heart failure. Eur J Heart Fail 2007;9:202–208. doi: 10.1016/j.ejheart.2006.06.001

46. Biteker M, Duran NE, Kaya H et al. Effect of levosimendan and predictors of recovery in patients with peripartum cardiomyopathy, a randomized clinical trial. Clin Res Cardiol 2011;100:571–577. doi:10.1007/s00392-010-0279-7

47. Kurt IH, Yavuzer K, Batur MK. Short-term effect of levosimendan on free light chain kappa and lambda levels in patients with decompensated chronic heart failure. Heart Vessels 2010;25:392–399. doi:10.1007/s00380-009-1216-4

48. Lomivorotov VV, Boboshko VA, Efremov SM et al. Levosimendan versus an intra-aortic balloon pump in high-risk cardiac patients. J Cardiothorac Vasc Anesth 2012;26:596–603. doi:10.1053/j.jvca.2011.09.006

49. Malfatto G, Della Rosa F, Villani A, Rella V, Branzi G, Facchini M, Parati G. Intermittent levosimendan infusions in advanced heart failure: favourable effects on left ventricular function, neurohormonal balance, and one-year survival. J Cardiovasc Pharmacol 2012;60:450–455

50. Mavrogeni S, Giamouzis G, Papadopoulou E et al. A 6-month follow-up of intermittent levosimendan administration effect on systolic function, specific activity questionnaire, and arrhythmia in advanced heart failure. J Card Fail 2007;13:556–559

51. Zemljic G, Bunc M, Yazdanbakhsh AP, Vrtovec B. Levosimendan improves renal function in patients with advanced chronic heart failure awaiting cardiac transplantation. J Card Fail 2007;13:417–421. doi:10.1016/j.cardfail.2007.03.005

52. Alhashemi JA, Alotaibi Q. Levosimendan versus dobutamine in septic shock. Crit Care 2009;13(Suppl 1):Abstract no: p171. doi:10.1186/cc7335

53. Al-Shawaf E, Ayed A, Vislocky I, Radomir B, Dehrab N, Tarazi R. Levosimendan or milrinone in the type 2 diabetic patient with low ejection fraction undergoing elective coronary artery surgery. J Cardiothorac Vasc Anesth 2006;20:353–357. doi:10.1053/j.jvca.2006.02.012

54. Alvarez J, Bouzada M, Fernandez A et al. Hemodynamic effects of levosimendan compared with dobutamine in patients with low cardiac output after cardiac surgery. Rev Esp Cardiol 2006;59:338–345. doi:10.1157/13087055

55 Baysal A, Yanartas M, Dogukan M, Gundogus N, Kocak T, Koksal C. Levosimendan improves renal outcome in cardiac surgery: a randomized trial. J Cardiothorac Vasc Anesth 2014; 28:586–594

56. Bergh CH, Andersson B, Dahlstrom U et al. Intravenous levosimendan vs. dobutamine in acute decompensated heart failure patients on beta-blockers. Eur J Heart Fail 2010;12:404–410. doi:10.1093/eurjhf/hfq032

57. Bonios MJ, Terrovitis JV, Drakos SG et al. Comparison of three different regimens of intermittent inotrope infusions for end stage heart failure. Int J Cardiol 2012;159:225–229. doi:10.1016/j.ijcard.2011.03.013

58. De Hert SG, Lorsomradee S, Cromheecke S, Van DL. The effects of levosimendan in cardiac surgery patients with poor left ventricular function. Anesth Analg 2007;104:766–773

59. Duygu H, Ozerkan F, Nalbantgil S et al. Effect of levosimendan on E/E' ratio in patients with ischemic heart failure. Int J Cardiol 2008;123:201–203

60. Duygu H, Turk U, Ozdogan O et al. Levosimendan versus dobutamine in heart failure patients treated chronically with carvedilol. Cardiovasc Ther 2008;26:182–188. doi:10.1111/j.1755-5922.2008.00050.x

61. Follath F, Cleland JG, Just H et al. Efficacy and safety of intravenous levosimendan compared with dobutamine in severe low-output heart failure (the LIDO study): a randomised double-blind trial. Lancet 2002;360:196–202

62. Fuhrmann JT, Schmeisser A, Schulze MR et al. Levosimendan is superior to enoximone in refractory cardiogenic shock complicating acute myocardial infarction [corrected] [published erratum appears in Crit Care Med 2008 Oct; 36(10):2966]. Crit Care Med 2008;36:2257–2266

63. Garcia-Gonzalez M, Dominguez-Rodriguez A, Ferrer-Hita J, Abreu-Gonzalez P, Munoz MB. Cardiogenic shock after primary percutaneous coronary intervention: effects of levosimendan compared with dobutamine on haemodynamics. Eur J Heart Fail 2006;8:723–728

64. Iyisoy A, Celik T, Celik M, Bugan B, Yaman H. Comparative effects of levosimendan and dobutamine infusion on p wave dispersion in patients with acute decompensated heart failure. Turk J Med Sci 2010;40:761–770

65. Levin RL, Degrange MA, Porcile R et al. The calcium sensitizer levosimendan gives superior results to dobutamine in postoperative low cardiac output syndrome. Rev Esp Cardiol 2008;61:471–479

66. Levin R, Degrange M, Del Mazo C et al. Superiority of Levosimendan over Dobutamine in low cardiac output syndrome after aortic valve replacement. Circulation 2008;118:E231

67. Mebazaa A, Nieminen MS, Packer M et al. Levosimendan vs dobutamine for patients with acute decompensated heart failure: the SURVIVE randomized trial. JAMA 2007;297:1883–1891

68. Memis D, Inal MT, Sut N. The effects of levosimendan vs dobutamine added to dopamine on liver functions assessed with noninvasive liver function monitoring in patients with septic shock. J Crit Care 2012;27:318.e1–318.e6. doi:10.1016/j.jcrc.2011.06.008

69. Morelli A, Castro S, Teboul JL et al. Effects of levosimendan on systemic and regional hemodynamics in septic myocardial depression. Intensive Care Med 2005;31:638–644. doi:10.1007/s00134-005-2619-z

70. Morelli A, Donati A, Ertmer C et al. Levosimendan for resuscitating the microcirculation in patients with septic shock: a randomized controlled study. Crit Care 2010;14:R232. doi:10.1186/cc9387

71. Vaitsis J, Michalopoulou H, Thomopoulos C, Massias S, Stamatis P. Use of levosimendan in myocardial dysfunction due to sepsis. Crit Care 2009;13:S67

72. Yilmaz MB, Yontar C, Erdem A et al. Comparative effects of levosimendan and dobutamine on right ventricular function in patients with biventricular heart failure. Heart Vessels 2009;24:16–21. doi:10.1007/s00380-008-1077-2

73. Yontar OC, Yilmaz MB, Yalta K, Erdem A, Tandogan I. Acute effects of levosimendan and dobutamine on QRS duration in patients with heart failure. Arq Bras Cardiol 2010l95:738–742

74. Savovic J, Jones HE, Altman DG et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med 2012;157:429–438. doi:10.7326/0003-4819-157-6-201209180-00537

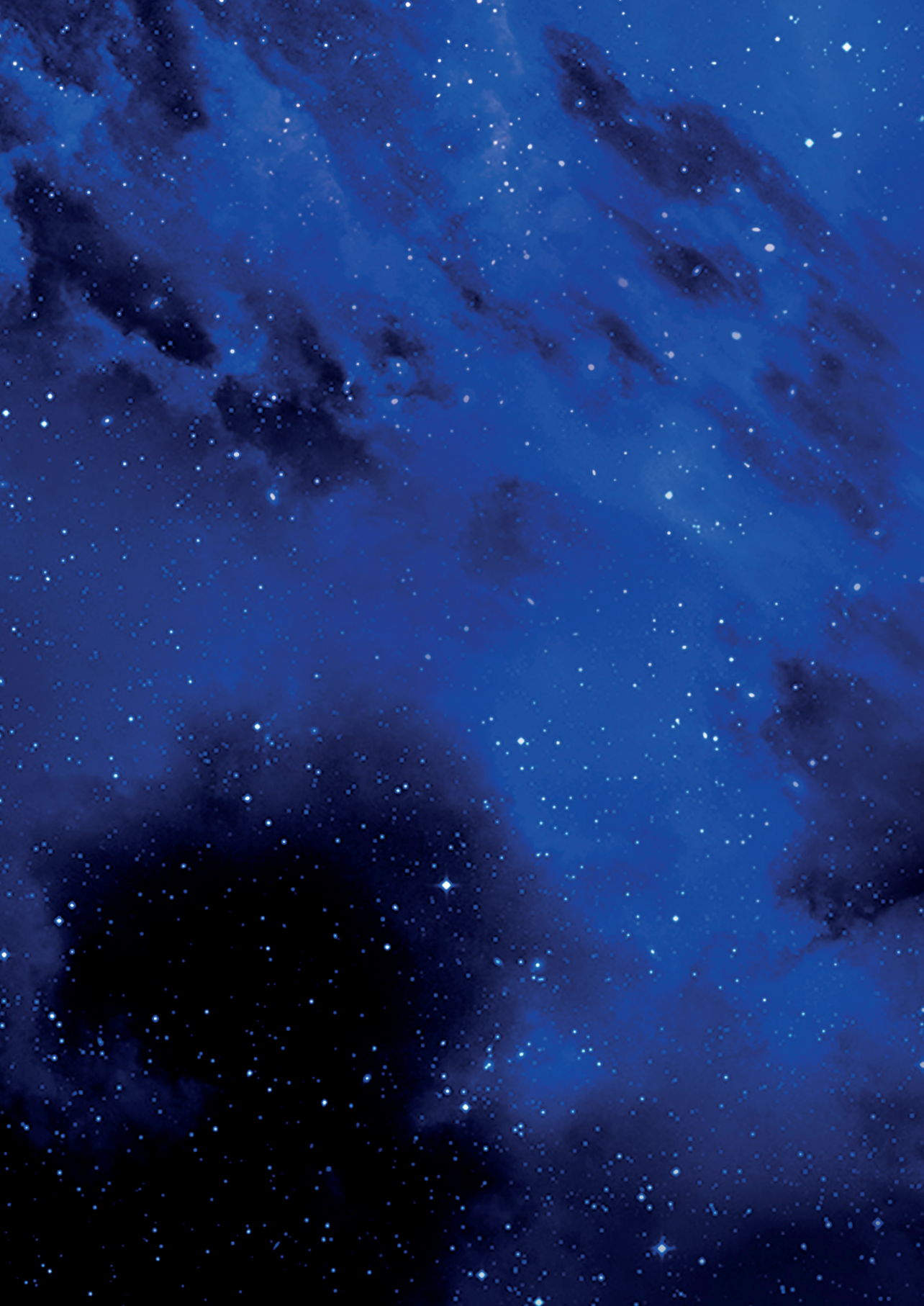
75. Wood L, Egger M, Gluud LL et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. BMJ 2008;336:601–605. doi:10.1136/bmj.39465.451748.AD

76. Kjaergard LL, Villumsen J, Gluud C. Reported methodologic quality and discrepancies between large and small randomized trials in metaanalyses. Ann Intern Med 2001;135:982–989

77. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. JAMA 1995;273:408–412

78. Chan A, Tetzlaff JM, Altman DG et al. SPIRIT 2013 Statement: defining Standard Protocol Items for Clinical Trials. Ann Intern Med 2013;158:200–207. doi:10.7326/0003-4819-158-3-201302050-00583

79. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMJ 2010;340:c332. doi:10.1136/bmj.c332



PART 2

Hemodynamic monitoring

5

Critical care ultrasonography in circulatory shock

Koster G, van der Horst ICC.
Curr Opin Crit Care 2017 Aug;23(4):326-333. doi: 10.1097/MCC.0000000000000428.

Abstract

Background

Assessment of any patient's hemodynamic profile based on clinical examination can be sufficient in several cases, but many times unclarities remain. Arterial catheters and central venous lines are commonly used in critically ill patients for practical reasons, and offer an opportunity for advanced hemodynamic monitoring. Critical care ultrasonography may add to the understanding of the hemodynamic profile at hand. Improvements in ultrasound techniques, for example, smaller devices and improved image quality, may reduce limitations and increase its value as a complementary tool. Critical care ultrasonography has great potential to guide decisions in the management of shock, but operators should be aware of limitations and pitfalls as well. Current evidence comes from cohort studies with heterogeneous design and outcomes. The objective was to define the role of ultrasound in the diagnosis and the management of circulatory shock by critical appraisal of the literature.

Conclusion

Use of ultrasonography for hemodynamic monitoring in critical care expands, probably because of absence of procedure-related adverse events. Easy applicability and the capacity of distinguishing different types of shock add to its increasing role, further supported by consensus statements promoting ultrasound as the preferred tool for diagnostics in circulatory shock.

Introduction

Circulatory shock affects about one third of patients admitted to an ICU [1]. Shock reduces oxygen and energy supply to organs and is associated with increased mortality [1,2]. Traditionally, four types of shock (i.e., the 'classic' shock states) have been distinguished by pathophysiological mechanism [3]. Critically ill patients frequently present with some combination(s) of one or more of these four types of shock. A large randomized trial found septic shock (as the most common form of distributive shock) in 62% of the patients, followed by cardiogenic and hypovolemic shock (both 16%), other types of distributive shock in 4%, and obstructive shock in 2% of the patients [4]. Such a priori incidences should be considered irrespective all diagnostic information available.

The ability of a clinician to detect or refute circulatory shock is fundamental for the care for the critically ill. A thorough physical examination is always the first step in the diagnosis of circulatory shock, despite its evidence base being weak [5]. Each additional physiological examination provides a collection of hemodynamic variables, but even merged, their ability to adequately assess a hemodynamic profile is considered weak [5,6].

The initial clinical management of the patient with circulatory shock may both include preliminary treatment based on initial diagnostic information [7,8] and use of additional diagnostic tools to inform the diagnosis for tailored treatment [2].

Additional diagnostic testing for a more accurate hemodynamic profile can either be invasive [e.g., pulmonary artery catheter (PAC)], less invasive [e.g., indicator dilution or arterial blood pressure (BP) waveform analysis], or noninvasive (e.g., ultrasonography) [9]. Many patients in the ICU have an arterial and venous catheter for standard care, so that less invasive tools require no additional invasive procedures. The evidence supporting the accuracy of less invasive monitoring in the management of critically ill varies with the technique [9]. A recent survey among German intensivists showed that ultrasonography was used by 99% as a tool to diagnose the type of circulatory shock [10]. Over 70% of all responders require that cardiac output (CO) is measured by ultrasonography in case of hemodynamic instability [10&].

Many extensive overviews have evaluated ultrasonography in critical and emergency care [11–18]. This opinion presents the evidence on the use of critical care ultrasonography (CCUS) for the diagnosis and management of circulatory shock [19].

ULTRASONOGRAPHY VERSUS PULMONARY ARTERY CATHETER

Ultrasonography is advocated as the preferred tool in the diagnostic process of the hemodynamic profile in shock, but also as tool for monitoring its management [2]. Still, the PAC is considered the gold standard for hemodynamic monitoring. PAC provides insight in nearly all hemodynamic variables, including continuous CO and mixed oxygen saturation measurements. Its use has declined because of absence of beneficial effects on outcomes [20,21], and because reasons for advanced monitoring decreased, like cardiogenic shock after acute myocardial infarction (MI) [22]. Further, interpretations of PAC measurements were shown difficult and might partially explain the lack of effect on outcome [23]. Additionally, the reliability and accuracy of the algorithms and CO measurements of PAC have been questioned as is the accuracy to assess fluid responsiveness [23]. The advocated criticism on the potential role of PAC in clinical practice also accounts for ultrasonography [24]. Evidence on the effect on outcome is lacking for CCUS as well and CO measurements with both techniques show no strong correlation [18,25,26,27]. Moreover, in studies evaluating ultrasonography to measure CO the required images are unobtainable in up to 20% of patients [28,29]. In summary, the PAC is not recommended as routine tool, but should be considered in patients with more complex cardiopulmonary conditions (e.g., pulmonary hypertension or after lung transplant) and if other monitoring tools, like CCUS, give insufficient measurements [30].

CRITICAL CARE ULTRASONOGRAPHY VERSUS ECHOCARDIOGRAPHY

Several factors account for the increasing use of CCUS [14]. First, its portability has increased dramatically to even tablet or smartphone dimensions [31]. Second, its noninvasiveness and absence of any adverse event reduces all thresholds. Third, image quality continues to improve. Fourth, images can be interpreted immediately.

Terminology (i.e. acronyms) and protocols of CCUS have been confusing [32–34]. The majority of the protocols are essentially alike. Varying terminology originates from emphasizing differences between CCUS and echocardiography. In this manuscript we defined focused echocardiography as CCUS.

Echocardiography is ultrasonography of the heart and large vessels either by transthoracic echocardiography (TTE) or transoesophageal echography (TOE). CCUS is ultrasonography of the heart and large vessels as well and may extend to the lungs and pleural cavities, abdomen, legs, and other [35]. Echocardiography acquires all standard acoustic views from all standard windows, whereas CCUS is by definition focused on the clinical question at hand. CCUS has shown to be superior in detecting cardiac abnormalities compared to standard physical examination

[36]. Indications for ordering a CCUS differ. In one study the indications were information on left ventricular function (46%), cause of hypotension (17%), and pulmonary oedema (14%), respectively [37].

Echocardiography is performed by a certified sonographer or cardiologist, whereas CCUS can be performed by any critical care personnel (including intensivists, fellows, residents, nurses, or students) [38]. Initiatives for training and certification of qualified personnel are emerging. Data suggests that only limited education and training of intensivists or even medical students suffices for obtaining adequate images and measurements [16,28]. Debate exists on the number of supervised examinations needed to be certified. Clear differences exist between the recommended minimal number necessary for CCUS compared with echocardiography [39]. Recently, certification for CCUS became more uniform and the European accreditation is available for advanced CCUS (European Diploma in advanced critical care EchoCardiography).

CCUS is hampered by suboptimal positioning and complicating patient conditions (i.e., postsurgery, emphysema, or chest tubes). Studies report success rates of sufficient image acquisition ranging from 72 to 99% [28,29,40,41,42,43]. Though, figures of 97% sufficient images should be interpreted with care, as this proportion was derived from cumulative views (58% subcostal, 80% apical, and 69% parasternal) [29]. Further, interpretability of acquired images varies between 80 and 100% [17,40,41,42]. In one study image quality was equally good or adequate in patients on mechanical ventilation as compared with spontaneously breathing patients (90 versus 91%) [37].

One major concern in CCUS is the risk of diagnostic and cognitive errors when interpreting CCUS data (both over and underdiagnosing [11,18,44,45]. This error risk is probably driven by differences in training between the echocardiography and CCUS operators. In one study image acquisition by noncardiologist intensivists was assessed successfully in 94% by an experienced cardiologist; and image interpretation of echocardiographic findings correctly in 84% [40]. A recent study also found no contradiction in results between CCUS and echocardiography, but the latter provided additional diagnostic information which confirms that the two modalities are different [43]. It is, however, reassuring that conclusions remained the same.

In summary, CCUS focusses on pattern recognition (i.e., the confirmation or rejection of hypotheses) and is suited for quickly informing the clinical question, even in urgent situations, by use of a minimal set of acoustic views, in contrast with echocardiography where the full examination is performed for informing the clinical question afterwards. As such, CCUS does never replace a comprehensive echocardiographic evaluation.

STATIC AND DYNAMIC PARAMETERS IN CRITICAL CARE ULTRASONOGRAPHY

CCUS is ideally suited for determining static and dynamic parameters depending on the clinical question. The measurement of static hemodynamic parameters (by CCUS) have not shown to be helpful in assessing fluid responsiveness and will not be discussed here [18].

Examples of quantitative dynamic parameters useful in daily practice include measurement of CO and vena cava collapsibility. CO can be measured by tracing of velocity time integral over the left ventricular outflow tract. The measurement itself is often validated [46] and data form studies before 1990 have shown acceptable comparability with PAC-derived CO [47]. Recently, a systematic review questioned this comparability [27]. Most recently, a new study on the correlation of TTE and PAC was published [48]. In 38 mechanically ventilated patients a total of 64 pairs of CO measurements were compared and a median bias of 0.2 L/min was observed. To our opinion this difference is acceptable for daily practice. In this study patients with arrhythmia and spontaneous ventilation were excluded. The measurement can be used repeatedly for observing trends in hemodynamics or assessing fluid responsiveness with or without passive leg raising [35]. Vena cava collapsibility/distensibility, which is associated with right atrial pressure/central venous pressure, can assess fluid responsiveness [15,35,36]. Most evidence is available for a collapse of less or greater than 50% as a sign of higher or lower right atrial pressure/central venous pressure, although a recent study found a cutoff of 25% to be more accurate [49]. Two important conditions for diagnoses made on inferior vena cava measurements are that data derived should be interpreted as one measure among others to define the likelihood of fluid responsiveness, like considering the size of the right ventricle, the tricuspid annular plane systolic excursion, and that the measures depend on the clinical scenario, like spontaneous versus mechanical ventilation, the presence of valvular disorder, and pericardial effusion. Finally, the vena cava inferior moves during breathing and all studies based their optimal cutoff on maximal diameters. Diameters of the inferior vena cava during CCUS can only be interpreted if the real diameters are measured.

So, dynamic measures can act as trigger for initiating therapy (i.e., start of inotrope with a low CO) or in case of fluid responsiveness in a septic patient to give fluid therapy [35,50] or observe trends in the hemodynamic profile (i.e., repetitive measurement of CO/velocity time integral).

EFFECT OF CRITICAL CARE ULTRASONOGRAPHY IN DIAGNOSIS AND MANAGEMENT

CCUS may assist the clinician in setting the correct diagnosis of the type of shock and further tailor its management so that diagnostic errors and potential harmful interventions can be prevented. Data on the effect of CCUS is increasing. To summarize the available evidence the effect can be described in time to diagnosis, change in diagnosis, and change in management. All these effects are focused on improving patients' outcome (Table 1). The evidence for establishing a diagnosis earlier comes from the emergency department. In a randomised trial of 184 patients CCUS resulted in a higher rate of diagnosis after 15 min in the group of immediate compared to delayed ultrasound [51]. Another retrospective study of 3834 patients confirmed that early use of CCUS was associated with decisions making earlier [52]. One small randomized trial showed that the early use of CCUS made clinical decision-making faster compared with delayed availability in the emergency department [52].

Multiple studies have shown that CCUS complementary to clinical examination can lead to frequent changes in diagnosis (ranging from 28 to 67%) [41,52–56]. Whether these changes were large or categorical, and/or relevant to patient important outcomes were not explicitly reported.

Different studies showed different changes in management based on CCUS. In a study management changed following TTE in 16% and TOE in 36% and comprised mostly of a change in vasopressors and fluid challenge [57]. Another study observed a change in management of 49% with TTE and 54% with TOE, and consisted of altered drug therapy in 22%, initiation or increase in inotropes in 14%, and either fluid loading (21%) or withholding (3%) [37]. A third study found management change in 41%, with 34% in patients with adequate versus 58% in patients with inadequate clinical data [55]. The changes were considered minor in the majority of cases, like addition of medication in 61%, discontinuation in 16%, initiation of new investigation in 11%, and consultation in 9%, respectively. Interestingly, in this study CCUS was even ordered in 73% of patients in whom a diagnosis and management plan had already been formulated by clinical assessment and physiologic measurements [55]. One study observed that a significant increase in CCUS use was associated with a significant decrease in echocardiography use without an increase in adverse outcomes [58,59]. In summary, mainly cohort studies provided observational data on the potential beneficial effects of CCUS on decision making in the critical care [17,18,41,60–63] (Table 1). CCUS is associated with earlier diagnosis, changes in diagnosis, and even change in management.

The question remains whether changes in management of individuals are associated with beneficial outcome [64]. After more than 30 years of availability of CCUS in the critical care there have not been any large, high-quality randomized trials performed to compare its use on patient relevant outcomes. The probable reason, no large randomized trial has been performed, is the difficulty to design a randomized trial to evaluate a(ny) technique on clinical decision-making. Moreover, the lack of well defined and proven treatment targets to be achieved and the lack of solid evidence for the individual treatments options, like fluid therapy, vasopressors, and inotropes hamper research [65].

To our opinion, randomized trials investigating the effect of solid clinical examination, complementary CCUS (or other more advanced hemodynamic monitoring), and evidence based, personalized treatment algorithms are needed. Currently, several study protocols on the role of CCUS are published [66–68].

CONSIDERATIONS FOR PRACTICE

Many clinicians apply CCUS in the ICU irrespective of the evidence. If the majority of all hemodynamic variables suggest one specific hemodynamic pattern or diagnosis, there seems no additive value associated with the use of CCUS. For instance, if a patient is coughing, has high fever, a low BP, tachycardia, warm extremities, mottling, hyperlactemia, oliguria, tachypnea, low SpO2, rhonchi over right thorax, and a chest X-ray showing an infiltrate, septic shock would probably be diagnosed by nearly 100% of all clinicians. Although some studies showed that diagnoses changed after CCUS, changes were mostly minor. In contrast, when the hemodynamic profile is less clear and/or some hemodynamic variables are seemingly in conflict with the diagnosis, or when treatments (i.e., vasopressors, inotropes, b-blockers) do not result in expected improvements, additional diagnostic testing is warranted for setting the correct diagnosis. For example, a patient with a history of MI on b-blocker therapy might present with confusion, low BP, a heart rate of 80 beats/min, normal skin temperature, oliguria, and prolonged capillary refill time. What is the diagnosis and which treatment should be initiated? Would CCUS be of any added value here? Opponents of CCUS would argue for obvious cardiac dysfunction post-MI. Proponents of CCUS would argue that data conflict and additional information on cardiac size, cardiac function, abnormalities, and a CO measurement would be informative for the correct diagnosis and optimal treatment. Should the clinician start inotropes and/or vasopressors without additional hemodynamic test? What should be the mean arterial pressure goal? These questions are among the most frequent in the ICU. Unfortunately, clinical profiles of patients are mostly less clear as patients frequently present with combinations of types of shock or their hemodynamic profiles develop during their sickness (i.e., any distributive shock because of sepsis

may evolve to cardiogenic shock when septic cardiomyopathy develops). The awareness that potentially several medical problems are concurrently present in one patient is often neglected [69,70]. Clinicians need education on how and when to use and interpret CCUS and need to be aware of the learning curve. Surprisingly only few studies on CCUS had their images judged by (independent) experienced cardiologists for quality control [40,42,55,60,62]. Especially given the doubts that exist with regard to the noncardiologist operator's expertise. The clinician using CCUS should realize that image acquisition and interpretation is more challenging compared to ultrasound in elective settings so that abnormalities can easily be missed [18,71]. Further, the clinician should be aware of the limitations associated with use of CCUS measurements concerning reliability, accuracy, and potential confounders (e.g., atrial fibrillation or aortic insufficiency/stenosis) just like any other diagnostic testing. One classic example of such a limitation is negative findings for tamponade by CCUS in case of circulatory shock after cardiothoracic surgery. TOE or PAC can be considered for advanced diagnostic testing when CCUS renders insufficient views or suboptimal quality [22,24,72].

TABLE 1. Critical care ultrasonography for hemodynamic monitoring on the ICU

First author + publication year	Centre	Patients	Eligible	CTS	MV	No of US	Operator	Population	Protocol
Retrospective cohort									
Vignon 1994	2	111	ns	16,5%	100%	128	Intensivists	All MV patients with a clinical problem	ns
Bernier-Jean 2015	3	968	ns	ns	51%	1215	Trained members of ICU team	All patients whom ICU team members performed an echo; 73% on the ICU	General CCUS and critical care echocardiography
Yin 2018	1	451	ns	ns	86%	451	Board certificated physician	All patients whom ICU board certified team members performed an echo	5 different 'points of echo view' which included the lung ultrasound score
Prospective cohort									
Jensen 2004	1	210	2200	~55%	65%	233	ns	Patients not making clinical progress after cardiopulmonary evaluation based on conventional monitoring	FATE
Joseph 2004	1	100	ns	ns	70%	100	ns	Patients in shock not having CTS within 7 days of inclusion; 85% on the ICU/CCU	Full echocardiographic exam
Vignon 2004	1	55	ns	ns	73%	110	Operators with level III training	Consecutive patients who required TTE. Main reasons: shock 27% and pulmonary oedema/ARDS 20%	TTE versus HCU
Manasia 2005	1	90	ns	ns	ns	90	4 experienced cardiac sonographers	Patients admitted to surgical ICU (79%) or hemodynamically decompensating (21%)	Limited echo-imaging

Outcome						
Acoustic window	**Image quality$**	**Independent judging of images**	**Problem solving**	**Abnormalities detected**	**Change in clinical diagnosis**	**Change in clinical management**
ns	TTE: 55% good, 23% suboptimal, 22% poor	ns	Clinical problem solved – TTE: 38%	ns	ns	16%
ns	94% adequate (cardiac views)	Blinded evaluated random sample of 100 images to ensure quality	Cardiac US large impact on diagnosis, cardiac US largest impact on management when compared to other US	ns	25% in all US versus 37% of cardiac US	44% in all US versus 58% of cardiac US
ns	ns	Results were doublechecked by other senior physicians and in case of 'abdnormal' images the examination was double checked with other phycisians and pathologists		34% LV dysfunction (10% moderate; 3% severe), 32% abnormal RV (moderate 26%; severe 6%), 17% pericardial effusion; 14% regional wall abnormalities, 7% pneumothorax	ns	ns
0/3 (2%), 1/3 (23%), 2/3 (41%), 3/3 (34%)	97%	ns	ns	ns	ns	ns
Parasternal views were generally poor	99%	Two board certified echocardiographers (94% agreement; k score of 0.86)	PPV of 98% and NPV 100% for detection of a cardiac cause of shcok	33% severe LV dysfunction and 17% tamponade	ns	51% total and 57% of total: medical therapy
5% no view	ns	ns	Diagnostic accuracy 90% for TTE compared with 80% for HCU	ns	ns	51% for TTE, 49% for HCU
99% succesfull	94% successful	Cardiologist	Valuable in 48% and no use in 15%	ns	ns	37%

First author + publication year	Centre	Patients	Eligible	CTS	MV	No of US	Operator	Population	Protocol
Stanko 2005	1	126	633	ns	ns	135	Trained members of ICU team (mostly residents and fellows)	Main indications: patients with hemodynamic compromise (64%) or an assessment of LV function (12%)	Full echocardiographic exam
Orme 2009	1	217	1576	ns	64%	187	Intensivists (83%)	Main indications: assessment of LV function (46%), search for cause of hypotension (17%) or pulmonary oedema (14%)	Full echocardiographic exam
Marcelino 2009	1	704	All consecutive	ns	1	682	Intensivist	All admitted patients	Full echocardiographic exam

Legend. CCU = coronary care unit, CCUS = critical care ultrasonography, CTS = cardiothoracic surgery, CVP = central venous pressure, FATE = focus assessed TTE; FICE = focused intensive care echo, HCU = hand-carried ultrasound, ICU = Intensive Care Unit, LV = left ventricle, MV = mechanical ventilation, No = number, ns = not specified, Ref = reference, RV = right ventricle, SB = spontaneously breathing, pts = patients, TTE = transthoracic echocardiography; US = ultrasound. $ Assessed as adequate.

Outcome						
Acoustic window	**Image quality$**	**Independent judging of images**	**Problem solving**	**Abnormalities detected**	**Change in clinical diagnosis**	**Change in clinical management**
ns	10% suboptimal	Two board certified cardiologist	ns	ns	29%	41%
ns	TTE in MV: 27% good, 62% adequate, TTE in SB: 27% good, 64% adequate	In case of doubt lead author or cardiologist	Support clinical decision-making 10%	ns	ns	49% total, 42% immediate change in treatment
0,6% no data, 5.4% no para-sternal view, 19% no IVC	ns	Whenever needed by a senior cardiologist (n=14)	ns	33% total, 7.5% severe previously unknown	ns	ns

Conclusion

CCUS is a highly available, widely used tool for diagnosis and hemodynamic monitoring. It seems irrelevant that CCUS-derived measurements are not highly correlated with PAC-derived values, provided that the operator is aware of the diagnostic test characteristics of CCUS. Both tools, CCUS and PAC, should be considered of potential additional value depending on the specific clinical question. Optimal use of CCUS includes repeated measures for monitoring and guiding treatment. Future studies on CCUS should focus on the diagnostic accuracy of different ultrasound applications [73] and on patient outcome [74]. Eventually, results of current and future studies should establish a more defined role for CCUS.

References

1. Sakr Y, Reinhart K, Vincent JL, et al. Does dopamine administration in shock influence outcome? Results of the Sepsis Occurrence in Acutely Ill Patients (SOAP) Study. Crit Care Med 2006; 34:589–597.
2. Cecconi M, De Backer D, Antonelli M, et al. Consensus on circulatory shock and hemodynamic monitoring. Task force of the European society of intensive care medicine. Intensive Care Med 2014; 40:1795–1815.
3. Weil MH. Personal commentary on the diagnosis and treatment of circulatory shock states. Curr Opin Crit Care 2004; 10:246–249.
4. De Backer D, Biston P, Devriendt J, et al. Comparison of dopamine and norepinephrine in the treatment of shock. N Engl J Med 2010; 362:779–789.
5. Hiemstra B, Eck RJ, Keus F, van der Horst IC. Clinical examination for diagnosing circulatory shock. Curr Opin Crit Care 2017; 23: article in press.
6. Saugel B, Ringmaier S, Holzapfel K, et al. Physical examination, central venous pressure, and chest radiography for the prediction of transpulmonary thermodilution-derived hemodynamic parameters in critically ill patients: a prospective trial. J Crit Care 2011; 26:402–410.
7. Hamzaoui O, Scheeren TW, Teboul JL. Norepinephrine in septic shock: when and how much? Curr Opin Crit Care 2017.
8. Joly HR, Weil MH. Temperature of the great toe as an indication of the severity of shock. Circulation 1969; 39:131–138.
9. Clement RP, Vos JJ, Scheeren TW. Minimally invasive cardiac output technologies in the ICU: putting it all together. Curr Opin Crit Care 2017; 23:article in press.
10. Saugel B, Reese PC, Wagner JY, et al. Advanced hemodynamic monitoring in intensive care medicine: a German web-based survey study. Med Klin Intensivmed Notfmed 2017. A survey study among 284 German intensivists showing that advanced hemodynamic assessment by CCUS is considered necessary in the diagnosis and management of circulatory shock.
11. Volpicelli G, Lamorte A, Tullio M, et al. Point-of-care multiorgan ultrasonography for the evaluation of undifferentiated hypotension in the emergency department. Intensive Care Med 2013; 39:1290–1298.
12. Strumwasser A, Frankel H, Murthi S, et al., American Association for the Surgery of Trauma Committee on Critical Care. Hemodynamic monitoring of the injured patient: From central venous pressure to focused echocardiography. J Trauma Acute Care Surg 2016; 80:499–510.
13. Oren-Grinberg A, Talmor D, Brown SM. Focused critical care echocardiography. Crit Care Med 2013; 41:2618–2626.
14. Whitson MR, Mayo PH. Ultrasonography in the emergency department. Crit Care 2016; 20:227.
15. McLean AS. Echocardiography in shock management. Crit Care 2016;20:275.
16. Boyd JH, Walley KR. The role of echocardiography in hemodynamic monitoring. Curr Opin Crit Care 2009; 15:239–243.
17. Heiberg J, El-Ansary D, Canty DJ, et al. Focused echocardiography: a systematic review of diagnostic and clinical decision-making in anaesthesia and critical care. Anaesthesia 2016; 71:1091–1100.
18. Brown SM, Kasal J. Bedside ultrasound in the intensive care unit: where is the evidence? Semin Respir Crit Care Med 2015; 36:878–889.

19. Goodwin M, Ito K, Gupta AH, Rivers EP. Protocolized care for early shock resuscitation. Curr Opin Crit Care 2016; 22:416–423.
20. Connors AF, Speroff T, Dawson NV, et al. The effectiveness of right heart catheterization in the initial care of critically ill patients. SUPPORT investigators. JAMA 1996; 276:889–897.
21. Harvey S, Harrison DA, Singer M, et al. Assessment of the clinical effectiveness of pulmonary artery catheters in management of patients in intensive care (PAC-Man): a randomised controlled trial. Lancet 2005; 366:472–477.
22. Weil MH. The assault on the Swan-Ganz catheter: a case history of constrained technology, constrained bedside clinicians, and constrained monetary expenditures. Chest 1998; 113:1379–1386.
23. Marik PE. Obituary: pulmonary artery catheter 1970 to 2013. Ann Intensive Care 2013; 3:38.
24. Reade MC, Angus DC. PAC-Man: game over for the pulmonary artery catheter? Crit Care 2006; 10:303.
25. Hyzy RC. Bedside ultrasonography for diagnosis of septic shock. JAMA 2016; 315:89.
26. Wetterslev M, Haase N, Johansen RR, Perner A. Predicting fluid responsiveness with transthoracic echocardiography is not yet evidence based. Acta Anaesthesiol Scand 2013; 57:692–697.
27. Wetterslev M, Moller-Sorensen H, Johansen RR, Perner A. Systematic review of cardiac output measurements by echocardiography vs. thermodilution: the techniques are not interchangeable. Intensive Care Med 2016; 42:1223–1233.
28. Marcelino PA, Marum SM, Fernandes AP, et al. Routine transthoracic echocardiography in a general intensive care unit: an 18 month survey in 704 patients. Eur J Intern Med 2009; 20:e37–e42.
29. Jensen MB, Sloth E, Larsen KM, Schmidt MB. Transthoracic echocardiography for cardiopulmonary monitoring in intensive care. Eur J Anaesthesiol 2004; 21:700–707.
30. Rhodes A, Evans LE, Alhazzani W, et al. Surviving sepsis campaign: international guidelines for management of sepsis and septic shock: 2016. Intensive Care Med 2017; 43:304–377.
31. Kimura BJ. Point-of-care cardiac ultrasound techniques in the physical examination: better at the bedside. Heart 2017.
32. Ahn JH, Jeon J, Toh HC, et al. SEARCH 8Es: a novel point of care ultrasound protocol for patients with chest pain, dyspnea or symptomatic hypotension in the emergency department. PLoS One 2017; 12:e0174581.
33. Mok KL. Make it SIMPLE: enhanced shock management by focused cardiac ultrasound. J Intensive Care 2016; 4:51.
34. Walley PE, Walley KR, Goodgame B, et al. A practical approach to goaldirected echocardiography in the critical care setting. Crit Care 2014;18:681.
35. Miller A, Mandeville J. Predicting and measuring fluid responsiveness with echocardiography. Echo Res Pract 2016; 3:G1–G12.
36. Feissel M, Michard F, Mangin I, et al. Respiratory changes in aortic blood velocity as an indicator of fluid responsiveness in ventilated patients with septic shock. Chest 2001; 119:867–873.
37. Orme RM, Oram MP, McKinstry CE. Impact of echocardiography on patient management in the intensive care unit: an audit of district general hospital practice. Br J Anaesth 2009; 102:340–344.
38. Caronia J, Kutnick R, Sarzynski A, et al. Focused transthoracic echocardiography performed and interpreted by medical residents in the critically ill. ICU Dir 2013; 4:177–182.

39. Neskovic AN, Edvardsen T, Galderisi M, et al., European Association of Cardiovascular Imaging Document Reviewers. Focus cardiac ultrasound: the European Association of Cardiovascular Imaging viewpoint. Eur Heart J Cardiovasc Imaging 2014; 15:956–960.
40. Manasia AR, Nagaraj HM, Kodali RB, et al. Feasibility and potential clinical utility of goal-directed transthoracic echocardiography performed by noncardiologist intensivists using a small hand-carried device (SonoHeart) in critically ill patients. J Cardiothorac Vasc Anesth 2005;19:155–159.
41. Bernier-Jean A, Albert M, Shiloh AL, et al. The diagnostic and therapeutic impact of point-of-care ultrasonography in the intensive care unit. J Intensive Care Med 2017; 32:197–203.
42. Joseph MX, Disney PJ, Da Costa R, Hutchison SJ. Transthoracic echocardiography to identify or exclude cardiac cause of shock. Chest 2004;126:1592–1597.
43. Hall DP, Jordan H, Alam S, Gillies MA. Intensive Care Echo protocol on the management of critically ill patients, and comparison with full echocardiographic studies by BSE-accredited sonographers. Hosp Med Clin 2017;EPub ahead of print.
44. Otto CM. Echocardiography: the transition from master of the craft to admiral of the fleet. Heart 2016; 102:899–901.
45. Blanco P, Volpicelli G. Common pitfalls in point-of-care ultrasound: a practical guide for emergency and critical care physicians. Crit Ultrasound J 2016;8:15.
46. Lewis JF, Kuo LC, Nelson JG, et al. Pulsed Doppler echocardiographic determination of stroke volume and cardiac output: clinical validation of two new methods using the apical window. Circulation 1984; 70:425–431.
47. Coats AJ. Doppler ultrasonic measurement of cardiac output: reproducibility and validation. Eur Heart J 1990; 11 (Suppl I):49–61.
48. Mercado P, Maizel J, Beyls C, et al. Transthoracic echocardiography: an accurate and precise method for estimating cardiac output in the critically ill patient. Crit Care 2017; 21:136.
49. Corl KA, George NR, Romanoff J, et al. Inferior vena cava collapsibility detects fluid responsiveness among spontaneously breathing critically-ill patients. J Crit Care 2017; 41:130–137.
50. Lee CW, Kory PD, Arntfield RT. Corrigendum 'Development of a fluid resuscitation protocol using inferior vena cava and lung ultrasound' journal of critical care 31(2016) 96-100. J Crit Care 2016; 33:267.
51. Jones AE, Tayal VS, Sullivan DM, Kline JA. Randomized, controlled trial of immediate versus delayed goal-directed ultrasound to identify the cause of nontraumatic hypotension in emergency department patients. Crit Care Med 2004; 32:1703–1708.
52. Hall MK, Taylor RA, Luty S, et al. Impact of point-of-care ultrasonography on ED time to disposition for patients with nontraumatic shock. Am J Emerg Med 2016; 34:1022–1030.
53. Alsaddique A, Royse AG, Royse CF, et al. Repeated monitoring with transthoracic echocardiography and lung ultrasound after cardiac surgery: feasibility and impact on diagnosis. J Cardiothorac Vasc Anesth 2016; 30:406–412.
54. Manno E, Navarra M, Faccio L, et al. Deep impact of ultrasound in the intensive care unit: the 'ICU-sound' protocol. Anesthesiology 2012; 117:801–809.
55. Stanko LK, Jacobsohn E, Tam JW, et al. Transthoracic echocardiography: impact on diagnosis and management in tertiary care intensive care units. Anaesth Intensive Care 2005; 33:492–496.
56. Cortellaro F, Ferrari L, Molteni F, et al. Accuracy of point of care ultrasound to identify the source of infection in septic patients: a prospective study. Intern Emerg Med 2017; 12:371–378.

57. Vignon P, Mentec H, Terre S, et al. Diagnostic accuracy and therapeutic impact of transthoracic and transesophageal echocardiography in mechanically ventilated patients in the ICU. Chest 1994; 106:1829–1834.

58. Shokoohi H, Boniface KS, Pourmand A, et al. Bedside ultrasound reduces diagnostic uncertainty and guides resuscitation in patients with undifferentiated hypotension. Crit Care Med 2015; 43:2562–2569.

59. Alherbish A, Priestap F, Arntfield R. The introduction of basic critical care echocardiography reduces the use of diagnostic echocardiography in the intensive care unit. J Crit Care 2015; 30:1419.e7–1419.e11.

60. Sekiguchi H, Harada Y, Villarraga HR, et al. Focused cardiac ultrasound in the early resuscitation of severe sepsis and septic shock: a prospective pilot study. J Anesth 2017;31(4):487-493. doi: 10.1007/s00540-017-2312-8.

61. Kanji HD, McCallum J, Sirounis D, et al. Limited echocardiography-guided therapy in subacute shock is associated with change in management and improved outcomes. J Crit Care 2014; 29:700–705.

62. Silva S, Biendel C, Ruiz J, et al. Usefulness of cardiothoracic chest ultrasound in the management of acute respiratory failure in critical care practice. Chest 2013; 144:859–865.

63. Vignon P, Frank MB, Lesage J, et al. Hand-held echocardiography with Doppler capability for the assessment of critically-ill patients: is it reliable? Intensive Care Med 2004; 30:718–723.

64. Saugel B, Vincent JL, Wagner J. Personalized hemodynamic management. Curr Opin Crit Care 2017; 23:; article in press.

65. Angus DC, Barnato AE, Bell D, et al. A systematic review and meta-analysis of early goal-directed therapy for septic shock: the ARISE, ProCESS and ProMISe Investigators. Intensive Care Med 2015; 41:1549–1560.

66. ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). Identifier NCT02974790, Diagnostic Performance of Echocardiography Performed by Emergency Physicians After a Basic Training (ECHOBEDSIDE). 2017; 2017.

67. ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). Identifier NCT03082326, The CCUS Based Characteristic of Hemodynamic and Lung Pathology in Early Stage of Shock in ICU: The Epidemic and Prognostic Value. 2017; 2017.

68. ClinicalTrials.gov [Internet]. Bethesda (MD): National Library of Medicine (US). Identifier NCT03093987, Critical Care Ultrasound Oriented Shock Treatment in ICU. 2017; 2017.

69. Dahake VT, Khedekar K, Faizal F, Majethia N. The role of autopsy in medical intensive care unit: comparison of clinical and postmortem diagnoses. J Evolution Med Dent Sci 2016; 5:1038–1042.

70. Roosen J, Frans E, Wilmer A, et al. Comparison of premortem clinical diagnoses in critically ill patients and subsequent autopsy findings. Mayo Clin Proc 2000; 75:562–567.

71. Colebourn C, Barber V, Salmon JB, Young D. The accuracy of diagnostic and haemodynamic data obtained by transthoracic echocardiography in critically ill adults: a systematic review. J Intensive Care Society 2008; 9:128–134.

72. Vignon P. Hemodynamic assessment of critically ill patients using echocardiography Doppler. Curr Opin Crit Care 2005; 11:227–234.

73. Favot M, Courage C, Ehrman R, et al. Strain echocardiography in acute cardiovascular diseases. West J Emerg Med 2016; 17:54–60.

74. Mayo P, Arntfield R, Balik M, et al. The ICM research agenda on critical care ultrasonography. Intensive Care Med 2017; 43(9):1257-1269. doi: 10.1007/s00134-017-4734-z

6

Feasibility of cardiac output measurements in critically ill patients by medical students

Koster G, Kaufmann T, Hiemstra B, Wiersema R, Vos ME, Dijkhuizen D, Wong A, Scheeren TWL, Hummel YM, Keus F, van der Horst ICC.
Ultrasound J 2020;12:1. https://doi.org/10.1186/s13089-020-0152-5

Abstract

Background

Critical care ultrasonography (CCUS) is increasingly applied also in the intensive care unit (ICU) and performed by non-experts, including even medical students. There is limited data on the training efforts necessary for novices to attain images of sufficient quality. There is no data on medical students performing CCUS for the measurement of cardiac output (CO), a hemodynamic variable of importance for daily critical care. The aim of this study was to explore the agreement of cardiac output measurements as well as the quality of images obtained by medical students in critically ill patients compared to the measurements obtained by experts in these images.

Methods

In a prospective observational cohort study, all acutely admitted adults with an expected ICU stay over 24 hours were included. CCUS was performed by students within 24 hours of admission. CCUS included the images required to measure the CO, i.e., the left ventricular outflow tract (LVOT) diameter and the velocity time integral (VTI) in the LVOT. Echocardiography experts were involved in the evaluation of the quality of images obtained and the quality of the CO measurements.

Results

There was an opportunity for a CCUS attempt in 1155 of the 1212 eligible patients (95%) and 1075 of the 1212 patients (89%) CCUS examination was performed by medical students. In 871 out of 1075 patients (81%) medical students measured CO. Experts measured CO in 783 patients (73%). In 760 patients (71%) CO was measured by both which allowed for comparison; bias of CO was 0.0 L·min-1 with limits of agreement of -2.6 L·min-1 to 2.7 L·min-1. The percentage error was 50%, reflecting poor agreement of the CO measurement by students compared with the experts CO measurement.

Conclusions

Medical students seem capable of obtaining sufficient quality CCUS images for CO measurement in the majority of critically ill patients. Measurements of CO by medical students, however, had poor agreement with expert measurements. Experts remain indispensable for reliable CO measurements.

Background

Critical care ultrasonography (CCUS) is a deliberately focused examination, aimed at rapidly answering straightforward clinical questions [1]. In the field of emergency and critical care medicine, CCUS is increasingly used to guide interventions in critically ill patients in various settings by experts and novices [2–14]. The training process required for users to attain competency in CCUS has varied widely between studies, reflecting the diversity in CCUS training between centers. Similarly, there is variability among statements from stakeholders regarding the type of training, the required number of hours spent and examinations performed by the trainee to achieve competency in CCUS [15–17]. However, besides these disparities, individual physicians struggle with barriers to its use, such as perceived difficulty in obtaining adequate technical skills [13], limitations in training, need (perceived and real), and costs [6, 14].

One valuable CCUS hemodynamic measurement is the determination of the cardiac output (CO), especially if the patient is in circulatory shock [18]. Circulatory shock occurs in one-third of patients admitted to the ICU [19], so being able to perform CCUS and measure CO is of importance. However, CO measurement by CCUS is considered an advanced level CCUS skill [20, 21]. Whether trained novices (e.g. medical students or other less experienced physicians) are able to obtain reliable CO measurements has not yet been investigated. In a convenience sample of 100 adult patients in the emergency department (ED), two ultrasound-naive ED physicians were able to measure CO by ultrasonography accurately [22]. Another study in the ED with a convenience sample of 80 patients, however, showed poor agreement in CO measurement by an emergency ultrasound fellow compared to an emergency cardiology fellow [23]. At the start of our study there were no data on medical students performing CO measurements by CCUS in critically ill patients, although medical students have been shown to be capable of performing CCUS after limited training [24]. To our knowledge, only one small study investigated CCUS by medical students on a (cardiac) intensive care unit, and CO was not measured (see supplements) [3].

The aim of this study was to explore the feasibility of a limited CCUS examination, consisting of CO measurements, performed by medical students in a protocolised manner, in critically ill patients. In addition, the quality of images required to calculate CO and the accuracy of CO measurements compared to those obtained by echocardiography experts were analyzed.

Methods

The Simple Intensive Care Studies (SICS)-I was a prospective, observational cohort study which followed a published protocol and statistical analysis plan (Clinicaltrials.gov; NCT02912624). The SICS-I was developed to unravel the diagnostic and prognostic value of a comprehensive selection of clinical, hemodynamic, and biochemical variables in critically ill patients, and details have been described elsewhere [25, 26]. All acutely admitted adults with an expected ICU stay over 24 hours were included. Patients were excluded when admission was planned and if clinical care interfered with acquiring research data (e.g., mechanical circulatory support). The local institutional review board approved the study (M15.168207).

Data collection and training

All patients underwent CCUS within 24 hours of ICU admission. Detailed information on the CCUS performed can be found in the supplements. Patients were enrolled by fourth-year to sixth-year medical students of a six-year medical school program. The training consisted of self-study on theoretical fundamentals and two practical sessions of at least two hours in total to learn how to operate the General Electric Vivid-S6 mobile ultrasonography machine using the cardiac phased-array probe (see appendix in supplements for detailed information). The theoretical self-study on how to perform CCUS and measure the CO consisted of study of the protocol (supplements), a website on the principles of echocardiography [27], and international guidelines [28, 29]. This information became available two weeks before participation of the medical students. During the practical sessions, medical students learned to obtain the parasternal long axis (PLAX), apical four-chamber (AP4CH), and apical five-chamber (AP5CH) views, among others. The medical students alternated with obtaining the views and measurements of CO during the practical sessions. All medical students received at least two hours hands-on training from cardiologist-intensivists (GK and IVDH). Views and images were obtained randomly during the respiratory cycle and/or phase of mechanical ventilation. In case of any arrhythmias, the average of multiple measurements over five heartbeats was taken.

The first 20 CCUS images and measurements of each medical student were supervised by medical students who had independently performed more than 50 CCUS examinations. After 20 scans, CCUS medical students were allowed to conduct/perform CCUS unsupervised, since previous studies showed acceptable capability for acquiring images beyond 20 exams [30].

Validation and definitions

For quality control, echocardiography technicians from an independent core laboratory (Groningen Image Core Lab, UMCG, Groningen, the Netherlands, www.g-icl.com) assessed all CCUS images and measurements obtained by the medical students according to the study protocol. If the images were obtained according to guideline standards, the LVOTd and VTI were independently remeasured and CO recalculated [28, 29]. Core laboratory technicians, which we refer to as experts throughout this report, were blinded to all other clinical measurements. The experts did not perform any CCUS examination.

Outcomes, index test and reference standard

The number of patients where CCUS could not be performed and reasons for unobtainable images by the medical students were reported. Patients were excluded from the analysis if, for research purposes, experts would also not be able to perform CCUS (i.e. drains, subcutaneous emphysema, surgical dressing/wounds). The number of patients in which CCUS images of PLAX or AP5CH were obtained was analyzed [28, 29]. Proportion of patients was reported wherein the CCUS images assessed by the experts was of insufficient quality for CO measurement.

We also evaluated the accuracy of CO measurements by medical students (COmedical student) compared to CO measurements by experts (COexpert). Moreover, the two components needed for CO calculation (i.e. LVOTd or VTI) were assessed to determine possible differences between medical students' and experts' measurements.

Sensitivity analyses were done with baseline characteristics to investigate reasons why experts could not measure a CO.

Sample size and missing data

Due to the observational nature of this study, no formal power calculation was performed. For the accuracy analysis on CO measurements, we only included patients if CO was measured by both medical students and experts.

Statistical analysis

Data were presented as mean with standard deviation (SD) when normally distributed or as median with interquartile ranges (IQR) in case of skewed data. Dichotomous and categorical data were presented in proportions. Intraclass correlation coefficients (ICC) were calculated to assess the concordance between

the measurements made by the medical students and the experts. Bland-Altman analysis was performed to assess agreement of medical student versus expert measurements by calculating mean and SD of the differences, the 95% limits of agreement (LOA) (= mean of the difference ± 1.96 × SD of the difference), and the percentage error [31]. In method comparison studies, a percentage error of 30% is considered acceptable if the error of the test and the reference method is 20%, which is the case when using the thermodilution method to calculate CO [32]. Since there is no reference for CCUS, and only one method was used with comparison between the observers, a percentage error of less than 20% was defined as clinically acceptable. This would mean that the CO difference between medical students and experts would be less than 0.5 L·min-1 in the lower end of the CO spectrum (e.g. when the experts measured a CO of 2.5 L·min-1, a CO of 2.0 – 3.0 L·min-1 by the medical student would be clinically acceptable). An alpha error of 0.05 was used to indicate statistical significance. Statistical analyses were conducted using STATA version 15.0 (StataCorp, College Station, USA).

Results

CCUS acquisition and images

Between March 27th, 2015 and July 22nd, 2017, sixteen medical students were involved in the study and 1212 patients fulfilled inclusion criteria. Of these, in a total of 1155 patients CCUS was performed, as in 40 patients there was interference with clinical care during the first 24 hours of admission (e.g., the patient was in severe hemodynamic instability or an intervention was being performed) and 17 patients had isolation restriction measures. Of these 1155 patients, in 80 patients, clinical conditions (i.e., thoracic drains, wounds, or subcutaneous emphysema) prohibited the image acquisition by CCUS, leaving 1075 patients with ultrasonography data (Figure 1).

The medical students deemed both LVOTd and VTI unmeasurable (i.e., images were of too low a quality and no or few structures could be identified) in 129 patients (12%), the LVOTd in 46 patients (4.2%), and the VTI in 29 patients (2.6%). The parasternal short axis view did not provide any additional measurements when the LVOTd was unmeasurable in the PLAX view. Thus, 204 patients (19%) out of 1075 had no CO measurement, leaving a total of 871 patients (81%) with a measured CO by medical students.

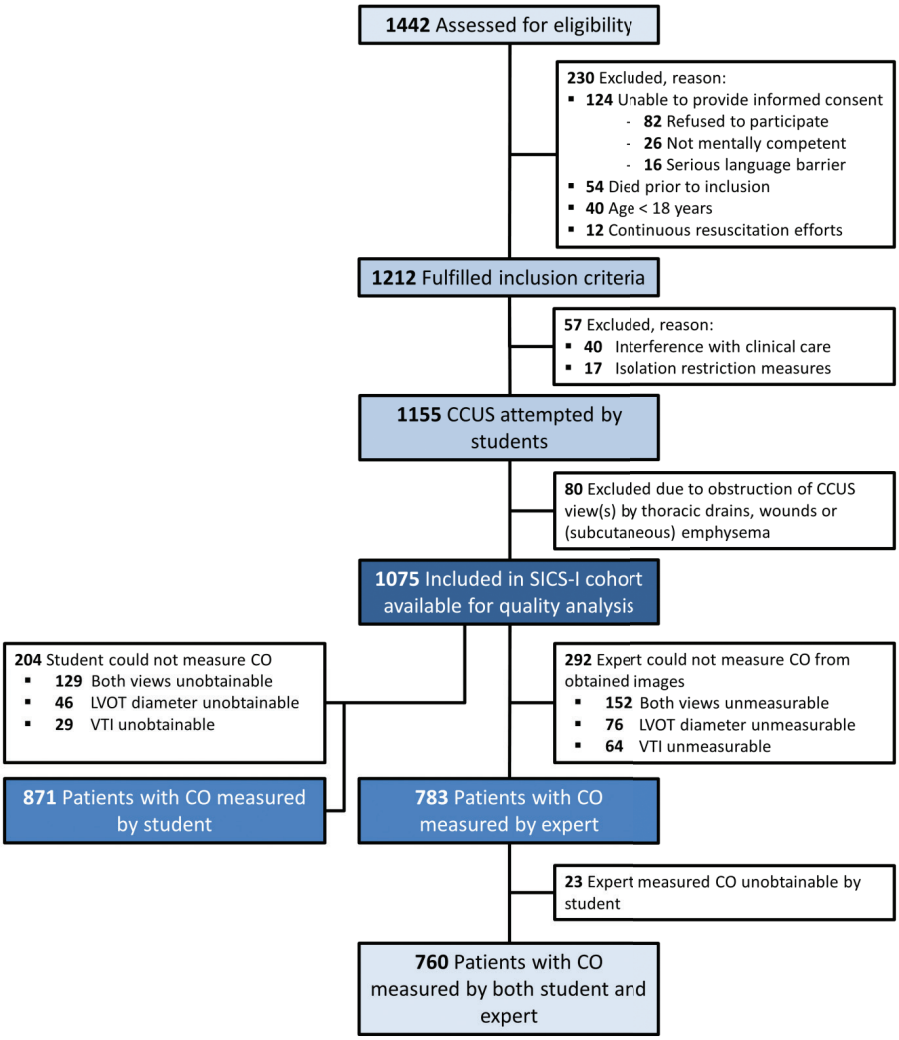


FIGURE 1. Flow diagram of the Simple Intensive Care Studies-I (SICS-I)
Legend. Abbreviations: ICU, intensive care unit; CCUS, critical care ultrasonography; CO, cardiac output; LVOT, left ventricular outflow tract; VTI, velocity time interval

CCUS quality of images

The experts used the images obtained by the medical students and were unable to measure both the LVOTd and VTI in 152/1075 (14%), LVOTd in 76/1075 (7.1%), and VTI in 64/1075 (6.0%). While the experts deemed more measurements to be impossible in the obtained images compared to the medical students, the experts

were also able to add 23 CO measurements in patients where medical students judged the images to be of too poor a quality and consequently did not perform the measurements. In total, the experts measured CO in 783 patients (73%). Comparisons of CO measurements by medical students and experts were possible in 760 (71%) out of 1075 patients in case of adequate image quality (Figure 1).
Differences in patient baseline characteristics were found between the group in which experts could measure a CO and the group in which experts could not measure a CO (see Table 1). Patients without CO measured by experts were characterized by older age, greater illness severity (reflected in higher APACHE IV scores), higher heart rate, greater prevalence of chronic obstructive pulmonary disease (COPD), higher rates of mechanical ventilation, greater likelihood of being post-operative, and higher vasopressor dose

Comparison of CO measurement by medical students and experts

The mean COmedical student was 5.2 ± 2.0 L·min-1 and COexpert was 5.2 ± 1.8 L·min-1 (p=0.44). Bland-Altman analysis demonstrated a bias of -0.0 L·min-1 (95% CI -0.06 – 0.13) with limits of agreement of -2.6 L·min-1 (95% CI -2.7 – -2.4) to 2.7 L·min-1 (95% CI 2.5 – 2.8) (Figure 2). Plotting a regression line in the Bland-Altman plot showed a proportional bias of 2%. The percentage error was 50% (95% CI 47 – 53). The ICC was 0.75 (95% CI 0.72 – 0.78).

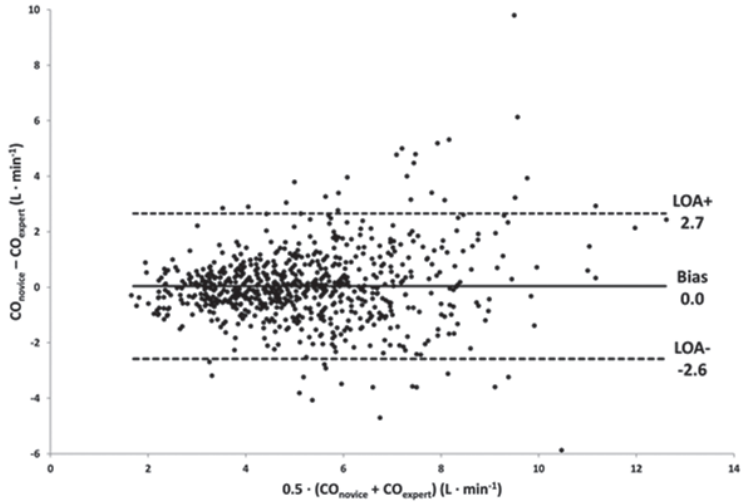


FIGURE 2. Bland-Altman plot showing the comparison between cardiac output measured by medical students (COmedical student) and core lab experts (COexpert).
Legend. The mean bias between COexpert and COmedical student and the upper and lower limits of agreement (LOA) are presented. The figure clearly shows the widening of the LOA in both directions with increasing CO.

Comparison of LVOTd and VTI measurements by medical students and experts

The medical students measured 900 LVOTd and the experts 847. There were 815 paired LVOTd measurements. Mean LVOTd by medical students (LVOTdmedical student) was 2.06 ± 0.24, whereas the mean of the LVOTd measured by experts (LVOTdexpert) was 2.09 ± 0.18 ($p<0.001$). Bland-Altman analysis showed a bias of 0.0 cm (95% CI 0.0 – 0.0) with limits of agreement of -0.5 cm (95% CI -0.5 – -0.4) to 0.4 cm (95% CI 0.4 – 0.4) (see supplements). The percentage error was 21% (95% CI 20 – 23). There was a proportional bias of 20% (0.41cm). The ICC was 0.43 (95% CI 0.37 – 0.48).

The medical students measured 917 VTI and the experts 859. There were 840 paired VTI measurements. Mean VTI by medical students (VTImedical student) was 19.0 ± 5.6 cm compared to 18.5 ± 5.4 cm of the experts (VTIexpert) ($p<0.001$). Bland-Altman analysis showed a bias of 0.5 cm (95% CI 0.4 – 0.7) with limits of agreement of -5.0 cm (95% CI -5.3 – -4.6) to 6.1 cm (95% CI 5.7 – 6.4) (see supplements). The percentage error was 30% (95% CI 28 – 31). The ICC was 0.86 (95% CI 0.84 – 0.88).

Discussion

In this large prospective ICU cohort study with CCUS, we found that, after dedicated training, medical students were able to acquire a CO measurement in three out of every four patients (871 of 1155 patients). This finding is of interest considering that the medical students were ultrasound naïve, the CO measurement is considered an advanced CCUS skill, and the ICU population is known for technical difficulties in acquiring ultrasound images. In a minority of ICU patients (80 of the 1155 patients) CCUS was not possible due to clinical conditions hampering image acquisition, leaving 1075 patients with ultrasonography data. The CCUS images obtained by medical students were assessed by experts and rated to be of adequate quality in 73%. Patients (292 of 1075 patients) in which no adequate image quality could be obtained were more often mechanically ventilated, admitted after cardiothoracic surgery or were more severely ill.

Although the students reached a reasonable percentage on image acquisition/ quality, our data do not support CO measurements by medical students (after limited training), as comparison to CO measurements by experts showed poor agreement. CCUS concerns more than acquiring the required images and any operator should be aware of the potential errors that can be made with ultrasonography, especially in complex critically ill patients [33]. It is important to note that education on ultrasonography should focus on specific training and quality control on all aspects of ultrasonography in order to achieve accurate measurements [17]. Our results are in line with recommendations by the European

Association of Cardiovascular Imaging (EACVI) on point-of-care, problem-oriented focus cardiac ultrasound examination (FoCUS), which state that supervision and quality control by experts are essential for proper and complete examination. Quality control in our study was performed by an accredited echocardiographic laboratory as is recommended in this viewpoint [15].

To be able to compare our results to those of other studies, it is of utmost importance that every step, from eligible patients to the number of patients in which a reliable CO measurement by CCUS is obtained, are presented. Currently these numbers are often lacking, and this leads to varying success rates on the feasibility of CCUS. If reported, results may vary based on differences in ultrasonography training and experience, which impedes a comparison of image acquisition and quality. We found four studies, on measuring CO in critically ill patients by non-experts to compare with our study (see supplements) [22, 23, 34, 35]. In two out of the three studies the operators had previous experience with ultrasonography, but training varied [23, 34]. The setting, sampling, and exclusion criteria may explain the reported high success rate in one study over another [22]. Whether images obtained are of sufficient quality should preferably be judged by independent experts, as two out of three studies did [22, 34]. In one study independent investigators assessed the quality, however, it is not clear if these were experts or not [35]. The percentage of adequate/good-quality images in our study was comparable with Dinh et al. In the study of Betcher et al. and Villavicencio et al., image quality was generally (judged) overall lower. Duration of training or differences in baseline characteristics might explain part of these differences.

The final step to obtain a reliable CO measurement is to measure LVOTd and VTI on images of sufficient quality. Dinh et al. and Lee et al. reported data on measurement quality, and, furthermore, Dinh et al. reported a low bias between sonographers and independent experts. These studies and ours showed lack in precision for CO measurement by novices. Villavicencio et al. compared ultrasonography derived CO with the transpulmonary thermodilution technique and concluded that there was an acceptable level of agreement between the techniques. Furthermore, they found a high inter- and intra-observer reliability.

Ultrasonography in the acute setting remains challenging, and data regarding novice-based CCUS are limited (see supplements). In our study we chose for medical students as novices (i.e. non-experts), since non-experts constitute the majority of ultrasound trained personnel in an IC and as students would not interfere with daily ICU care. Five studies reported on medical students performing CCUS in critically ill patients (3 in ED setting, 1 in operating theatre and 1 in ICU) [3, 7–10]. Four out of the five studies showed that images could be acquired in a promising 82-98% of cases. The studies reporting on image quality showed

percentages of (at least) adequate imaging ranging from 89 to 98%, unfortunately by non-independent judging [3, 7]. Furthermore, after training, medical students can adequately interpret images with a very simplified or binary assessment [36]. A number of previous studies employed training curricula for medical students on ultrasonography protocols [37–39]. Four other studies used a point-of-care ultrasonography training program to determine diagnostic performance in various clinical scenarios [36, 40–42]. All studies showed feasibility to train medical students to perform ultrasonography after a relatively short amount of training, which is comparable to the training medical students received in our study.

In previous manuscripts on SICS study data we reported a higher percentage of images judged to be of sufficient quality [25, 26]. The current results showed the percentage of measurements of CO considered of sufficient quality by a core-laboratory and not images with a LVOT and VTI. The high(er) level of quality considered necessary is according to internal protocol and is independently monitored.

Limitations

First, the proportion of patients with an acoustic window was based on the results of CCUS by medical students only. We did not check if more experienced sonographers were able to retrieve images in these cases, because the design of our study was to obtain images outside patient care. We believe image quality can only be assessed if the observers are blinded for all other study data and are not involved in the patient's clinical care. Ideally, independent experts perform ultrasonography themselves and make a direct comparison with the medical student. The availability of time and staff outside clinical care in our center was limited, leading us to include all consecutive patients and allow trained medical students to run the study.

Second, we did not check for interindividual variation of skills and quality of CCUS in each medical student who participated in the study, mainly to limit the time of investigation at the bedside.

Third, CCUS of the heart was limited to 2D imaging of the LVOTd, - of the AP5CH and pulse wave Doppler imaging of the LVOT. Therefore, valvular disease could have been missed.

Conclusions

Medical students as novices were capable of performing CCUS with adequate image acquisition in the majority of an ICU population of acutely admitted critically ill patients. However, they cannot accurately measure a CCUS derived cardiac

output after limited training. Cardiac output measurements with CCUS in research and daily care should be interpreted with caution if not validated by experts; this is in concordance with the viewpoint of the EACVI on CCUS.

TABLE 1. Patient characteristics separated on the presence or absence of an expert measured cardiac output (n = 1075)

		Patients without CO measurement (n = 292)	**Patients with CO measurement (n = 783)**	**p-values**
Age (years)		64 ± 13	61 ± 15	0.004
Male gender		190 (65%)	484 (62%)	0.33
BMI (kg m-2)		26.9 ± 5.3	26.9 ± 5.6	0.96
Respiratory rate (bpm)		18 ± 6	18 ± 6	0.88
Mechanical ventilation		194 (66%)	438 (56%)	0.002
PEEP (cm H2O)		7 (5, 8)	7 (5, 8)	0.83
SBP (mmHg)		113 ± 25	120 ± 25	<0.001
DBP (mmHg)		59 ± 12	60 ± 12	0.44
MAP (mmHg)		76 ± 14	79 ± 14	0.014
Heart rate (bpm)		91 ± 22	87 ± 21	0.002
Atrial fibrillation		22 (8%)	56 (7%)	0.91
Norepinephrine		168 (58%)	361 (46%)	<0.001
CVP (mmHg)		9 (4 – 12)	9 (5 – 13)	0.84
Lactate (mmol L-1)		1.5 (1.0 – 2.5)	1.3 (0.9 – 2.1)	<0.001
Consciousness	Alert	75 (26%)	254 (32%)	0.018
	reacting to Voice	49 (17%)	154 (20%)	
	reacting to Pain	22 (8%)	67 (9%)	
	Unresponsive	146 (49%)	308 (39%)	
COPD		54 (18%)	88 (11%)	0.002
Acute surgery		108 (37%)	230 (29%)#	0.017
Post cardiothoracic surgery		40 (14%)	48 (6%)	<0.001
SAPS-II		49 ± 17	46 ± 17	0.004
APACHE IV score		80 ± 30	75 ± 29	0.017
90-day mortality		80 (27%)	217 (28%)	0.97

Abbreviations: APACHE; acute physiology and chronic health evaluation, BMI; body mass index, bpm; beats per minute, CO; cardiac output, CVP; central venous pressure, DBP; diastolic blood pressure, MAP; mean arterial pressure, PEEP; positive end-expiratory pressure, SAPS; simple acute physiology score, SBP; systolic blood pressure. # Significant overlap with cardiothoracic surgery

References

1. Mayo PH, Beaulieu Y, Doelken P, et.al. American College of Chest Physicians/La Société de Réanimation de Langue Française statement on competence in critical care ultrasonography. Chest 2009;135:1050–1060. https://doi.org/10.1378/chest.08-2305
2. Koster G, Van Der Horst ICC. Critical care ultrasonography in circulatory shock. Curr Opin Crit Care 2017;23:326–333. https://doi.org/10.1097/MCC.0000000000000428
3. Filipiak-Strzecka D, John B, Kasprzak J, Michalski B, Lipiec P. Pocket-size echocardiograph - a valuable tool for non-experts or just a portable device for echocardiographers? Adv Med Sci 2013;58:67–72. https://doi.org/10.2478/v10039-012-0054-2
4. AlEassa EM, Ziesmann MT, Kirkpatrick AW, Wurster CL, Gillman LM. Point of care ultrasonography use and training among trauma providers across Canada. Can J Surg 2016;59:6–8. https://doi.org/10.1503/cjs.010215
5. Funcke S, Sander M, Goepfert MS, et.al. Practice of hemodynamic monitoring and management in German, Austrian, and Swiss intensive care units: the multicenter cross-sectional ICU-CardioMan Study. Ann Intensive Care 2016;6:49. https://doi.org/10.1186/s13613-016-0148-2
6. Yang Y, Royse C, Royse A, Williams K, Canty D. Survey of the training and use of echocardiography and lung ultrasound in Australasian intensive care units. Crit Care 2016;20:339. https://doi.org/10.1186/s13054-016-1444-9
7. Panoulas VF, Daigeler A-L, Malaweera ASN, Lota AS, Baskaran D, Rahman S, Nihoyannopoulos P. Pocket-size hand-held cardiac ultrasound as an adjunct to clinical examination in the hands of medical students and junior doctors. Eur Hear J - Cardiovasc Imaging 2013;14:323–330. https://doi.org/10.1093/ehjci/jes140
8. Ho AM-H, Critchley LAH, Leung JYC, et.al. Introducing Final-Year Medical Students to Pocket-Sized Ultrasound Imaging: Teaching Transthoracic Echocardiography on a 2-Week Anesthesia Rotation. Teach Learn Med 2015;27:307–313. https://doi.org/10.1080/10401334.2015.1044657
9. Krause C, Krause R, Krause R, Gomez N, Jafry Z, Dinh VA. Effectiveness of a 1-Hour Extended Focused Assessment With Sonography in Trauma Session in the Medical Student Surgery Clerkship. J Surg Educ 2017;74:968–974. https://doi.org/10.1016/j.jsurg.2017.03.007
10. Udrea DS, Sumnicht A, Lo D, et.al. Effects of Student-Performed Point-of-Care Ultrasound on Physician Diagnosis and Management of Patients in the Emergency Department. J Emerg Med 2017;53:102–109. https://doi.org/10.1016/j.jemermed.2017.01.021
11. Bobbia X, Chabannon M, Chevallier T, et.al. Assessment of five different probes for lung ultrasound in critically ill patients: A pilot study. Am J Emerg Med 2018;36:1265–1269. https://doi.org/10.1016/j.ajem.2018.03.077
12. Rice BT, Vu H, Tran LD, Vo QX, Mowafi H. Survey of point of care ultrasound usage in emergency medicine by Vietnamese physicians. Emerg Med Australas 2015;27:580–583 . https://doi.org/10.1111/1742-6723.12476
13. Nielsen K, Lauridsen JRM, Laursen CB, Brabrand M. Physicians using ultrasound in Danish emergency departments are mostly summoned specialists. Scand J Trauma Resusc Emerg Med 2015;23:51. https://doi.org/10.1186/s13049-015-0131-1
14. Sanders J, Noble V, Raja A, Sullivan A, Camargo C. Access to and Use of Point-of-Care Ultrasound in the Emergency Department. West J Emerg Med 2015;16:747–752 . https://doi.org/10.5811/westjem.2015.7.27216

15. Neskovic AN, Skinner H, Price S, et.al. Focus cardiac ultrasound core curriculum and core syllabus of the European Association of Cardiovascular Imaging. Eur Heart J Cardiovasc Imaging 2018;19:475–481. https://doi.org/10.1093/ehjci/jey006
16. International expert statement on training standards for critical care ultrasonography. Intensive Care Med 2011;37:1077–1083. https://doi.org/10.1007/s00134-011-2246-9
17. Price S, Via G, Sloth E, Guarracino F, Breitkreutz R, Catena E, Talmor D. Echocardiography practice, training and accreditation in the intensive care: document for the World Interactive Network Focused on Critical Ultrasound (WINFOCUS). Cardiovasc Ultrasound 2008;6:49. https://doi.org/10.1186/1476-7120-6-49
18. Cecconi M, De Backer D, Antonelli M, et.al. Consensus on circulatory shock and hemodynamic monitoring. Task force of the European Society of Intensive Care Medicine. Intensive Care Med 2014;40:1795–1815. https://doi.org/10.1007/s00134-014-3525-z
19. Sakr Y, Reinhart K, Vincent J-L, et.al. Does dopamine administration in shock influence outcome? Results of the Sepsis Occurrence in Acutely Ill Patients (SOAP) Study. Crit Care Med 2006;34:589–97. https://doi.org/10.1097/01.CCM.0000201896.45809.E3
20. Vieillard-Baron A, Mayo PH, Vignon P, et.al. International consensus statement on training standards for advanced critical care echocardiography. Intensive Care Med 2014;40:654–666. https://doi.org/10.1007/s00134-014-3228-5
21. Millington SJ, Goffi A, Arntfield RT. Critical care echocardiography: a certification pathway for advanced users. Can J Anesth 2018;65:345–349 . https://doi.org/10.1007/s12630-018-1061-y
22. Dinh VA, Ko HS, Rao R, Bansal RC, Smith DD, Kim TE, Nguyen HB. Measuring cardiac index with a focused cardiac ultrasound examination in the ED. Am J Emerg Med 2012;30:1845–51. https://doi.org/10.1016/j.ajem.2012.03.025
23. Lee L, Tewalde S, Allison M, Witting M, Euerle B. Inter-Rater Reliability of Doppler Echocardiography Assessment of Cardiac Output by Emergency Physicians. Ann Emerg Med 2014;64:S126. https://doi.org/10.1016/j.annemergmed.2014.07.384
24. Andersen GN, Viset A, Mjølstad OC, Salvesen Ø, Dalen H, Haugen BO. Feasibility and accuracy of point-of-care pocket-size ultrasonography performed by medical students. BMC Med Educ 2014;14:156. https://doi.org/10.1186/1472-6920-14-156
25. Hiemstra B, Eck RJ, Koster G, et.al. Clinical examination, critical care ultrasonography and outcomes in the critically ill: cohort profile of the Simple Intensive Care Studies-I. BMJ Open 2017;7:e017170. https://doi.org/10.1136/bmjopen-2017-017170
26. Hiemstra B, Koster G, Wiersema R, et.al. The diagnostic accuracy of clinical examination for estimating cardiac index in critically ill patients: the Simple Intensive Care Studies-I. Intensive Care Med 2019;45:190–200. https://doi.org/10.1007/s00134-019-05527-y
27. Qasim A. Echocardiographer.org. http://echocardiographer.org/. Accessed 25 Oct 2019
28. Baumgartner H, Hung J, Bermejo J, et.al. Recommendations on the echocardiographic assessment of aortic valve stenosis: a focused update from the European Association of Cardiovascular Imaging and the American Society of Echocardiography. Eur Heart J Cardiovasc Imaging 2017;18:254–275 . https://doi.org/10.1093/ehjci/jew335
29. Quiñones MA, Otto CM, Stoddard M, Waggoner A, Zoghbi WA. Recommendations for quantification of Doppler echocardiography: a report from the Doppler Quantification Task Force of the Nomenclature and Standards Committee of the American Society of Echocardiography. J Am Soc Echocardiogr 2002;15:167–84

30. Royse CF, Seah JL, Donelan L, Royse AG. Point of care ultrasound for basic haemodynamic assessment: novice compared with an expert operator. Anaesthesia 2006;61:849–855. https://doi.org/10.1111/j.1365-2044.2006.04746.x

31. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999;8:135–60. https://doi.org/10.1177/096228029900800204

32. Critchley LAH, Critchley JAJH. A meta-analysis of studies using bias and precision statistics to compare cardiac output measurement techniques. J Clin Monit Comput 1999;15:85–91. https://doi.org/10.1023/A:1009982611386

33. Orde S, Slama M, Hilton A, Yastrebov K, McLean A. Pearls and pitfalls in comprehensive critical care echocardiography. Crit Care 2017;21:279. https://doi.org/10.1186/s13054-017-1866-z

34. Betcher J, Majkrzak A, Cranford J, Kessler R, Theyyunni N, Huang R. Feasibility study of advanced focused cardiac measurements within the emergency department. Crit Ultrasound J 2018;10:10. https://doi.org/10.1186/s13089-018-0093-4

35. Villavicencio C, Leache J, Marin J, Oliva I, Rodriguez A, Bodí M, Soni NJ. Basic critical care echocardiography training of intensivists allows reproducible and reliable measurements of cardiac output. Ultrasound J 2019;11:5. https://doi.org/10.1186/s13089-019-0120-0

36. Lebeau R, Sas G, Rayes M El, Serban A. Left ventricular ejection fraction assessment by non-cardiologists from transverse views using a simplified wall motion score index. Echo Res Pract 2015;2(1):1-8 https://doi.org/10.1530/ERP-14-0003

37. Heiberg J, Hansen LS, Wemmelund K, et.al. Point-of-Care Clinical Ultrasound for Medical Students. Ultrasound Int open 2015;1:E58-66. https://doi.org/10.1055/s-0035-1565173

38. Kobal SL, Lior Y, Ben-Sasson A, Liel-Cohen N, Galante O, Fuchs L. The feasibility and efficacy of implementing a focused cardiac ultrasound course into a medical school curriculum. BMC Med Educ 2017;17:94. https://doi.org/10.1186/s12909-017-0928-x

39. Wong I, Jayatilleke T, Kendall R, Atkinson P. Feasibility of a focused ultrasound training programme for medical undergraduate students. Clin Teach 2011;8:3–7. https://doi.org/10.1111/j.1743-498X.2010.00416.x

40. Parks AR, Verheul G, Leblanc-duchin D, Atkinson P. Effect of a point-of-care ultrasound protocol on the diagnostic performance of medical learners during simulated cardiorespiratory scenarios. CJEM 2019;17(3):263–269. https://doi.org/10.1017/cem.2014.41

41. Arora S, Cheung AC, Tarique U, Agarwal A, Firdouse M, Ailon J. First-year medical students use of ultrasound or physical examination to diagnose hepatomegaly and ascites: a randomized controlled trial. J Ultrasound 2017;20:199–204. https://doi.org/10.1007/s40477-017-0261-6

42. Ang J, Doyle B, Allen P, Cheek C. Teaching bedside ultrasound to medical students. Clin Teach 2018;15:331–335. https://doi.org/10.1111/tct.12692

7

Should the probe replace your stethoscope? A SICS-I sub-study comparing auscultation and lung ultrasound in the critically ill

Cox EGM, Koster G, MD, Baron A, Kaufmann T, Eck RJ, Veenstra TC, Hiemstra B, Wong A, Kwee TC, Tulleken JE, Keus F, Wiersema R, van der Horst ICC; SICS Study Group.
Crit Care. 2020 Jan 13;24(1):14. doi: 10.1186/s13054-019-2719-8.

Abstract

Background

In critically ill patients, auscultation might be challenging as dorsal lung fields are difficult to reach in supine positioned patients and the environment is often noisy. In recent years clinicians have started to consider lung ultrasound as a useful diagnostic tool for a variety of pulmonary pathologies, including pulmonary edema. The aim of this study is to compare lung ultrasound and pulmonary auscultation for detecting pulmonary edema in critically ill patients.

Methods

This study was a planned sub-study of the Simple Intensive Care Studies-I, a single-center, prospective observational study. All acutely admitted patients who were 18 years and older with an expected ICU stay of at least 24 hours were eligible for inclusion. All patients underwent clinical examination combined with lung ultrasound, conducted by researchers not involved in patient care. Clinical examination included auscultation of bilateral regions for crepitations and rhonchi. Lung ultrasound was conducted according to the Bedside Lung Ultrasound in Emergency-protocol. Pulmonary edema was defined as three or more B-lines in at least two (bilateral) scan-sites. Agreement was described by using the Cohen κ-coefficient, sensitivity, specificity, negative predictive value, positive predictive value, and overall accuracy. Sensitivity analyses were performed in patients who were not mechanically ventilated.

Results

The Simple Intensive Care Studies-I cohort included 1075 patients, of whom 926 (86%) were eligible for inclusion in this analysis. 307 of the 926 patients (33%) fulfilled the criteria for pulmonary edema on lung ultrasound. In 156 (51%) of these patients, auscultation was normal. A total of 302 patients (32%) had audible crepitations or rhonchi upon auscultation. From 130 patients with crepitations, 86 patients (66%) had pulmonary edema on lung ultrasound; and from 209 patients with rhonchi, 96 patients (46%) had pulmonary edema on lung ultrasound. Agreement between auscultation findings and lung ultrasound diagnosis was poor (κ - statistic 0.25). Sensitivity analyses showed that diagnostic accuracy of auscultation was better in non-ventilated than in ventilated patients.

Conclusions

The agreement between lung ultrasound and auscultation is poor.

Introduction

Physicians are trained to use auscultation as part of clinical examination in routine care for critically ill patients. Auscultation is accepted as one of the essential components of the clinical examination. Frequent pathologies encountered in the critically ill are pulmonary edema and pneumonia; both present with an increase in alveolar fluid and often coexist. Crepitations and rhonchi can be present in patients with pulmonary edema [1]. In recent years clinicians have started to consider lung ultrasound (LUS) as a useful diagnostic tool for a variety of pulmonary pathologies [2–4]. An increasing body of evidence supports the use of LUS in diagnosing pulmonary edema and/or pneumonia [5]. Several studies have shown the diagnostic value of LUS in patients with dyspnea or specific diagnoses, such as pneumothorax, high altitude pulmonary edema, and cardiogenic pulmonary edema [6–10]. LUS has even been suggested to be superior to chest radiography (X-ray), and comparable to chest computed tomography (CT)-scan for the diagnosis of pulmonary edema and increased alveolar fluid (commonly referred to as interstitial syndrome) [3,8]. However, few studies have compared LUS to pulmonary auscultation, even while the stethoscope still constitutes the majority of contemporary practice [11–13].

In critically ill patients, auscultation might be challenging as dorsal lung fields are difficult to reach in supine positioned patients and the environment is often noisy. No studies have prospectively compared auscultation with LUS in the intensive care unit (ICU) setting. Accordingly, the aim was to compare the agreement of LUS against pulmonary auscultation for the detection of pulmonary edema in acutely admitted ICU patients. We hypothesized that auscultation for pulmonary edema would have insufficient agreement compared to LUS.

Methods

Design and setting

This was a planned sub-study of the Simple Intensive Care Studies-I (SICS-I), a single-center, prospective observational study designed to evaluate the diagnostic and prognostic value of combinations of clinical examination and critical care ultrasonography (CCUS), in critically ill patients [14]. This sub-study and prespecified hypothesis was added to the SICS-I study [14]. The local institutional review board (Medisch Ethische Toetsingscommissie of the University Medical Center Groningen; UMCG) approved the study (M15.168207). This manuscript was reported according to the Standards for Reporting of Diagnostic Accuracy Studies guidelines [15].

Participants

All acutely admitted patients who were 18 years and older with an expected ICU stay of at least 24 hours were eligible for inclusion. Patients were excluded if their ICU admission was planned, if acquiring research data interfered with clinical care due to e.g., continuous resuscitation efforts (e.g. mechanical circulatory support), or if consent was not obtained. In this sub-study we selected a convenience sample of patients who had bilateral LUS images in at least two scan sites.

Variables

All included patients underwent clinical examination followed by CCUS within the first 24 hours of their ICU admission. The researchers were senior medical students and junior residents trained by cardiologist-intensivists for both clinical examination and CCUS before contributing to the study. Training included self-study of theory on how to perform auscultation and lung ultrasound, at least two hours hands-on training from cardiologists-intensivists, practice on healthy individuals during practical sessions, and supervised clinical examination and CCUS in their first twenty patients.

Data from clinical examination was prospectively collected based on definitions in the protocol, including the presence of crepitations and rhonchi [14]. Abnormal auscultation was defined as the presence of crepitations and/or rhonchi at any of the sites. Pulmonary edema was defined as the presence of three or more B-lines; diffuse pulmonary edema was defined as edema in two or more scan sites of LUS bilaterally [16].

Auscultation was performed of the anterior and axillary lung fields in each hemithorax with the patient in a supine position. Subsequently, CCUS was performed following a predefined protocol using a phased array probe (M3S or M4S) set at a frequency of 3.6 MHz, a depth of 15 cm, and maximal image width [17] (Vivid-S6, GE Healthcare, London, UK). LUS was performed using the Bedside Lung Ultrasound in Emergency (BLUE)-protocol, assessing six scan sites per patient (superior, inferior and lateral, bilaterally) (Figure 1). In each scan site, the numbers of B-lines (0-5) were recorded [18]. Measurements were subsequently conducted by researchers, who were not involved in patient care. Researchers were instructed not to share their findings with the attending physicians, so that these were used for research purposes only.

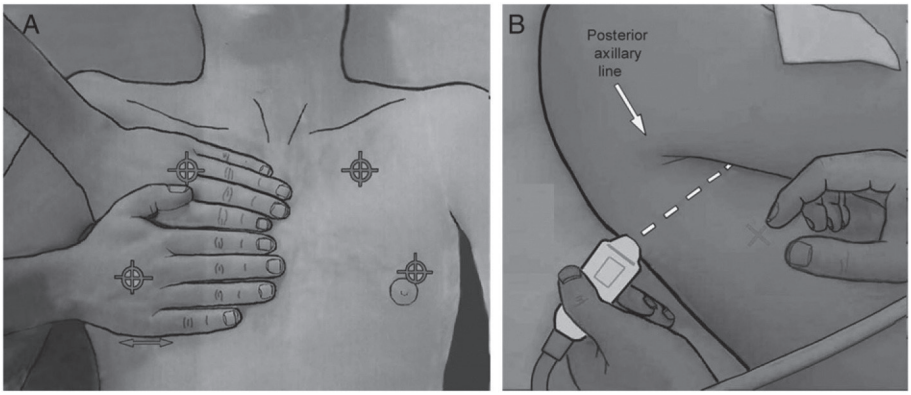


FIGURE 1. The six scan sites according to the BLUE-protocol [18].

Statistical analyses

The overall statistical methods were described in the predefined statistical analysis plan (SAP) of the main study (NCT02912624). Continuous variables were reported as means with standard deviation (SD) or median with interquartile range (IQR) depending on distributions. Categorical data were presented in proportions. Student's T-test, Mann-Whitney U test or the Chi-square tests were used as appropriate. The agreement between LUS and auscultation for pulmonary edema was described by using the Cohen k-coefficient. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV) and diagnostic accuracy of lung ultrasound against auscultation to detect pulmonary edema were calculated. Analyses were performed on complete cases using Stata version 15 (StataCorp, CollegeStation, TX, USA). A subgroup analysis was performed to assess whether these results were robust in patients who were not mechanically ventilated. We performed a sensitivity analysis to assess the agreement and diagnostic accuracy of LUS for pulmonary edema on chest X-ray, in patients where a chest X-ray was available shortly before or after study inclusion (i.e., on the same day).

The SICS-I was designed to address multiple hypotheses on six different outcomes and, therefore, the pulmonary edema outcome was adjusted for multiple hypothesis testing. We refer to our SAP for more details, but in short a p-value of 0.015 indicated statistical significance and p-values between 0.015 and 0.05 indicated suggestive significance with an increased family-wise error rate [19]. For secondary or sensitivity analyses, a p-value below 0.05 indicated statistical significance due to the hypothesis-generating purpose. Accordingly, the primary analyses are presented with 98.5% CIs and secondary (subgroup) analyses with 95% CIs.

Results

This SICS-I sub-study started on September 15th, 2015 and continued until July 22nd, 2017, during which 1009 patients were included. A total of 149 patients (15%) were excluded because no bilateral or less than two scan sites were scanned due to emphysema, drains, or wound dressings hampering the ultrasound windows; leaving 926 patients (85%) for analysis (Figure 2). Baseline characteristics of all patients are shown in Table 1.

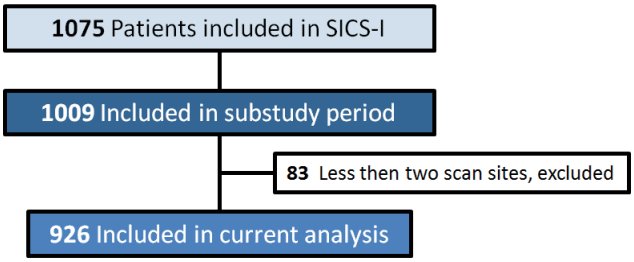


FIGURE 2. Flowchart *

*Less than two scan sites meaning if less than two out of six scan sites or no bilateral scan sites of LUS were available, the presence of pulmonary edema could not be assessed.

TABLE 1. Baseline characteristics of all included patients

	N = 926
Age, years (SD)	62 (14)
Gender, male (%)	598 (64)
Height, cm (SD)	176 (10)
Weight, kg (SD)	83 (18)
Mechanical ventilation, n (%)	537 (57)
Vasoactive medication, n (%)	461 (49)
APACHE IV - score, mean (SD)	76 (29)
Admission type - Surgical, n (%) - Medical, n (%)	 292 (31) 645 (69)
Outcomes - Length of stay, days - 90-day mortality, n (%)	 3.3 (1.9-6.8) 249 (27)

Findings of lung ultrasound and auscultation

The criteria for pulmonary edema diagnosed by LUS were met in 307 of 926 patients (33%). In 156 of these patients (51%), auscultation was normal. A total of 302 of 926 patients (32%) had pulmonary edema diagnosed by pulmonary auscultation. From these patients, 151 patients (50%) had pulmonary edema on LUS. Of the 302 patients with pulmonary edema on auscultation, 130 patients had crepitations and 209 patients had rhonchi.

From 130 patients with crepitations, 86 patients (66%) had pulmonary edema on LUS; and of the 209 patients with rhonchi, 96 patients (46%) had pulmonary edema on LUS. The agreement between auscultation and LUS was poor (κ - statistic 0.25).

Diagnostic performance

Diagnostic performance measures of crepitations, rhonchi and auscultation for detection of pulmonary edema are displayed in Table 2. The sensitivity of crepitations was 66% (98.5% CI 55-76), specificity was 71% (98.5% CI 67-75), positive predictive value was 28% (98.5% CI 22-34) and negative predictive value was 93% (98.5% CI 90-95). Overall diagnostic accuracy of crepitations was 72% (98.5% CI 69-74). The sensitivity of rhonchi was 47% (98.5% CI 39-56), specificity was 69% (98.5% CI 65-74), positive predictive value was 31% (98.5% CI 25-38) and the negative predictive value was 82% (98.5% CI 77-85). Overall diagnostic accuracy of rhonchi was 64% (98.5% CI 61-67).

The sensitivity of abnormal auscultation overall was 52% (98.5% CI 45-59), specificity was 74% (98.5% CI 70-79), positive predictive value was 49% (98.5% CI 42-56) and the negative predictive value was 76% (98.5% CI 72-80). Overall diagnostic accuracy of auscultation was 67% (98.5% CI 64-70).

TABLE 2. Test characteristics of specific findings compared to LUS in all patients

	Abnormal	Total	Diagnostic performance in% (98.5% confidence intervals)				
	N	N	Sensitivity	Specificity	PPV	NPV	Diagnostic accuracy
Crepitations	130	917	66 (55-76)	71 (67-75)	28 (22-34)	93 (90-95)	72 (69-74)
Rhonchi	209	913	47 (39-56)	69 (65-74)	31 (25-38)	82 (77-85)	64 (61-67)
Auscultation	302	926	52 (45-59)	74 (70-79)	49 (42-56)	76 (72-80)	67 (64-70)

Abnormal auscultation was defined as the presence of crepitations and/or rhonchi at any of the sites.

Sensitivity analysis

Diagnostic accuracy of auscultation improved if patients were not mechanically ventilated (Table 3). Overall accuracy for auscultation was 69% (95% CI 64-74) in non-mechanically ventilated patients and 67% (98.5% CI 64-70) in all patients ($p<0.001$). Overall accuracy for crepitations was 71% (95% CI 67-76) and for rhonchi 66% (95% CI 61-71) in non-ventilated patients. The agreement between auscultation and LUS improved in non-mechanically ventilated patients (κ - statistic 0.31).

TABLE 3. Test characteristics of specific findings compared to LUS in non-mechanically ventilated patients

	Abnormal	Total	Diagnostic performance in% (95% CI confidence intervals)				
	N	N	Sensitivity	Specificity	PPV	NPV	Diagnostic accuracy
Crepitations	73	387	36 (28-45)	90 (85-94)	66 (55-75)	73 (70-75)	71 (67-76)
Rhonchi	70	384	28 (21-36)	87 (82-91)	54 (44-64)	69 (66-71)	66 (61-71)
Auscultation	124	391	51 (43-60)	79 (73-84)	56 (49-63)	75 (72-79)	69 (64-74)

Radiologists' reports assessing the chest X-ray were analyzed in a subset of 315 patients. Baseline characteristics of these patients were comparable to the overall population (E-Table 1). Median time lag between LUS and chest X-ray was 4 hours (2-7 hours). In 89 of these patients (28%) the radiologist reported the diagnosis of edema, in 6 patients (2%) it was unclear, and in 220 patients (70%) there was no pulmonary edema on chest X-ray according to the radiologist (E-Table 2). The agreement and diagnostic accuracy of LUS for pulmonary edema as diagnosed on chest X-ray were limited (κ - statistic 0.12; E-Table 3).

Discussion

In this prospective observational study, we found poor agreement between auscultation and LUS for the diagnosis of pulmonary edema in acutely admitted critically ill patients.

Several previous studies focused on the diagnostic accuracy of LUS compared to other imaging modalities, such as chest X-ray and CT-scan [4,10,20]. However, few studies have compared the diagnostic accuracy of LUS with the stethoscope, one of the most frequently used instruments at the bedside. Lichtenstein et al. prospectively compared the diagnostic performance of auscultation, LUS, and chest X-ray, for detecting alveolar consolidation and alveolar–pulmonary edema with CT-scan in 32 patients with acute respiratory distress syndrome and in 10 healthy volunteers [13]. The authors found that auscultation had a diagnostic

accuracy of 55% for alveolar–pulmonary edema, which corresponds fairly to the 64% accuracy in our study [13]. In that study LUS had a diagnostic accuracy of 97% for alveolar consolidation and 95% for alveolar–pulmonary edemas and chest X-ray had a diagnostic accuracy of 75% for alveolar consolidation and 72% for alveolar–pulmonary edema [13]. In a sensitivity analysis we observed that the agreement and diagnostic accuracy of LUS for pulmonary edema was limited when compared to chest X-ray, which is in line with other studies [1].

Another study by Torino et al. prospectively investigated the agreement between auscultation and LUS in non-admitted patients before and after undergoing hemodialysis [11]. The authors similarly found a very poor agreement (κ - statistic 0.16, in this study κ - statistic 0.25) between the presence of crepitations on auscultation and the presence of B-lines on LUS in a total of 1106 measurements in 79 patients [11]. Although their population seems different to ours, patients receiving dialysis may also suffer from pulmonary edema as a consequence of fluid overload. Their results and conclusions are similar to ours and therefore, these observations may be generalizable to populations beyond the critically ill.

We found that the diagnostic accuracy of auscultation improved if patients were not mechanically ventilated, no previous study has reported this finding. Acoustic disturbances caused by the ventilators might explain the complicated appreciation of subtle auscultation findings.

Implications and generalizability

Improved diagnostic accuracy for detecting pulmonary edema could lead to improved treatment leading to increased benefits and decreased harms for the patient. In critically ill patients typically multiple pathophysiological processes are co-occurring at the same time, which hampers the extrapolation of tests characteristics for diagnosing abnormalities in these patients, such as pulmonary edema. As some physicians still use auscultation to detect pulmonary edema, we think our study clarifies that auscultation may not be as reliable for detecting pulmonary edema as classically perceived, especially in the ICU. Ultrasonography becomes increasingly available and our data add nuance to the discussion surrounding how this technology might be properly integrated into clinical practice in the care of the critically ill. These observations encourage further research of LUS, the need for external validation remains to increase the generalizability of this diagnostic modality.

Limitations

Several limitations of this study must be acknowledged. First, the clinical examination and ultrasonography were conducted as early as possible after ICU

admission which limits applicability of use in patients with prolonged admission. Further studies should explicate how auscultation and LUS compare in other departments and more specifically, other pathologies such as a pneumothorax. Second, we were not able to validate all our LUS assessments by experts, also because there are no reference standards for the interpretation of LUS. Chest radiography is another diagnostic method that is frequently used for the assessment of pulmonary edema. However, previous studies have suggested that LUS is superior to chest X-ray for diagnosing pulmonary edema [3,8], which made us decide not to use this modality as a reference standard. We limited LUS reporting to the number of B-lines per field and did not use further qualitative commentary. Third, during clinical examination researchers performed both the auscultation of the lung fields and did the LUS scan sites, however, they only specified whether they heard significant crepitation or rhonchi. Other abnormal breathing sounds were not recorded, and only documented the overall presence or absence of abnormal breath sounds, we are unable to compare auscultation with LUS for each specific scan sites. In addition, ideally, we ask the patient to cough to distinguish between rhonchi and/or crepitations. Unfortunately, the large majority of the patients in the ICU are not cooperative to this request. Fourth, even though the researchers who performed the measurements were not involved in patient care, they were not blinded for patient information, such as admission diagnoses and other clinical variables and the results of auscultation when performing the CCUS. However, as ultrasonography was always performed after auscultation, we believe it is proper to discuss this potential source of bias but do not believe that it substantially influenced our results due to the objective nature of B-line appearance. Fifth, since researchers were senior medical students and junior residents, auscultation by more experienced medical doctors could potentially improve diagnostic accuracy. Last, 83 (8%) patients were excluded from analyses due to the absence of LUS- or auscultation data. However, the relatively small proportion of this excluded patient group makes it unlikely that excluded patients would have altered conclusions. Despite potential biases and limitations, we showed that the agreement between auscultation and lung ultrasound was poor. This is important as currently data is scarce on the diagnostic value of new non-invasive bed tools such as CCUS, especially in comparison with clinical examination in critically ill patients.

Conclusions

The agreement between auscultation and LUS for detecting pulmonary edema is poor. Further studies are necessary to explicate whether the use of LUS should be preferred over the use of auscultation for the detection of pulmonary edema in critically ill patients.

References

1. Assaad S, Kratzert WB, Shelley B, Friedman MB, Perrino A. Assessment of Pulmonary Edema: Principles and Practice. J Cardiothorac Vasc Anesth 2018;32(2):901-914
2. Mojoli F, Bouhemad B, Mongodi S, Lichtenstein D. Lung Ultrasound for Critically Ill Patients. Am J Respir Crit Care Med 2019;199(6):701-714
3. Shrestha GS, Weeratunga D, Baker K. Point-of-Care Lung Ultrasound in Critically ill Patients. Rev Recent Clin Trials 2018;13(1):15-26
4. Xirouchaki N, Magkanas E, Vaporidi K, et al. Lung ultrasound in critically ill patients: comparison with bedside chest radiography. Intensive Care Med 2011;37(9):1488-1493
5. Miglioranza MH, Picano E, Badano LP, et al. Pulmonary congestion evaluated by lung ultrasound predicts decompensation in heart failure outpatients. Int J Cardiol 2017;240:271–8
6. Saigal S, Joshi R, Sharma JP, Pandey V, Pakhare A. Lung Ultrasound and Blood Gas-Based Classification of Critically Ill Patients with Dyspnea: A Pathophysiologic Approach. Indian J Crit Care Med 2018;22(11):789
7. Amir R, Knio ZO, Mahmood F, et al. Ultrasound as a Screening Tool for Central Venous Catheter Positioning and Exclusion of Pneumothorax. Crit Care Med 2017;45(7):1192-1198
8. Yang W, Wang Y, Qiu Z, et al. Lung Ultrasound Is Accurate for the Diagnosis of High-Altitude Pulmonary Edema: A Prospective Study. Can Respir J 2018
9. Cortellaro F, Ceriani E, Spinelli M, et al. Lung ultrasound for monitoring cardiogenic pulmonary edema. Intern Emerg Med 2017;12(7):1011-1017
10. Martindale JL, Wakai A, Collins SP, et al. Diagnosing Acute Heart Failure in the Emergency Department: A Systematic Review and Meta-analysis. Acad Emerg Med 2016;23(3):223-242
11. Torino C, Gargani L, Sicari R, et al. The Agreement between Auscultation and Lung Ultrasound in Hemodialysis Patients: The LUST Study. Clin J Am Soc Nephrol 2016;11(11):2005-2011
12. Tasci O, Hatipoglu ON, Cagli B, Ermis V. Sonography of the chest using linear-array versus sector transducers: Correlation with auscultation, chest radiography, and computed tomography. J Clin Ultrasound 2016;44(6):383-389
13. Lichtenstein D, Goldstein I, Mourgeon E, Cluzel P, Grenier P, Rouby J-J. Comparative diagnostic performances of auscultation, chest radiography, and lung ultrasonography in acute respiratory distress syndrome. Anesthesiology 2004;100(1):9-15
14. Hiemstra B, Eck RJ, Koster G, et al. Clinical examination, critical care ultrasonography and outcomes in the critically ill: cohort profile of the Simple Intensive Care Studies-I. BMJ Open 2017;7(9):e017170
15. Cohen JF, Korevaar DA, Altman DG, et al. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open 2016;6(11):e012799
16. Marini F, Frigieri FC, Guarducci D. Lung and Pleural Ultrasonography in Emergency and Intensive Care. Textb Echocardiogr Intensivists Emerg Physicians. Cham: Springer International Publishing; 2019
17. Wiersema R, Castela Forte JN, Kaufmann T, et al. Observational Study Protocol for Repeated Clinical Examination and Critical Care Ultrasonography Within the Simple Intensive Care Studies. J Vis Exp 2019;143:e58802

18. Lichtenstein DA. BLUE-Protocol and FALLS-Protocol: Two Applications of Lung Ultrasound in the Critically Ill. Chest 2015;147(6):1659-1670
19. Jakobsen JC, Wetterslev J, Winkel P, Lange T, Gluud C. Thresholds for statistical and clinical significance in systematic reviews with meta-analytic methods. BMC Med Res Methodol 2014;14(1):1-13
20. Martelius L, Heldt H, Lauerma K. B-Lines on Pediatric Lung Sonography. J Ultrasound Med 2016;35(1):153-157

8

Discussion and Future Perspectives

PART 1: Hemodynamic interventions

Medicine has come a long way from its early trial and error to its modern large multicentre randomised clinical trials (RCTs). Like the early days, it is still of utmost importance for each caregiver to know the potential benefits of an intervention, and equally important, to know the potential harms of that intervention.

This thesis was unable to reach conclusions supporting or refuting the use of any of the inotropes. Still, many other papers, including reviews, guidelines and society statements, recommend using inotropes in critically ill patients. How can we explain the differences between our systematic reviews (SRs) and these statements, and which concepts are important to consider?

Risk of bias

Previous SRs and meta-analyses concluded that certain inotropes were beneficial compared to placebo or other inotropes. The conclusions of the SRs and meta-analyses included in this thesis contrasted substantially with previous SRs. Some methodological aspects of SR conduct explain the differences between our and previous findings. Most previous SRs did not assess the risks of bias and/or did not incorporate the bias risk assessment in their results and adapted the conclusions accordingly. Risk of bias assessment is done by evaluating the domains in which bias is associated with the trial design and - conduct. In the first version of the risk of bias assessment tool of Cochrane (ROB 1), seven domains were considered/ identified, i.e., sequence generation, allocation sequence concealment, blinding of participants and personnel, blinding of outcome assessment, incomplete outcome data, selective outcome reporting and other potential threats to validity (i.e. academic or funding bias). Each bias domain can be assessed as high risk of bias, uncertain or low risk of bias; uncertain bias assessments are categorised under high risk of bias. Pooling the individual RCT results without incorporating the risk of bias evaluations is likely to overestimate beneficial effects and underestimate the harmful effects of the evaluated interventions.[1,2] The overestimation is driven by the RCTs with a high risk of bias which more often shows a beneficial effect.[1]

The previously suggested beneficial intervention effects of some inotropes are likely explained by the high risk of bias of many of the included RCTs. Our SRs identified only one RCT with an overall low risk of bias (levosimendan n=1 out of 49 RCTs; milrinone n=0 out of 16 RCTs; dopamine n=0 out of 17 RCTs), implicating that all randomised RCTs except one were at high risks of bias. Of the total 574 bias domains (seven bias domains in 82 RCTs), which we evaluated, only 32% of the domains were assessed as being low risk. One incentive to conduct an SR is to learn what is already known on a specific intervention and learn from others' mistakes on the same subject in the past. Every clinician needs to understand

that the quality of an SR[3], and thus the validity of its conclusions, is trivial and that it depends heavily on the choices of the author(s) and their adherence to the protocols and methodological standards, such as the Cochrane Handbook for Systematic Reviews. Critical appraisal of the scientific literature, including assessing the risk of bias associated with the conduct of SRs and meta-analyses, is complex, a skill that every doctor(-to-be) ought to possess. Future RCT's need to adhere to CONSORT/ICH guidelines and should comply with all bias reducing requirements, if possible. We owe it to our patients that we continuously keep on improving, and that starts by taking advantage of the opportunity to learn from previous mistakes; there is much to be learned from others. We are aware of the efforts and difficulties of conducting an RCT in critically ill patients. Certain biases, such as lack of blinding, are challenging to prevent, especially in an environment where many caregivers continuously monitor the patient. Fortunately, the lack of blinding does not affect mortality rates in RCTs of ICU interventions while other biases do.[4]

Risk of random error

Previous SRs on inotropes did not account for the risks of random error. The random error refers to imprecision or 'the play of chance', meaning that multiple replications of the same study will produce different effect estimates simply because of sampling variation even if they would, on average, give the true answer. So, some meta-analyses with 'positive' findings may be due to the play of chance rather than a 'true' intervention effect, and vice versa, in case of 'negative' or 'neutral' findings.[5,6] This can especially be the case in a meta-analysis with a small number of RCTs and a small number of randomised patients risking spurious findings and premature conclusions on the beneficial effects of an intervention. So, the amount of available evidence influences the precision. One way of controlling for sparse data, repeated testing and the 'play of chance' in meta-analyses is by conducting trial sequential analysis (TSA); a statistical method comparable with interim analyses in a single RCT.[7] TSA controls for the risk of random error by widening the thresholds (i.e. conventional boundaries) for statistical significance by calculation of a diversity-adjusted required information size estimate (DARIS), trial sequential monitoring boundaries and an adjusted confidence interval around the intervention effect estimate. The DARIS is comparable to a sample size calculation for an RCT and visualises the number of patients required for conclusions with a low risk of random error. The TSA adjusted confidence intervals represent the precision around the effect estimate, controlled for the information size. The above is calculated based on a control event rate, a realistic relative risk reduction (RRR), an overall percentage for the risk of a type I error (alpha) and a power (1-beta).

Conventional meta-analyses are at risk for random errors with typically false claims of effectiveness of a specific intervention, if not considered in the perspective of the amount of data necessary for strong conclusions.[3] Of all the TSA analyses conducted with realistic RRR, in our SRs with meta-analyses evaluating inotropes, *none* of the TSAs had sufficient patients included meeting the DARIS estimates.

Due to sparse data, TSA could either often not be conducted, or TSA showed that many more patients needed to be randomised before a definitive answer could be given. In our SRs, TSA mostly suggested that there was no significant intervention effect, which was typically in contrast with the reported significant p-values in conventional analyses (levosimendan in the non-cardiac surgery setting) or TSA suggested that further studies would not change the conclusions on outcomes (dopamine post-hoc analysis; futility). So, future RCTs and SRs with meta-analyses should be conducted according to the highest standard; that is, they should acknowledge and consider the risks of random error beyond standard conventional p-values. Any meta-analyses' information size should at least be as large as the sample size of a well-powered sufficiently large single (low risk of bias) RCT with a *realistic* RRR.

Clinical heterogeneity

All SRs in this thesis had multiple sources of clinical heterogeneity. This resulted in not pooling the data in one review (levosimendan) and in too small sample sizes in subgroups, prohibiting the exploration of differential treatment effects according to predefined subgroups.

Clinical heterogeneity – the patients

The major problem with research of ICU patients is the clinical differences, including all the confounders in the diseases and pathophysiological processes of the patients, and specifically in this thesis, the nature of their cardiac dysfunction. First, patients may differ in age and comorbidities. Second, cardiac failure may, for instance, result from an acute myocardial infarction, a previous myocardial infarction, sepsis or post-operative cardiac stunning. Third, the severity of cardiac dysfunction may differ: patients with a left ventricular ejection fraction (LVEF) <20% differ from patients with an LVEF <40%. Fourth, the method of measuring LVEF may differ. If outcomes between trials vary, this may originate from clinical heterogeneity of patients or differences in inotrope interventions, or differences in outcome measurements or random error. It may also be that inotropes are indeed beneficial in a specific homogenous patient group and not in all others, which could explain the frequently observed average intervention effect close to zero in the meta-analyses. Such differential effects may be explored in predefined subgroup analyses provided that RCTs are sufficiently large.

Another aspect to consider is the difference in intervention effects between centres in multicentre RCTs, i.e. smaller versus larger and general hospitals versus highly specialised academic centres. Centres will select different patients depending on their capacities and ambitions. Pooling all these patients into one meta-analytic overall effect estimate could easily dilute potential differences in treatment effects. This should be addressed a priori and accounted for in the trial design (stratification or selection) or the meta-analysis design (predefined subgroup analyses).

Clinical heterogeneity – the interventions

RCTs apply highly varying protocols regarding algorithms, doses and durations of inotrope interventions. This results in severe clinical heterogeneity when conducting an SR since the initiation, dosing, duration and halting of the intervention differs between the RCTs. Although one could explore such discrepancies in predefined subgroups, for instance, as we attempted in the dopamine review, not all variants can be explored in sufficiently large subgroups, and all subgroup analyses are by definition only with explorative intentions. The fundamental problem with all inotropes is that the ideal dosing regimen remains unknown based on these inherent limitations.

The dosing regimen can be fixed or adapted according to certain variables. Variables can be general, such as age, weight or body-mass index, or specific hemodynamic monitoring variables. The latter variables include blood pressure or cardiac output, or cardiac index. Other variables can be a normalisation of the diuresis, a decrease in the difference between peripheral to central temperature, a decrease in mottling or capillary refill time or an increase in the central or mixed venous oxygen saturation or an increase in lactate clearance.

Further, all these variables could either serve as a trigger to initiate or halt treatment or as a target to guide inotrope dosing. In addition, each variable can be used in isolation or combination with a few or multiple other variables, or as a complex treatment algorithm, or it may be left to the discretion of the treating physician. Unfortunately, none of the reviewed RCTs did assess all these variables. The trials frequently even did not measure the cardiac output continuously; when interpreting results of RCTs and meta-analyses the registration of these trigger or target variables (and the accuracy and precision of their measurements) is frequently omitted.

Clinical heterogeneity – the comparators and co-interventions

Ideally, trialists have used placebo as a comparator to evaluate the true potential benefit of the inotrope intervention. In addition, the use of any co-interventions

should be according to a protocol. Both the comparator and differential use of co-interventions may affect the estimate of the evaluated intervention effect. A placebo was used as a comparator in 25 of the 82 RCTs (dopamine n=1 out of 17 RCTs; milrinone n=5 out of 16 RCTs; levosimendan n=19 out of 49 RCTs), although co-interventions (especially other inotropes) were mostly allowed. Typically, these co-interventions were not predefined and described with details, making many of them add-on trials. Especially in inotropes, this is tricky since another inotrope may act as an effect modifier of the investigated inotrope.

Clinical heterogeneity – the outcomes

The outcomes in the SRs were chosen based on the patient's perspective, according to GRADE.[8] Unfortunately, many RCTs on inotropes investigated surrogate outcomes without registering any patient-important outcome data, and therefore, these RCTs did not contribute to the meta-analyses. For inotropes, surrogate outcomes may include cardiac output data or any other variables related to organ dysfunction interpreted as a pathophysiologic sign of insufficient organ perfusion. However, every intervention by definition comes with a risk, especially medications with all their side effects. Inotropes are associated with myocardial ischemia and/or arrhythmias. Therefore, RCTs which only evaluate surrogate data do not add to the evidence base. So, future RCTs and SRs should focus on patient-relevant outcomes. Hopefully, the Core Outcome Measures of Effectiveness Trials (COMET) initiative will develop a set of standardised outcomes for RCTs with critically ill patients having cardiac dysfunction.[9]

The precise definitions of outcomes typically differ between the RCTs, which will result in either under- or overreporting of (adverse) events. Typically, event rates of myocardial infarction, arrhythmias and hypotension heavily depend on their precise definitions.

In conclusion, the risks of bias, risks of random error, and clinical heterogeneity challenge the conduct of individual RCTs and the design and interpretation of an SR with or without meta-analysis and its related questions on the generalisability of the data. Not even acknowledging all these important validity issues reduces the level of the evidence. More critical: ignoring the heterogeneity issues in SRs and blind acceptance of review conclusions without a critical appraisal may harm our patients. The numbers of patients included in our reviews were simply too small to explore sources of clinical heterogeneity. This lack of power will continue to be an issue, especially in the tailor-made 'personalised medicine' machine learning era.[10] This is what is called the precision medicine paradox: numbers need to be sufficiently large before interventions can be accepted with a sufficiently high level of evidence; at the same time, the interventions need to be tailored to the individual patient with his/her specific characteristics in order to be considered personalised medicine.

The SRs and meta-analyses on inotropes in this thesis show that the evidence base for inotropes is lacking. Proper evaluation of inotropes is the responsibility of research, but proper critical appraisal of the evidence and the adoption or rejection of inotrope interventions in clinical practice is the responsibility of each intensivist. In due time, inotropes will likely be adopted in practice, and care must be taken to do no further harm.

PART 2: Hemodynamic monitoring – ultrasonography

The use of ultrasonography in Intensive Care Medicine has become standard. Alike therapeutic interventions, applying diagnostic tools in the ICU, such as ultrasonography, should have a proper evidence base. The body of evidence on ultrasonography in critical care is exponentially growing. However, a large number of studies do not guarantee high-quality data. Part 2 of this thesis adds to the evidence of this field. To start, it should be emphasised that this part of this thesis focuses on ultrasonography of patients in the ICU setting (Critical Care Ultrasound; CCUS), that is, not in a cardiology setting. This is important to realise as these two settings differ in training and perspective, which is reviewed in chapter 5.

CCUS can be used in Intensive Care Medicine for multiple purposes: diagnostic, monitoring, and guiding invasive procedures. The initial technical limitations in the early days (size, costs and image quality) prohibited its widespread use in the ICU. Today, only patient-related factors limit its use in ICU patients, and ultrasonography is now part of the daily routine in ICU care. Still, many challenges remain with its use, for instance, learning curves, availability, certification, potential confounders, and expert validation. One of the major limitations of ultrasonography is its operator-dependency as the US is not easy to use. The operator should have had minimal theoretical and practical training in order to 1. be able to operate the ultrasonography machine, 2. know how to obtain standardised images, 3. interpret the acquired images, 4. perform accurate measurements on the images obtained and 5. most important, be aware of the potential pitfalls in CCUS.[11,12]

Training and image acquisition

The ICU operator faces a more challenging setting than any other controlled elective outpatient setting concerning patient positioning, image acquisition, image quality, and image interpretation. Ultrasonography is not a technique that is feasible in each ICU patient. Training and expertise improve image acquisition quality; however, even the most experienced ultrasonographer cannot acquire images in each critically ill patient due to patient-related factors (i.e., chest tubes, wounds). This limitation is illustrated by the large variety in successful image acquisition percentages in CCUS studies.[13]

There are multiple possible strategies for an operator to become highly skilled in ultrasonography technique. There is scarce literature on which training program is either sufficient or superior. The training or the trainer (cardiologist or CCUS expert) probably does not matter that much as long as the training includes theoretical concepts and practical hands-on training. Certain factors are important to acknowledge before becoming a capable CCUS expert, including the desired level of competency, training duration, supervision during training and, probably more critical, afterwards. However, there are no clear guidelines on how long the training should be for each level of competency: for the 'basic level' structured approach CCUS there are no recommendations, while expert statements suggest a minimum of 10 hours theoretical and 30 to 50 or even 100 ultrasonography examinations for level 1 competency in basic transthoracic CCUS of the heart.[14-16]

The teacher should know each trainee's ultrasound competency to facilitate learning, irrespective of whether being a student, fellow or colleague. This can be achieved by standardising some basic level of competency[16], defined by the successful completion of a CCUS course with a structured approach of a critically ill patient.[17-19] Similar to obtaining your driver's licence: after achieving the basic competency, you are now capable of using the technique in a 'standardised manner', but still need many miles to go (preferably with feedback) before becoming a skilled and reliable driver. One reaches a certain level of competency and may grow from there: the so-called individual learning curve. However, feedback is needed to improve further and become an expert, preferably by immediate supervision in the post-certification period. Ultrasound experts should facilitate not only the ultrasonography course but also this post-course supervision. This is still a gap in nearly all ultrasonography courses and organisations where many courses are provided without specific 'aftercare'. Unsurprising, this gap is equally a limitation in all research involving ultrasonography. Only a few studies on CCUS included and facilitated immediate supervision.[20,21] In our cohort study, we had organised immediate supervision by ultrasonography experienced students and indirect supervision by experts.

Interpretation of images and performing measurements

Being capable of acquiring the image does not automatically mean that one can also interpret the image. The same applies to performing measurements on the images obtained. It appears that it takes time and experience to learn what is expected and what is abnormal, what is clinically acceptable and what is not; one needs to build a clinical framework of reference images. This all refers back to the number of patients in which the US was used to obtain images and interpret them. Experience is built over time, and after a specific time, an expert trusts the novice. The necessary time for this learning curve severely depends on the background

and experience of the novice: how much time it takes to build this frame of reference, including the most frequently observed disease states in the ICU. While immediate supervision helps the initial learning curve, indirect supervision can also achieve progress, especially in the later phase.

In a research setting, adequate image interpretation and measurements with a low risk of bias can only be achieved by independent expert validation. A few studies on CCUS had their images judged by (independent) experienced cardiologists for quality control. In our cohort study, an independent core echo laboratory evaluated all images and measurements. Future studies should apply independent data validation to increase the validity of their ultrasonography observations.

Clinical practice

How should the data obtained by the studies in this thesis affect our daily practice in the ICU? The one most important lesson learned is that no firm statements can be made referring to the evidence base: there are far more questions than conclusions. However, accumulating data suggests that it is better to not use dopamine as the first-line inotrope for hemodynamic support in patients with shock. This is likely due to the high doses given in this setting, exposing patients to detrimental harm (tachyarrhythmias and high myocardial oxygen demand) in the absence of substantial benefit. For other conditions and settings of critically ill patients with cardiac dysfunction and the use of inotropes, including milrinone and levosimendan, no recommendations can be made.

Ultrasonography is a promising tool and certainly has added value in the ICU, for instance, in identifying the cause of shock by assessing cardiac (dys)function and detecting pulmonary oedema. It is essential to realise that CCUS is not a standalone technique but should always be considered one of the many clinical variables when assessing the patient's condition. So, every ICU physician should be sufficiently skilled to use the technique adequately at some basic level of competency, just as any other device in the ICU, e.g. a ventilator or a pulmonary artery catheter measurement. In obtaining standardised, high-quality images and adequate measurements, it is a good development (for novices) that manufacturers implement artificial intelligence (AI) in their ultrasound machines.[22]

What should we do with the other inotropes? Should we go back to trial and error? The problem with the trial-and-error approach is the limited time available for physicians to judge the effects on the conditions of the patient. Nevertheless, there is still the philosophical argument: why burden the patient in his/her last living hours with additional harm in the absence of any proven benefit? A failing cardio(vascular) system is bound to end in death by the cascade of multi-organ failure. So, there

is no time to wait for a (supposed) beneficial effect, in addition to the risk to prescribe another inotrope after the lack of effect of the previous with continued deterioration. Maybe this approach may stabilise the patient, but more likely, it does not. In a sequence of deterioration and multiple interventions simultaneously, the actual problem is that one does not know which intervention was helpful and which intervention was harmful, also blurred by the tsunami of confounders present. Any confidence in a causal interpretation of effect is unjustified as each patient with cardiac dysfunction is different. So, does the approach, which was supposed to have worked last time, work now? Nobody knows.

Should we employ a pathophysiological rationale approach in daily clinical practice? This is essentially the expert opinion that sets us back to the era of trial and error. With this approach in mind, one can equally well argue against the use of any inotropes in cardiac dysfunction from the perspective of the uselessness of 'whipping an exhausted horse' (i.e., the failing heart). Should we just skip the inotropes and apply temporary mechanical circulatory support devices earlier? Regarding the latter, there is equally a lack of evidence to support any such approach.[23] And what about the reliability of the variables and devices used to evaluate and interpret the (beneficial) effects of inotropes. Literature is lacking to support daily clinical practice. Experts recently came to the same conclusion and arrived at recommendations in line with the data in this thesis.[24] More importantly, they launched a research agenda which could lead to future answers on which variable to use as a trigger and/or guidance, a set of patient-important outcomes and which inotrope to use in which patients in shock. Whether the answers will arrive in the coming years is heavily dependent on the willingness of the stakeholders to collaborate in this field.

Eventually, four recommendations should be made regarding the use of inotropes and the application of ultrasonography in the setting of suspected cardiac dysfunction in critically ill patients.[25] These are: 1. use ultrasonography to assess if there is cardiac dysfunction and evaluate the potential cause of the cardiac dysfunction, 2. only prescribe inotropes based on sound arguments, that is a measured low cardiac output state, and 3. use any device to measure and monitor the one thing that an inotrope is supposed to enhance: the cardiac output. It probably does not matter that much which device you use (a pulmonary artery catheter, PICCO or (repeated) ultrasound), depending on your personal preferences, just as long as the device is used correctly. 4. always consider not to use inotropes in order to protect your patient from additional harm.

This thesis may guide us towards acknowledging uncertainties and respecting the first principle of doing no further harm to patients. Restricted use of inotropes and ultrasound might, therefore, be more beneficial than one would expect.

References

1. Higgins JPT, Savović J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomised trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 (updated July 2019). Cochrane, 2019. Available from www.training.cochrane.org/handbook
2. Page MJ, Higgins JPT, Clayton G, Sterne JA, Hróbjartsson A, Savović J. Empirical evidence of study design biases in randomised trials: systematic review of meta-epidemiological studies. PLoS One 2016;11:e0159267. doi:10.1371/journal.pone.0159267 pmid:27398997
3. Koster TM, Wetterslev J, Gluud C, Keus F, van der Horst ICC. Systematic overview and critical appraisal of meta-analyses of interventions in intensive care medicine. Acta Anaesthesiol Scand 2018;62:1041–1049. DOI: 10.1111/aas.13147
4. Anthon CT, Granholm A, Perner A, Laake JH, Møller MH. No firm evidence that lack of blinding affects estimates of mortality in randomized clinical trials of intensive care interventions: a systematic review and meta-analysis. J Clin Epidemiol 2018 Aug;100:71-81. doi: 10.1016/j.jclinepi.2018.04.016. Epub 2018 Apr 26. PMID: 29705093
5. Brok J, Thorlund K, Wetterslev J, Gluud C. Apparently conclusive meta-analyses may be inconclusive--Trial sequential analysis adjustment of random error risk due to repetitive testing of accumulating data in apparently conclusive neonatal meta-analyses. Int J Epidemiol 2009 Feb;38(1):287-98. doi: 10.1093/ije/dyn188. Epub 2008 Sep 29. PMID: 18824466.
6. Koster TM, Wetterslev J, Gluud C, Jakobsen JC, Kaufmann T, Eck RJ, Koster G, et al. Apparently conclusive meta-analyses on interventions in critical care may be inconclusive-a meta-epidemiological study. J Clin Epidemiol 2019 Oct;114:1-10. doi: 10.1016/j.jclinepi.2019.05.011. Epub 2019 Jun 11. PMID: 31200004
7. Wetterslev J, Jakobsen JC, Gluud C. Trial sequential analysis in systematic reviews with meta-analysis. BMC Med Res Methodol 2017;17(1):7.
8. Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ, GRADE Working Group. What is "quality of evidence" and why is it important to clinicians? BMJ 2008;336:995–998. doi:10.1136/bmj.39490.551019.BE
9. Gargon E, Gorst SL, Williamson PR. Choosing important health outcomes for comparative effectiveness research: 5th annual update to a systematic review of core outcome sets for research. PLoS One 2019;14(12):e0225980.
10. Wilkinson J, Arnold KF, Murray EJ, van Smeden M, Carr K, Sippy R, de Kamps M, et al. Time to reality check the promises of machine learning-powered precision medicine. Lancet Digit Health 2020 Dec;2(12):e677-e680. doi.org/10.1016/S2589-7500(20)30200-4
11. Blanco P, Volpicelli G. Common pitfalls in point-of-care ultrasound: a practical guide for emergency and critical care physicians. Crit Ultrasound J 2016;8:15
12. Orde S, Slama M, Hilton A, Yastrebov K, McLean A. Pearls and pitfalls in comprehensive critical care echocardiography. Crit Care 2017;21:279 https://doi.org/10.1186/s13054-017-1866-z
13. Koster G, van der Horst ICC. Critical care ultrasonography in circulatory shock. Curr Opin Crit Care 2017;23(4):326-333 doi: 10.1097/MCC.0000000000000428
14. Neskovic AN, Skinner H, Price S, Via G, De Hert S, Stankovic I, Galderisi M, et al. Focus cardiac ultrasound core curriculum and core syllabus of the European Association of Cardiovascular Imaging. Eur Heart J Cardiovasc Imaging 2018;19:475–481 https://doi.org/10.1093/ehjci/jey006

15. International expert statement on training standards for critical care ultrasonography. Intensive Care Med 2011;37:1077–1083 . https://doi.org/10.1007/s00134-011-2246-9

16. Price S, Via G, Sloth E, Guarracino F, Breitkreutz R, Catena E, Talmor D. Echocardiography practice, training and accreditation in the intensive care: document for the World Interactive Network Focused on Critical Ultrasound (WINFOCUS). Cardiovasc Ultrasound 2008;6:49 https://doi.org/10.1186/1476-7120-6-49

17. Perera P, Mailhot T, Riley D, Mandavia D. The RUSH exam: rapid ultrasound in SHock in the evaluation of the critically lll. Emerg Med Clin North Am 2010;28(1):29–56

18. Jensen MB, Sloth E, Larsen KM, Schmidt MB. Transthoracic echocardiography for cardiopulmonary monitoring in intensive care. Eur J Anaesthesiol 2004;21:700-707

19. Breitkreutz R, Walcher F, Seeger F. Focused echocardiographic evaluation in resuscitation management: Concept of an advanced life support-conformed algorithm. Crit Care Med 2007, 35: S150-S161

20. Manasia AR, Nagaraj HM, Kodali RB, Croft LB, Oropello JM, Kohli-Seth R, Leibowitz AB, et al. Feasibility and potential clinical utility of goal-directed transthoracic echocardiography performed by noncardiologist intensivists using a small hand-carried device (SonoHeart) in critically ill patients. J Cardiothorac Vasc Anesth 2005;19:155–159

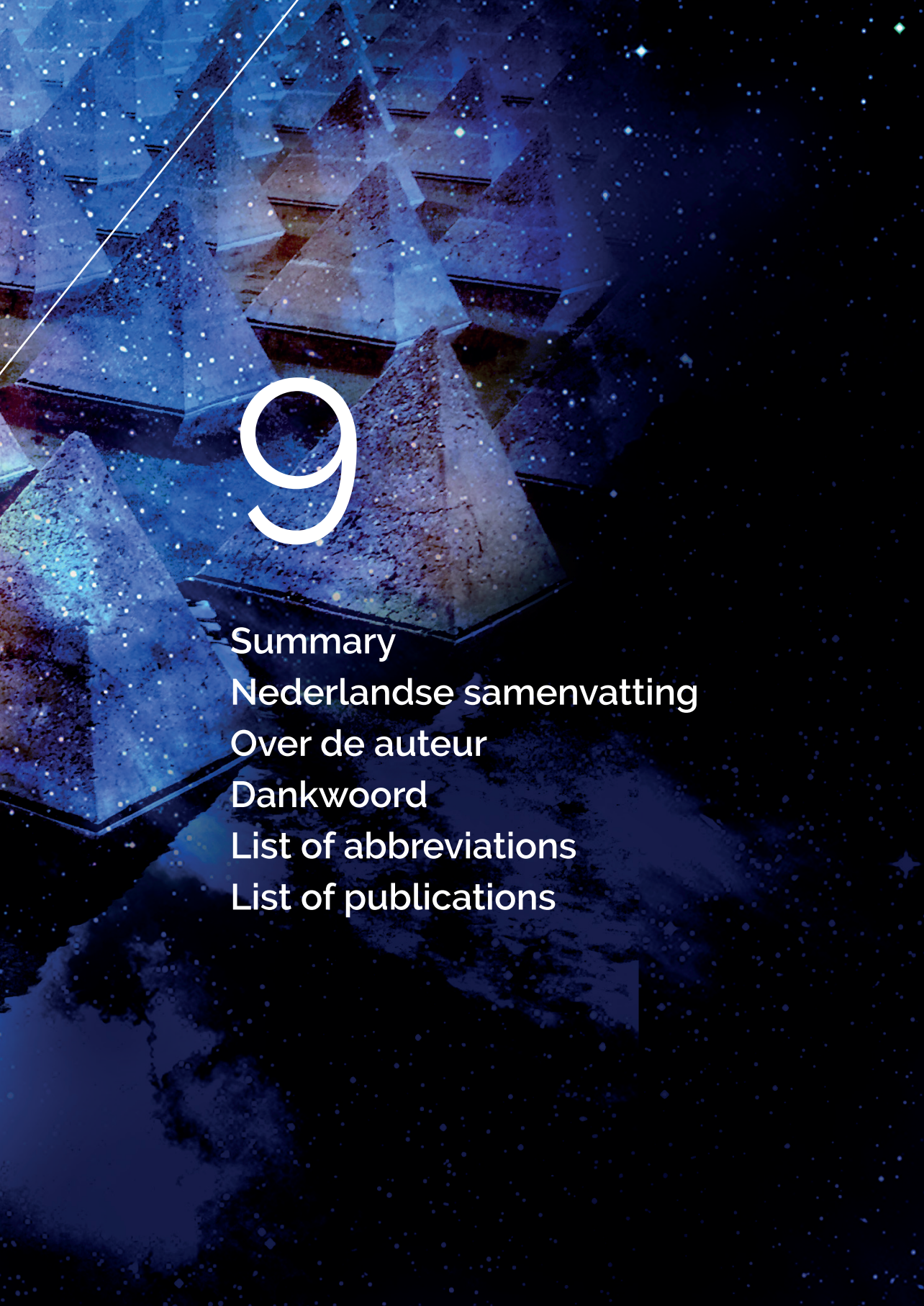
21. Alsaddique A, Royse AG, Royse CF, Mobeirek A, El Shaer F, AlBackr H, Fouda M, et al. Repeated monitoring with transthoracic echocardiography and lung ultrasound after cardiac surgery: feasibility and impact on diagnosis. J Cardiothorac Vasc Anesth 2016; 30:406–412

22. Muse ED, Topol EJ. Guiding ultrasound image capture with artificial intelligence. The Lancet 2020;396(10253):P749. doi: https://doi.org/10.1016/S0140-6736(20)31875-4

23. Ni hIci T, Boardman HM, Baig K, Stafford JL, Cernei C, Bodger O, Westaby S. Mechanical assist devices for acute cardiogenic shock. Cochrane Database Syst Rev 2020 Jun 4;6(6):CD013002 doi: 10.1002/14651858.CD013002.pub2

24. Scheeren TWL, Bakker J, Kaufmann T, Annane D, Asfar P, Boerma EC, Cecconi M, et.al. Current use of inotropes in circulatory shock. Ann Intensive Care 2021 Jan 29;11(1):21. doi: 10.1186/s13613-021-00806-8.

25. Farmakis D, Agostoni P, Baholli L, Bautin A, Comin-Colet J, Crespo-Leiro MG, Fedele F, et al. A pragmatic approach to the use of inotropes for the management of acute and advanced heart failure: An expert panel consensus. Int J Cardiol 2019;297:83-90



9

Summary
Nederlandse samenvatting
Over de auteur
Dankwoord
List of abbreviations
List of publications

Summary

Overall aim

This thesis evaluated the evidence on inotropes in critically ill patients and the evidence for monitoring hemodynamic variables using ultrasonography in critically ill patients. Systematic reviews with meta-analyses were conducted to evaluate the effects of inotropes, and substudies of a cohort study were performed to evaluate hemodynamic monitoring using ultrasonography.

Evidence on the use of inotropes in critically ill patients with cardiac dysfunction

Inotropes for critically ill patients' treatment are common; however, high-quality data on the beneficial and harmful effects on patient-relevant outcomes are scarce. Interpretation of data is hampered by heterogeneity in (amongst others) selection criteria, interventions (dosing and timing of inotrope), clinical settings, and outcome definitions. A systematic review with a meta-analysis on dopamine in critically ill patients with cardiac dysfunction showed that the trials had high risks of systematic and random errors and that dopamine was not associated with beneficial nor harmful effects [chapter 2]. A post-hoc analysis of dopamine administration in critically ill patients without documented cardiac failure, however, suggested an association with harm. In the systematic review on milrinone, a more recent developed inotrope, similar results on primary and secondary analyses were observed: data was limited, the included trials had high risks of bias, and they showed no beneficial or harmful effects in the population at interest [chapter 3]. Levosimendan, one of the latest class of inotropes, was also evaluated in a systematic review with meta-analyses. The amount of data on levosimendan was considerably larger compared to the other inotropes [chapter 4]. However, data was mostly of low quality, similar to the older data of the other inotropes; only one trial had a low risk of bias. The results were inconclusive about the beneficial or harmful effects, irrespective of the clinical setting, and analyses suggested that many more patients need to be accrued to arrive at more robust conclusions.

Overall, current data does not support nor refute the use of inotropes in critically ill patients with cardiac dysfunction: clinicians face a difficult choice in patient management.

Current evidence on the use of ultrasonography for hemodynamic assessment in critically ill patients

As with any novel diagnostic tool, ultrasonography has come a long way since its early beginning with the development of B-mode ultrasound in the 1950s. Predominantly used by cardiologists, radiologists and gynaecologist/obstetricians, the technique has become available to any medical worker who is willing to

become a trained ultrasonographer. The speed with which miniaturisation has evolved is astonishing, and this advancement opened up whole new territories for its use. Along with expanding territories for ultrasonography, concerns and questions have been raised by the 'sitting authorities' on who is allowed to use ultrasonography, i.e., who is an expert and who is not? Despite these concerns, it has now been recognised that the technique can no longer be restricted to highly specialised personnel. Restraints in time, money and/or staffing prohibit selective use of ultrasonography in all inpatients by cardiologist and/or radiologists.

Intensive care is an ideal setting for a diagnostic tool that is fast, non-ionising and delivers immediate answers to binary clinical questions. Therefore, it is unsurprising that many intensivists have embraced ultrasonography, resulting in many courses and – protocols endorsed by national and international intensive care societies. However, no quality control programmes have been developed yet, which is problematic since ultrasonography is an operator-dependent technique. Authorities recognise the importance of such quality control programmes, but implementation in the field still has to follow. This lack of quality control is a limitation that can also be identified in all the conducted trials where independent quality control of data is minimal. This makes the comparability of image acquisition and results between the trials difficult [chapter 5].

In addition, image acquisition in critically ill patients is hampered by multiple factors prohibiting the use of ultrasonography in each patient [chapter 5,6,7]. Ultrasonography data should be considered complementary to all the information collected at the bedside, including laboratory and all other hemodynamic monitoring data [chapter 5]. It is essential to realise that ultrasonography alone should never on itself impact patient-relevant outcomes. Instead, ultrasonography should be considered a valuable asset, potentially delivering answers quickly and providing individual patient-tailored management options. Still, this has yet to be proven.

Chapter 6 evaluated the use of ultrasonographically measured cardiac output (CO) by trained medical students. Medical students seemed capable of obtaining sufficient quality CCUS images for CO measurement in the majority of critically ill patients. The CO measurements themselves by medical students, however, had a poor agreement with expert measurements. Experts remain indispensable for reliable CO measurements [chapter 6].

Physical examination is still a cornerstone of medical practice; however, the overall diagnostic accuracy is poor. In chapter 7, a planned sub-study of the SICS showed poor agreement between lung ultrasound and auscultation for detecting pulmonary oedema. It seems that the ultrasound probe outperforms the stethoscope. How this result will influence medical practice remains to be seen since many ICU physicians cannot perform ultrasonography.

Nederlandse samenvatting

Doel

Dit proefschrift evalueerde het wetenschappelijke bewijs voor hemodynamische interventies met inotropica en hemodynamische monitoring middels 'critical care ultrasonography' (CCUS) in een kritisch zieke patiëntenpopulatie. Ten behoeve van het informeren van clinici over bovenstaande doelen zijn er meerdere systematische reviews met meta-analyses uitgevoerd en zijn er substudies van een cohort studie verricht naar het gebruik van echografie.

Huidig bewijs voor het gebruik van intropica in kritisch zieke patiënten met cardiale dysfunctie.

Het toepassen van inotropica in de behandeling van kritisch ziekte patiënten is dagelijkse klinische praktijk. Echter wetenschappelijk bewijs van hoge kwaliteit over zowel de voor- als nadelige effecten van inotropica op patiënt relevante uitkomsten, ontbreekt. De beschikbare literatuur wordt gehinderd door heterogeniciteit in onder andere inclusie criteria, interventies, inotropie doseringen, klinische setting, en primaire uitkomstmaat.

De systematische review met meta-analyse over toediening van dopamine in kritisch zieke patiënten met cardiale dysfunctie toonde aan dat er weinig data beschikbaar was, en de data die beschikbaar was van lage methodologische kwaliteit was (m.a.w. hoog risico op systematische - en toevals fouten). Het gebruik van dopamine bleek op basis van deze data niet geassocieerd te zijn met voor- noch nadelige effecten [hoofdstuk 2]. Een post-hoc analyse liet echter zien dat in kritisch zieke patiënten waarbij de meerderheid geen gedocumenteerde cardiale dysfunctie had, dopamine geassocieerd was met schade en indien er meer data bij zou komen het effect op z'n best nog futiel zou kunnen zijn.

Voor milrinone, een meer recenter ontwikkeld inotropicum, was eveneens niet veel data beschikbaar, en de data die er was werd eveneens gehinderd door systematische fouten en toonde geen positieve effecten in de kritisch zieke patiënten met cardiale dysfunctie [hoofdstuk 3].

Levosimendan, een medicament uit de tot nog toe nieuwste klasse inotropica, werd geëvalueerd in een systematische review met meta-analyse en er bleek meer relevante data beschikbaar te zijn in vergelijking met de oudere inotropica [hoofdstuk 4]. Echter, in lijn met de literatuur van de oudere inotropica, was ook deze data grotendeels van lage methodologische kwaliteit; slechts 1 studie werd als laag risico op systematische fouten beoordeeld. Uit de data bleek dat het gebruik van levosimendan in kritisch zieke patiënten met cardiale dysfunctie niet

geassocieerd was met positieve danwel negatieve effecten, ongeacht de klinische setting. In overeenstemming met de andere onderzochte inotropica, dienen er nog vele patiënten in studies geïncludeerd te worden voordat een definitief antwoord gegeven kan worden.

De eindconclusie is derhalve dat de tot nog toe beschikbare literatuur ten aanzien van de onderzochte inotropica, het gebruik van deze medicijnen in kritisch zieke patiënten met cardiale dysfunctie noch kan ondersteunen noch weerleggen: dat brengt clinici in een lastig pakket ten aanzien van de management van deze patiënten.

Huidig bewijs voor het gebruik van echografie ten behoeve van de hemodynamische beoordeling van de kritisch zieke patiënt

Echografie heeft een lange weg afgelegd als nieuw diagnostisch instrument sinds het begin van B-mode echo(cardio)grafie in de jaren 50. Initieel werd echografie gebruik door cardiologen, radiologen, gynaecologen/obstetristen, maar de techniek is nu feitelijk beschikbaar voor iedere dokter die zich wil bekwamen in echografie. De snelheid waarmee technische verbeteringen en miniaturisatie hebben plaatsgevonden in de echografie is verbazingwekkend en heeft ertoe geleid dat er vele nieuwe toepassingsgebieden zijn. Zoals wel vaker gaat de uitbreiding sneller dan dat de 'gevestigde orde' kan bijhouden en ontstaan er zorgen over wie het echoapparaat mag bedienen. Vragen als wie is expert en wie niet blijven nog onbeantwoord. Hedendaags is men er wel over uit dat echografie niet meer is voorbehouden aan hooggekwalificeerd personeel. Beperkingen in tijd, geld en personeel verhinderen simpelweg dat alle patiënten in een afzienbare tijd geëchood kunnen worden door een cardioloog of radioloog.

De intensive care (IC) is bij uitstek de plek voor een diagnostische tool die snel inzetbaar is, niet schadelijk is en in staat is vragen van clinici direct te beantwoorden. Echografie wordt dan ook door vele intensivisten omarmd. Er zijn vele echografie cursussen en – protocollen die gevolgd kunnen worden bij zowel nationale als internationale intensive care organisaties. Er is echter geen controle op de kwaliteit, wat grote gevolgen kan hebben aangezien echografie een techniek betreft die erg afhankelijk is van degene die het uitvoert. Dit wordt erkend door autoriteiten, echter implementatie van kwaliteitscontrole programma's om dit te borgen in het veld staan nog in de kinderschoenen. Dit is eveneens zichtbaar in de CCUS onderzoeken waar onafhankelijke data controle erg zeldzaam is. Hiermee wordt het onderling vergelijken van studies ten aanzien van de beelden acquisitie en resultaten bemoeilijkt [hoofdstuk 5].

Verder dienen we ons te realiseren dat beeld acquisitie in kritisch zieke patiënten wordt beperkt door verschillende factoren wat ervoor zorgt dat we niet bij elke

IC patiënt echografisch beeldmateriaal kunnen verkrijgen [hoofdstuk 5,6,7]. Echografische data dient beschouwd te worden als complementair zijnde aan andere bedside -, laboratorium informatie en andere hemodynamische monitoring [hoofdstuk 5]. Het is belangrijk te beseffen dat echografie nooit op zichzelf in staat is om een grote impact te hebben op patiënt relevante uitkomsten. Het kan echter wel van waarde zijn bij het snel geven van antwoorden, waarna de behandeling op maat kan worden gegeven. Dit alles zal echter nog bewezen moeten worden.

In hoofdstuk 6 hebben we in een substudie van een grote cohort studie op de IC de inzet van getrainde medisch studenten onderzocht in het echocardiografisch verkrijgen van een cardiac output (CO), een belangrijke hemodynamische variabele. Medisch studenten bleken goed in staat om in de meerderheid van de kritisch zieke patiënten echografische beelden te verkrijgen van voldoende kwaliteit voor een CO meting. De CO metingen door de studenten bleken echter slecht overeen te komen met de metingen van de expert. Dit bevestigt dat experts (ofwel een 'expert deskundigheidsniveau') absoluut noodzakelijk blijven voor de validiteit van de CO meting [hoofdstuk 6].

Lichamelijk onderzoek is nog altijd een vast onderdeel in de medische praktijk, ondanks de beperkte sensitiviteit en specificiteit ten aanzien van het stellen van een diagnose. In hoofdstuk 7 staat de uitgevoerde substudie van een cohort onderzoek beschreven welke aantoonde dat er een slechte overeenkomst is tussen longechografie en auscultatie van de longen voor het detecteren van longoedeem. Het lijkt verstandiger om voor deze vraag de echoprobe te pakken dan de stethoscoop. Hoe deze resultaten de medische praktijk gaan veranderen is onzeker aangezien er nog veel IC artsen zijn die echografie als vaardigheid niet beheersen.

Over de auteur

Geert Koster werd geboren op 24 augustus 1977 te Leiden. In 1979 verhuisde het gezin naar Leeuwarden. In 1995 behaalde hij zijn VWO diploma aan de Stedelijke Scholen Gemeenschap te Leeuwarden. Na uitgeloot te zijn voor de studie Geneeskunde besloot hij Biologie aan de Rijksuniversiteit Groningen te studeren en behaalde in 1996 zijn propedeuse. Omdat hij voor de tweede maal werd uitgeloot, besloot hij om werkervaring op te doen via een uitzendbureau bij diverse bedrijven en in diverse functies. In 1997 werd hij uiteindelijk ingeloot en begon hij met de studie Geneeskunde in Groningen. In 2003 behaalde hij het artsexamen. Zijn eerste baan als arts was als poortarts voor de interne geneeskunde in de voormalige locatie Van Ketwich van het Martini Ziekenhuis te Groningen, na een half jaar gevolgd door arts-niet in opleiding bij de afdeling cardiologie in het Martini Ziekenhuis te Groningen. In 2004 volgde de overstap naar de afdeling cardiologie in het Universitair Medisch Centrum Groningen (UMCG), waarbij hij op de diverse afdelingen van de cardiologie werkzaam was, inclusief het revalidatiecentrum Beatrixoord. In 2007 begon hij aan zijn opleiding tot cardioloog, waarbij hij gedurende zijn opleiding wisselend in het Martini Ziekenhuis en het UMCG stages gelopen heeft. In 2012 rondde hij zijn opleiding tot cardioloog af en besloot zich daarna verder te specialiseren in Intensive Care Geneeskunde. Dit fellowship vond plaats op de Intensive Care van het UMCG. Aldaar ontstond ook het idee om onderzoek te doen en onder de bezielende begeleiding van Eric Keus en Iwan van der Horst leidde dit tot dit proefschrift. In 2015 ronde hij het fellowship af waarna hij zich cardioloog-intensivist mocht noemen. Sindsdien is hij als staflid werkzaam op de Intensive Care Volwassenen van het UMCG. Naast zijn klinische werkzaamheden aldaar is hij, als stagebegeleider en plaatsvervangend opleider, betrokken bij de opleiding van diverse functies binnen de afdeling, waaronder fellows in opleiding tot intensivist. Verder zet hij zich in voor (multidisciplinair) onderwijs in allerlei vormen en het verbeteren van de afdeling op het gebied van onderwijs, opleiding, samenwerking en ontwikkeling.

Dankwoord

In tegenstelling tot wat men normaliter van mij verwacht, houd ik het hier kort en wil ik iedereen bedanken voor zijn/haar bijdrage aan dit proefschrift.

List of abbreviations

AHF	Acute heart failure
AP5CH	Apical five chamber view
BLUE-protocol	Bedside Lung Ultrasound in Emergency – protocol
CABG	Coronary artery bypass graft
CCUS	Critical care ultrasonography
CHF	Chronic heart failure
CI	Confidence interval
CO	Cardiac output
CPB	Cardiopulmonary bypass
DARIS or DIS	Diversity adjusted required information size
EACVI	European Association of Cardiovascular Imaging
GRADE	Grade of Recommendations Assessment, Development and Evaluation
IABP	Intra-aortic balloon pump
ICU	Intensive Care Unit
IQR	Interquartile ranges
LCO	Low cardiac output
LOA	Limits of agreement
LVEF	Left ventricular ejection fraction
LVOT(d)	Left ventricular outflow tract (diameter)
LUS	Lung ultrasound
MI	Myocardial infarction
PLAX	Parasternal long axis
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
RCT	Randomised controlled trial
SAE	Serious adverse events
SD	Standard deviation
SE	Standard error
SICS	Simple Intensive Care Studies
SVT	Supraventricular tachyarrhythmia
TSA	Trial sequential analysis
VF	Ventricular fibrillation
VT	Ventricular tachycardia
VTI	Velocity time integral

List of publications

1. Scheeren TWL, Bakker J, Kaufmann T, Annane D, Asfar P, Boerma EC, Cecconi M, Chew MS, Cholley B, Cronhjort M, De Backer D, Dubin A, Dünser MW, Duranteau J, Gordon AC, Hajjar LA, Hamzaoui O, Hernandez G, Kanoore Edul V, Koster G, Landoni G, Leone M, Levy B, Martin C, Mebazaa A, Monnet X, Morelli A, Payen D, Pearse RM, Pinsky MR, Radermacher P, Reuter DA, Sakr Y, Sander M, Saugel B, Singer M, Squara P, Vieillard-Baron A, Vignon P, Vincent JL, van der Horst ICC, Vistisen ST, Teboul JL. Current use of inotropes in circulatory shock. Ann Intensive Care. 2021 Jan 29;11(1):21. doi: 10.1186/s13613-021-00806-8.
2. Cox EGM, Wiersema R, Wong A, van der Horst ICC; SICS Study Group. Six versus eight and twenty-eight scan sites for B-line assessment: differences in examination time and findings. Intensive Care Med. 2020 May;46(5):1063-1064. doi: 10.1007/s00134-020-06004-7. Epub 2020 Mar 19.
3. Kaufmann T, Cox EGM, Wiersema R, Hiemstra B, Eck RJ, Koster G, Scheeren TWL, Keus F, Saugel B, van der Horst ICC; SICS Study Group. Non-invasive oscillometric versus invasive arterial blood pressure measurements in critically ill patients: A post hoc analysis of a prospective observational study. J Crit Care. 2020 Jun;57:118-123. doi: 10.1016/j.jcrc.2020.02.013. Epub 2020 Feb 22.
4. Cox EGM, Koster G, Baron A, Kaufmann T, Eck RJ, Veenstra TC, Hiemstra B, Wong A, Kwee TC, Tulleken JE, Keus F, Wiersema R, van der Horst ICC; SICS Study Group. Should the ultrasound probe replace your stethoscope? A SICS-I sub-study comparing lung ultrasound and pulmonary auscultation in the critically ill. Crit Care. 2020 Jan 13;24(1):14. doi: 10.1186/s13054-019-2719-8.
5. Koster G, Kaufmann T, Hiemstra B, Wiersema R, Vos ME, Dijkhuizen D, Wong A, Scheeren TWL, Hummel YM, Keus F, van der Horst ICC. Feasibility of cardiac output measurements in critically ill patients by medical students. Ultrasound J. 2020 Jan 8;12(1):1. doi: 10.1186/s13089-020-0152-5.
6. Kaufmann T, Castela Forte J, Hiemstra B, Wiering MA, Grzegorczyk M, Epema AH, van der Horst ICC; SICS Study Group. A Bayesian Network Analysis of the Diagnostic Process and Its Accuracy to Determine How Clinicians Estimate Cardiac Function in Critically Ill Patients: Prospective Observational Cohort Study. JMIR Med Inform. 2019 Oct 30;7(4):e15358. doi: 10.2196/15358.
7. Hiemstra B, Eck RJ, Wiersema R, Kaufmann T, Koster G, Scheeren TWL, Snieder H, Perner A, Pettilä V, Wetterslev J, Keus F, van der Horst ICC; SICS Study Group. Clinical Examination for the Prediction of Mortality in the Critically Ill: The Simple Intensive Care Studies-I. Crit Care Med. 2019 Oct;47(10):1301-1309. doi: 10.1097/CCM.0000000000003897.
8. Wiersema R, Koeze J, Eck RJ, Kaufmann T, Hiemstra B, Koster G, Franssen CFM, Vaara ST, Keus F, Van der Horst ICC. Clinical examination findings as predictors of acute kidney injury in critically ill patients. Acta Anaesthesiol Scand. 2020 Jan;64(1):69-74. doi: 10.1111/aas.13465. Epub 2019 Sep 30.
9. Koster TM, Wetterslev J, Gluud C, Jakobsen JC, Kaufmann T, Eck RJ, Koster G, Hiemstra B, van der Horst ICC, Keus E. Apparently conclusive meta-analyses on interventions in critical care may be inconclusive-a meta-epidemiological study. J Clin Epidemiol. 2019 Oct;114:1-10. doi: 10.1016/j.jclinepi.2019.05.011. Epub 2019 Jun 11.

10. Kaufmann T, Clement RP, Hiemstra B, Vos JJ, Scheeren TWL, Keus F, van der Horst ICC; SICS Study Group. Disagreement in cardiac output measurements between fourth-generation FloTrac and critical care ultrasonography in patients with circulatory shock: a prospective observational study. J Intensive Care. 2019 Apr 11;7:21. doi: 10.1186/s40560-019-0373-5. eCollection 2019.
11. Hiemstra B, Koster G, Wiersema R, Hummel YM, van der Harst P, Snieder H, Eck RJ, Kaufmann T, Scheeren TWL, Perner A, Wetterslev J, de Smet AMGA, Keus F, van der Horst ICC; SICS Study Group. The diagnostic accuracy of clinical examination for estimating cardiac index in critically ill patients: the Simple Intensive Care Studies-I. Intensive Care Med. 2019 Feb;45(2):190-200. doi: 10.1007/s00134-019-05527-y. Epub 2019 Jan 31.
12. Wiersema R, Castela Forte JN, Kaufmann T, de Haas RJ, Koster G, Hummel YM, Koeze J, Franssen CFM, Vos ME, Hiemstra B, Keus F, van der Horst ICC. Observational Study Protocol for Repeated Clinical Examination and Critical Care Ultrasonography Within the Simple Intensive Care Studies. J Vis Exp. 2019 Jan 16;(143). doi: 10.3791/58802.
13. Hiemstra B, Koster G, Wetterslev J, Gluud C, Jakobsen JC, Scheeren TWL, Keus F, van der Horst ICC. Dopamine in critically ill patients with cardiac dysfunction: A systematic review with meta-analysis and trial sequential analysis. Acta Anaesthesiol Scand. 2019 Apr;63(4):424-437. doi: 10.1111/aas.13294. Epub 2018 Dec 4.
14. Hiemstra B, Eck RJ, Koster G, Wetterslev J, Perner A, Pettilä V, Snieder H, Hummel YM, Wiersema R, de Smet AMGA, Keus F, van der Horst ICC; SICS Study Group. Clinical examination, critical care ultrasonography and outcomes in the critically ill: cohort profile of the Simple Intensive Care Studies-I. BMJ Open. 2017 Sep 27;7(9):e017170. doi: 10.1136/bmjopen-2017-017170.
15. Koster G, van der Horst ICC. Critical care ultrasonography in circulatory shock. Curr Opin Crit Care. 2017 Aug;23(4):326-333. doi: 10.1097/MCC.0000000000000428.
16. Koster G, Bekema HJ, Wetterslev J, Gluud C, Keus F, van der Horst IC. Milrinone for cardiac dysfunction in critically ill adult patients: a systematic review of randomised clinical trials with meta-analysis and trial sequential analysis. Intensive Care Med. 2016 Sep;42(9):1322-35. doi: 10.1007/s00134-016-4449-6. Epub 2016 Jul 22.
17. Koster G, Wetterslev J, Gluud C, Zijlstra JG, Scheeren TW, van der Horst IC, Keus F. Effects of levosimendan for low cardiac output syndrome in critically ill patients: systematic review with meta-analysis and trial sequential analysis. Intensive Care Med. 2015 Feb;41(2):203-21. doi: 10.1007/s00134-014-3604-1. Epub 2014 Dec 18.