Title

Accuracy and precision of non-invasive cardiac output monitoring devices in perioperative medicine: a systematic review and meta-analysis † † This Article is accompanied by Editorial Aew442.

Journal

British Journal of Anaesthesia, 118(3)

ISSN

0007-0912

Authors

Joosten, A
Desebbe, O
Suehiro, K
et al.

Publication Date

2017-03-01

DOI

10.1093/bja/aew461

BJA

British Journal of Anaesthesia, 118 (3): 298–310 (2017)

doi: 10.1093/bja/aew461
Review Article

REVIEW ARTICLE

Accuracy and precision of non-invasive cardiac output monitoring devices in perioperative medicine: a systematic review and meta-analysis†

A. Joosten[1,*], O. Desebbe[2], K. Suehiro[3], L. S.-L. Murphy[4], M. Essiet[5], B. Alexander[6], M.-O. Fischer[7,8], L. Barvais[1], L. Van Obbergh[1], D. Maucort-Boulch[9] and M. Cannesson[10]

[1]Department of Anesthesiology, CUB Erasme University Hospital, Université Libre de Bruxelles, 808 Route de Lennick, 1070 Brussels, Belgium, [2]Department of Anesthesiology and Critical Care, Clinique de la Sauvegarde, Lyon, France, [3]Department of Anesthesiology, Osaka City University Graduate School of Medicine, Osaka, Japan, [4]Ayala Science Library Reference Department, University of California, Irvine, CA, USA, [5]Department of Anesthesiology and Perioperative Care, University of California Irvine, CA, USA, [6]Department of Anesthesiology, University of California San Diego, CA, USA, [7]CHU de Caen, Pôle Réanimations Anesthésie SAMU/SMUR, Avenue de la Côte de Nacre, CS 30001, Caen, F-14 000, France, [8]EA 4650, Université de Caen Basse-Normandie, Esplanade de la Paix, CS 14 032, Caen, F-14 000, France, [9]Department of Biostatistics, Hospices Civils de Lyon, France and [10]Department of Anesthesiology and Perioperative Medicine, University of California Los Angeles, CA, USA

*Corresponding author. E-mail: alexandre.joosten@erasme.ulb.ac.be

Abstract

Cardiac output (CO) measurement is crucial for the guidance of therapeutic decisions in critically ill and high-risk surgical patients. Newly developed completely non-invasive CO technologies are commercially available; however, their accuracy and precision have not recently been evaluated in a meta-analysis. We conducted a systematic search using PubMed, Cochrane Library of Clinical Trials, Scopus, and Web of Science to review published data comparing CO measured by bolus thermodilution with commercially available non-invasive technologies including pulse wave transit time, non-invasive pulse contour analysis, thoracic electrical bioimpedance/bioreactance, and CO_2 rebreathing. The non-invasive CO technology was considered acceptable if the pooled estimate of percentage error was <30%, as previously recommended. Using a random-effects model, SD, pooled mean bias, and mean percentage error were calculated. An I^2 statistic was also used to evaluate the inter-study heterogeneity. A total of 37 studies (1543 patients) were included. Mean CO of both methods was 4.78 litres min^{-1}. Bias was presented as the reference method minus the tested methods in 15 studies. Only six studies assessed the random error (repeatability) of the tested device. The overall random-effects pooled bias (limits of agreement) and the percentage error were −0,13 [−2.38 , 2.12] litres min^{-1} and 47%, respectively. Inter-study sensitivity heterogeneity was high (I^2=83%, P<0.001). With a wide percentage error, completely non-invasive CO devices are not interchangeable with bolus thermodilution. Additional studies are warranted to demonstrate their role in improving the quality of care.

†This Article is accompanied by Editorial Aew442.

†This Article is accompanied by Editorial Aew442.

Key words: cardiac output; meta-analysis

Editor's key points

- Advances in non-invasive cardiac output technologies offer simpler perioperative monitoring, but their accuracy is questioned.
- This meta-analysis found modest agreement and inadequate percentage error for most technologies.
- Novel non-invasive cardiac output technologies are typically developed in relatively healthy populations; their internal algorithms may thus be inappropriate to major surgery or critical illness.
- Percentage error and trending are important variables in the evaluation of non-invasive cardiac output technologies.

Cardiac output (CO) is a fundamental component of oxygen delivery to end organs and is therefore an important variable in adequate management of critically ill and high-risk surgical patients with therapeutic actions (fluid, vasopressor, or inotrope administration). Although its value is beyond dispute for anaesthetists and critical care physicians, this variable has been shown to be undermonitored in clinical settings.[1] Within the last decade, CO has been measured with less invasive methods using newly developed completely non-invasive monitoring devices. Several techniques have emerged recently, such as pulse wave transit time (PWTT) and non-invasive pulse contour analysis (niPCA). These are in addition to older non-invasive approaches, such as thoracic electrical bioimpedance (TEB) and partial CO_2 rebreathing (CO_2r). Bolus thermodilution (TD) is still the most accepted reference method as there is still no clearly established gold standard for CO measurement in human studies.[2 3] That explains why the accuracy (bias, corresponding to the systematic error between both techniques) and the precision (SD of the bias, corresponding to the random error between both techniques) of CO measurements are usually assessed with a Bland–Altman graph, which does not statistically determine superiority of one device over another. Moreover, the precision of both methods is quantitatively dependent on the value of the mean CO. As a result, percentage of error (PE) has been proposed as an additional measured variable, and is obtained by dividing the 95% confidence interval (CI) of the bias (limits of agreement, LOA) by the mean value of both CO methods.[4]

It is important to highlight the fact that successful cardiac output monitoring relies on two things: accuracy of measurements; and the ability to detect short term changes in cardiac output value. However, we decided to focus our meta-analysis on Bland–Altman results because this is the analysis that is now most commonly performed and presented in published method-comparison studies (the trend analysis being less often performed, but equally important).

In 2010, Peyton and Chong[5] conducted a meta-analysis comparing the accuracy and precision of four minimally invasive technologies (calibrated and uncalibrated pulse contour techniques, oesophageal Doppler, CO_2r and transthoracic bioimpedance) with bolus TD, concluding that none achieved appropriate agreement with bolus thermodilution (PE<30%). Given that the conclusion was drawn in 2010 and no recent meta-analysis has been performed on the topic (despite two recent contemporary reviews),[2 6] we sought to update the meta-analysis of Peyton and Chong[5] by focusing only on non-invasive CO techniques. As a result, we have included in our meta-analysis PWTT, niPCA, bioimpedance, and CO_2r technologies (while excluding some minimally invasive or intrusive technologies that were used in the meta-analysis of Peyton and Chong,[5] such as pulse contour monitoring using an a-line, Doppler, or both). Therefore, the goal of this systematic review and meta-analysis was to assess the agreement and precision of completely non-invasive CO monitoring devices against bolus TD.

Our objectives were two-fold. Firstly, we systematically reviewed all available studies comparing CO measured with a commercially available and completely non-invasive CO monitoring device (tested method) against bolus TD (reference method) in the operating room (OR), intensive care unit (ICU), and emergency department (ED) as long as they assessed agreement and precision in adults. Secondly, we aimed to conduct a meta-analysis using the extracted data from the systematic review in order to calculate the following four variables: (i) the pooled estimate of the mean difference between the tested and reference method (bias); (ii) the pooled estimate of the SD (precision) of the bias; (iii) the pooled estimate of the 95% CI of the bias (LOA); and (iv) the pooled estimate of the PE. An acceptable agreement between the tested and the reference method was defined as a pooled estimate PE <30%, as previously recommended.[4]

Methods

This systematic review and meta-analysis was conducted following the established guidelines from the 'Preferred Reporting Items for Systematic Reviews and Meta-Analyses' (PRISMA).[7]

Eligibility criteria

We deliberately chose to focus on completely non-invasive CO devices that are commercially available in order for us to compare devices that are easy to use at the bedside by clinicians. Of note, we excluded Doppler-based devices, such as the USCOM and oesophageal Doppler. We made this decision because the USCOM is not a continuous CO monitoring device and the oesophageal Doppler is relatively intrusive when compared with other technologies available. Additionally, as we wanted to focus on non-invasive technologies, we did not include pulse contour monitoring devices (calibrated or uncalibrated), such as the Vigileo Flo Trac, PICCO, and LIDCO systems. Although our patients are completely anaesthetized and there should be no barrier to the use of minimally invasive methods, such as oesophageal Doppler and minimally invasive pulse contour technologies, our meta-analysis focuses on newer technologies that have recently been introduced to the market and not included in recent meta-analyses.

We defined the study eligibility criteria as follows: (i) published manuscripts that compared CO values measured using a commercially available and completely non-invasive CO monitoring device with CO measured by bolus TD (with either transpulmonary or right heart bolus TD); (ii) manuscripts

documenting extractable biases for CO, SD of the bias or LOA, and PE or reporting sufficient data to calculate PE; (iii) participants must have been adults (age >18 yr) and manuscripts must have reported clear patient characteristics (age, weight, height, sex, setting, sample size, and number of data points collected); and (iv) studies performed in the OR, ICU, or ED. Our exclusion criteria included prototypical technologies and all the conditions that lead to inaccurate measurements of CO by bolus TD; this essentially consisted of any patient population with intracardiac shunts or pulmonary and tricuspid valve insufficiencies. Owing to a lack of funding for translation, we excluded non-English studies except French, as three co-authors are fluent (A.J., O.D., and M.C.). Published data that were not categorized as traditional journal articles (e.g. editorials, letters, and conference papers) were also excluded.

Literature search

We systematically searched the following four databases: PubMed, Scopus (which includes coverage of EMBASE from 1996 to the present), Web of Science, and the Cochrane Library of Clinical Trials, for prospective studies published as a manuscript in English or French from January 1, 2000 to April 22, 2015. We deliberately chose to start the inclusion in 2000 because of the publication of the PE by Critchley and Critchley[4] in 1999. The full search strategy for PubMed is depicted as an example in Appendix 1. The variable search strategies for the other databases can be provided upon request. We examined the references from included studies in order to find other potentially relevant studies that were missed by the literature search (hand-search strategy). Additionally, we contacted manufacturers of commercially available monitors for potential unpublished studies. Manufacturers included Nexfin™ (BMEYE, Amsterdam, The Netherlands), NICOM™ (Cheetah Medical, Tel Aviv, Israel), Aesculon™/ICON™ (Osypka Medical, Inc., La Jolla, CA, USA), BioZ™ (CardioDynamics, San Diego, CA, USA), and NICO™ (Novametrix Medical Systems, Wallingford, CT, USA). One of us (L.S.-L.M.) is a health information specialist who designed a unique search strategy for each of the four databases. All relevant search results were imported into EndNote (Thomson Reuters, Philadelphia, PA, USA), wherein duplicate studies were removed. Lastly, we contacted every first and last author of potentially included studies, with 1 month reminders, to acquire additional information as necessary, especially when haemodynamic data were missing.

Study selection

Initially, three reviewers (A.J., K.S., and O.D.) independently screened eligible studies gathered using the above approach, also known as the eligibility assessment checklist. The second step consisted of assessing every selected study by reading the entire text. The decision to include each study was reached as follows: abstracts were classified as relevant, potentially relevant, or not relevant. Relevant abstracts were selected for a full review of the article (eligibility). Any discrepancy between the three reviewers was resolved with a final decision from the last investigator (M.C.). The flowchart for this study selection process is shown in Fig. 1.

Data-collection process and data items

A standardized Excel spreadsheet™ form was used to save information from each article. Two authors (A.J. and M.E.) collected data separately; two authors (K.S. and O.D.) checked all data abstractions. Data items of included studies are presented in Table 1. Additionally, we extracted data about authors presenting a conflict of interest with the tested device, the specified studied time plot, mean CO [interquartile range (IQR) or SD], device repeatability, coefficient of determination (R^2) and distribution of the bias (normally or not normally distributed, or unknown). The selected reviewers (O.D., K.S., and A.J.) then independently interpreted and summarized the entire text of each remaining article. Once their summaries were completed, a consensus was agreed between the three independent reports. All data were separately transferred to a standardized Excel spreadsheet™. In order for a study to qualify for inclusion in the analysis, it must have minimally presented a mean bias, the SD of the bias, and the mean PE.

The percentage error was defined in this meta-analysis as follows:

$$Percentage\ error = \frac{1.96 \times (SD\ of\ bias\ between\ both\ methods)}{0.5 \times (Mean\ non\text{-}invasive\ CO + mean\ bolus\ TD\ CO)}$$

If CO was displayed for specific time points only, we calculated the mean CO for the whole study according to the number of data per time plot. If bias and SD or the mean PE was not calculated in this manner, it was recalculated using our definition. Our definition of bias was the value of CO measured using the non-invasive method subtracted from the value of CO measured via bolus TD. When only bias and LOA were documented by a study, SD was then recalculated as follows: upper LOA−bias)/1.96. If necessary for our calculations, we attempted to contact the authors of included studies. Furthermore, we randomly (using a randomizing selection tool from random.org) extracted 10 studies and recalculated the distribution of the bias from the Bland–Altman graph displayed by the authors in order to verify the distribution of the bias (i.e. we scanned the graph, recalculated the y-axis of the coordinates of the plots, with a zero reference as the value of the bias; OOoDigitzer™, Libreoffice™, 4.4.2.2, Openoffice.org). Lastly, we reviewed the number of studies assessing CO trending.

Risk of bias in individual studies

The quality of individual studies was assessed according to Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) guidelines, modified in order to be specific for studies of method comparison. This approach was developed and used by our group in previously published meta-analyses on method-comparison studies.[45–47] O.D., K.S., A.J., and M.C. adjusted the quality assessment form to make it more applicable to this specific meta-analysis. This amended form was used by O.D., A.J., and K.S. to perform quality assessment on each included study. Each manuscript's risk for bias and applicability was systematically classified as low, high, or unclear. At the end of the assessment, the three reviewers met together to obtain the same evaluation for each article.

Summary measures

We defined the various summary measures of our analysis as follows: (i) accuracy of the measured CO monitoring device (defined as the pooled estimate bias between the tested method and the bolus TD method); (ii) pooled estimate precision of the tested CO monitoring device (defined as SD of the bias); (iii) pooled estimate of the LOA (95% CI of the bias); and (iv) the pooled estimate PE. We defined the estimate pooled PE as our

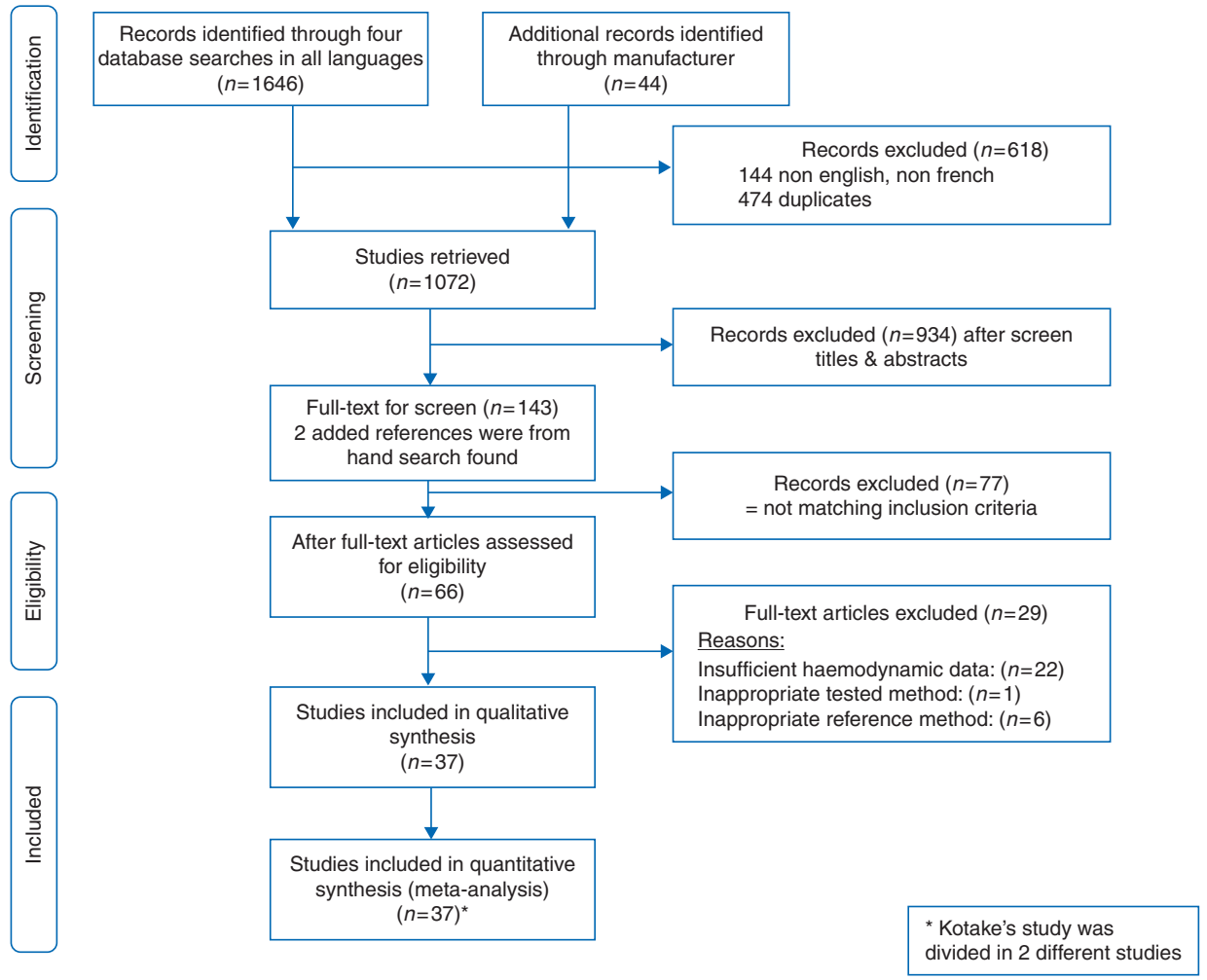

Fig 1 Flow diagram of the search strategy.

primary end point. The PE was considered to be acceptable when <30%, as previously recommended by Critchley and Critchley.[4]

Synthesis of the results

Heterogeneity could be present in bias and SD. We therefore used random-effect models to make a synthesis of pooled bias and SD.[48] Pooled PE was calculated as follows:

$$Pooled\ percentage\ of\ error = \frac{1.96 \times Pooled\ estimate(SD\ of\ the\ bias)}{Mean(Tested\ method\ CO + TD\ CO)}$$

Heterogeneity of bias and SD for the evaluated studies was calculated using a Q test and described as an I^2 index (25–50%, low heterogeneity; 50–75%, moderate heterogeneity; >75%, high heterogeneity).[49] If moderate or high heterogeneity (I^2>50%) was discovered, sensitivity analysis and meta-regression were subsequently performed. As a result, forest plots were created to view and understand the pooled estimate bias with the 95% CI (LOA of the bias).

Risk of publication bias across studies

Publication bias of the included studies was examined using standard funnel plots. These plots represent the bias of CO *vs* standard error in each study. Egger regression tests were then performed to assess the asymmetry of the Funnel plot using a significance level of 0.1, as the sample size was relatively small.[50]

Additional analysis; meta-bias

To assess the potential influences of study heterogeneity, subgroup analyses were conducted according to several characteristics, as follows: technology of the device; type of the reference method (right heart or transpulmonary bolus TD); tested method (PWTT, niPCA, CO_2r, or bioimpedance); and setting (OR, ICU, or OR+ICU). Additionally, meta-regression analyses were

Table 1 Summary of included studies. CO, cardiac output; ICU, intensive care unit; NC, not calculated; NR, not reported; OR, operating room; PAC, pulmonary artery catheter; PE*, percentage error=limits of agreement (1.96×SD)/mean CO of both methods; PE, percentage error=limits of agreement (1.96×SD)/mean CO of the reference method; PICCO, PCO2 RB, partial carbon dioxide rebreathing method; PWTT, pulse wave transit time; RAAT, radial artery applanation tonometry; RM, reference method; TEB, transthoracic electrical bioimpedance; VCM, volume clamp method

Authors	Year	Device type	Device name	RM	Sample size	Population	Setting	Mean CO of RM (SD) [range] (litres min^{-1})	Bias (SD) (litres min^{-1})	PE* (%)	PE (%)
Ameloot and colleagues[8]	2013	VCM	Nexfin	PICCO	45	Critically ill patients	ICU	6.6 (2.2)	−0.4 (1.2)	37	36
Ball and colleagues[9]	2013	PWTT	esCCO	PAC	35	Cardiac surgery patients	OR	5.4 (1.6)	−0.8 (1.43)	48	53.8
Bernstein and colleagues[10]	2005	TEB	ICG	PAC	106	Cardiac surgery patients	ICU	5.97 (1.41)	0.1 (0.94)	31	31
Bogert and colleagues[11]	2010	VCM	Nexfin	PAC	25	Cardiac surgery patients	ICU	5.45 (NC)	0.39 (0.82)	28.6	27.7
Broch and colleagues[12]	2012	VCM	Nexfin	PICCO	40	Cardiac surgery patients	OR	5.17 (1.4)	0.1 (0.62)	24	24
Bubenek-Turconi and colleagues[13]	2013	VCM	Nexfin	PAC	28	Cardiac surgery patients	ICU	5.2 (1.2)	0 (1)	38	38
Critchley and colleagues[14]	2000	TEB	NCCOM3	PAC	24	Critically ill patients	ICU	6.7 [3.6–12.9]	−1.49 (2.12)	69	62
Engoren and colleagues[15]	2005	TEB	BIOZ	PAC	46	Patients requiring a PAC	ICU	6.3 (2.2)	−1 (1.33)	43.7	41.3
Fischer and colleagues[16]	2012	VCM	Nexfin	PICCO	24	Cardiac surgery patients	ICU	4.62 (1.3)	0 (1.17)	49	50
Fischer and colleagues[17]	2013	VCM	Nexfin	PICCO	37	Cardiac surgery patients	ICU	4.3 (1.1)	−0.2 (1.18)	53	54
Gerhardt and colleagues[18]	2000	VCM	PORTAPRES	PAC	46	Patients requiring a PAC	ICU	6.1 (2.2)	−2.1 (2.55)	99	82.9
Heringlake and colleagues[19]	2007	TEB	Aesculon	PAC	29	Cardiac surgery patients	OR/ICU	4.65 (NC)	0.02 (1.73)	73.4	73.1
Hofhuizen and colleagues[20]	2014	VCM	Nexfin	PICCO	19	Cardiac surgery patients	ICU	5.9 (NC)	−0.26 (1.12)	37.9	37.2
Ishihara and colleagues[21]	2012	PWTT	esCCO	PAC	207	Cardiac surgery patients	OR/ICU	4.22 [1.3–15.5]	0.34 (1.5)	66.9	69.6
Kotake and colleagues[22]	2003	PCO2 RB	NICO 3.1	PAC	28	Aortic reconstruction	OR	4.68 (NC)	−0.58 (0.9)	40.4	37.6
Kotake and colleagues[23]	2009	PCO2 RB	NICO 4.2	PAC	21	Aortic reconstruction	OR	5.02 (NR)	0.18 (0.88)	33.4	34.3
			NICO 5		21			4.82 (NR)	0.18 (0.83)	33.2	33.8
Kothari and colleagues[24]	2003	PCO2 RB	NICO	PAC	23	Cardiac surgery patients	OR	4.41 [1.97–10.6]	0.15 (1.24)	55.3	56.2
Mekis and colleagues[25]	2008	TEB	Aesculon	PAC	16	Cardiac surgery patients	OR/ICU	NR [2.2–7.5]	−0.21 (0.78)	40	NR
Mielck and colleagues[26]	2003	PCO2 RB	NICO	PAC	22	Cardiac surgery patients	ICU	6.09 [3.3–11.8]	0.64 (1.39)	43	45.6
Murias and colleagues[27]	2002	PCO2 RB	NICO	PAC	22	Critically ill patients	ICU	6.14 (NR)	0.18 (1.39)	43	44.3
Ng and colleagues[28]	2007	PCO2 RB	NICO	PAC	12	Thoracic surgery	OR	4.57 (NR)	0.29 (0.67)	29.6	28.7
Nilsson and colleagues[29]	2001	PCO2 RB	NICO	PAC	30	Cardiac surgery patients	ICU	4.4 (0.9) [2.7–6.1]	0.16 (0.92)	40.2	40.9
Palmers and colleagues[30]	2012	PCO2 RB	NICO	PAC	23	Critically ill patients	ICU	6.7 (1.6) [3.8–11]	−1.3 (1.7)	56.2	50.8
Raue and colleagues[31]	2009	TEB	Aesculon	PICCO	30	Critically ill patients	ICU	7.04 (NR)	−0.3 (1.9)	51.8	52.9
Rocco and colleagues[32]	2004	PCO2 RB	NICO	PAC	12	Critically ill patients	ICU	7.27 (2.42)	−1.2 (1.5)	43.9	40.4
Sharma and colleagues[33]	2011	TEB	NICOMON	PAC	46	Cardiac surgery patients	ICU	4.56 (NC)	0.02 (0.84)	36.1	36.2
Simon and colleagues[34]	2009	TEB	ICG	PAC	13	Cardiac surgery patients	ICU	5.3 (1.6)	−0.5 (1.3)	50.5	48.1
Squara and colleagues[35]	2009	Bioreactance	NICOM	PICCO	20	Cardiac surgery patients	ICU	4.23 (NR)	0.22 (0.85)	38.5	39.5
Tachibana and colleagues[36]	2002	PCO2 RB	NICO	PAC	25	Cardiac surgery patients	ICU	5.28 (NC)	−0.74 (1.05)	38.9	39
Tachibana and colleagues[37]	2003	PCO2 RB	NICO	PAC	25	Cardiac surgery patients	ICU	5.96 (NC)	−0.31 (0.86)	28.4	28.3
Tachibana and colleagues[38]	2005	PCO2 RB	NICO	PAC	13	Cardiac surgery patients	ICU	5.61 (NC)	−0.16 (1.07)	38.1	37.4
Thonnerieux and colleagues[39]	2015	PWTT	esCCO	PAC	27	Cardiac surgery patients	ICU	4.87 (1.59)	0.7 (2.74)	53	56.3
Valiatti and colleagues[40]	2004	PCO2 RB	NICO	PAC	20	Acute hypoxic insufficiency	ICU	8.12 (NC)	−1.33 (2.37)	62.1	67.3
Wacharasint and colleagues[41]	2014	PWTT	esCCO	PAC	50	Cardiac surgery patients	OR	4.8 (NR)	1.2 (1.59)	57.7	65
Wagner and colleagues[42]	2015	RAAT	T-Line	PAC	50	Cardiac surgery patients	ICU	4.7 (1.2)	−0.2 (1.57)	33	33
Yamada and colleagues[43]	2012	PWTT	esCCO	PAC	213	Patients requiring a PAC	OR/ICU	3.9 [3.2–5.1]	0.13 (1.15)	56.8	57.7
Zoremba and colleagues[44]	2007	TEB	Aesculon	PAC	25	Critically ill patients	ICU	5.25 (NR)	0.05 (0.71)	26.4	26.5

performed to determine whether the weighted mean differences between the reference and tested methods were affected by study characteristics (sample size, number of measurements, publication year, and conflict of interest).

All the statistical analyses were conducted using Microsoft Excel software 2010 (Microsoft Corporation, Redmond, WA, USA), R software (version 3.0.2; R Development Core Team; *R: A Language Environment for Statistical Computing*. Vienna, Austria. ISBN 3-900051-07-0. URL: http://www.R-project.org 2013). For all analyses, $P<0.05$ was considered as statistically significant. The 95% CI was calculated for all values.

Results

Data retrieval

The initial search resulted in 1646 articles in all languages. A flow diagram outlining the study selection process is provided in Fig. 1. Of note, among the 37 selected studies, one was split into two different sections (the study by Kotake and colleagues[23] was divided into two different sections in the forest plots and in Table 1 because there were two different versions of the same tested device). It is important to note that 29 articles were excluded for failure to meet our inclusion criteria or insufficient data (Appendix 2).

Study characteristics

A total of 1543 patients were included, with a median of 97 (IQR 36–158) pairs of data per study. The median number of patients was 28 (range 21–45). The majority of studies were conducted in perioperative cardiac surgery, with the remaining studies analysing critically ill patients. Of these, 27 studies were conducted in an ICU setting, eight in the OR, and three in both ICU and OR. Parial CO_2 rebreathing was used by one device (NICO), although there were four different versions noted. Thoracic electrical bioimpedance corresponded to three similar technologies (bioimpedance, bioreactance, and electrical velocimetry) used by six devices (NICOMON™, ICG™, BoMedNCCOM3™, BioZ™, NICOM™, and Aesculon™). Non-invasive pulse contour analysis corresponded to two similar technologies used by three devices (volume clamp method with the Portapres™, the Nexfin™, and arterial tonometry with the T-Line system™). Pulse wave transit time technology corresponded to one technology and one device, the esCCO™ (Nihon Kohden, Tokyo, Japan). Technologies using CO_2r and TEB were the most represented, followed by niPCA, and then PWTT. Concerning the reference methods, 31 were conducted using right heart TD and seven using transpulmonary TD. Patient characteristics of included studies are shown in Table 1. The mean CO of the reference method was not cited seven times and its dispersion (SD or IQR) was not cited in 14 instances. The PE was calculated by the authors in 19 instances, and modified in one instance.[33] Three studies, [9 25 40] among the 10 randomly tested, [9 16 19 22 23 25 27 29 33 40] demonstrated a non-normal distribution of the bias. No study clearly stated that bias was normally distributed. Owing to a lack of abstracted data, 31 authors were contacted, 13 answered, and only eight provided their data.

Subgroup analysis by device type

Lastly, some breakdown of the bioimpedance group into its three components (classic impedance, bioreactance, and velocimetry) was done (see Appendix 3) but did not demonstrate any further differences.

Risk of bias in individual studies

Appendix 4 contains the quality assessment analysis for the included studies (QUADAS-2). In eight studies, flow and timing characteristics presented a high risk of bias.

Synthesis of results

Overall meta-analysis

Bias, SD and LOA of the bias, and PE for all included studies are shown in Table 1. The overall random-effects pooled bias (limits of agreement) and the PE were −0.13 (2.23) litres min^{-1} and 47%, respectively. Significant inter-study heterogeneity was detected for bias ($P<0.001$, $I^2=83\%$), see Fig. 2. Symmetrical aspect of the funnel plot (Appendix 5) of mean bias against standard error was confirmed by the Egger test ($P=0.13$).

Additional analysis

Sensitivity and subgroup analyses

The results of subgroup analyses are shown in Figs 3–5. All the subgroup analyses revealed large heterogeneity.

Figure 3 demonstrates the results of the subgroup analysis by device type. The pooled mean bias, in litres per minute, [95% LoA], and PE (%) were as follows: 0.31 litres min^{-1}, [−2.45, −3.07], and 62% for PWTT; −0.20 litres min^{-1}, [−2.32 −1.95], and 45% for niPCA; −0.20 litres min^{-1}, [−2.40 −2], and 40% for CO_2r; and −0.22 litres min^{-1} [−2.43 −1.99], and 42% for TEB. Significant inter-study heterogeneity was detected for all tested technologies, as follows: $I^2=90.6\%$ ($P<0.001$) for PWTT; $I^2=82.6\%$ ($P<0.001$) for niPCA; $I^2=8.7\%$ ($P<0.001$) for CO_2r; and $I^2=79.2\%$ ($P<0.001$) for TEB.

Meta-regression

Declared conflict of interest ($P=0.026$) and publication year ($P=0.001$) were significantly associated with mean bias (Appendix 6). However, no significant heterogeneity remained after accounting for those factors.

Number of studies presenting a trend analysis

Only 10 studies assessed CO trending, and there was very limited information to perform any additional meta-analysis.

Discussion

Summary of evidence

The main results of this meta-analysis assessing the agreement and precision between non-invasive CO monitoring devices and bolus TD monitoring are as follows: (i) the estimate pooled bias [95% CI] were −0,13 [−2.38 , 2.12] litres min^{-1}; (ii) the estimate pooled LOA is ±2.23 litres min^{-1}; (iii) the estimate pooled PE is 47%, well above the threshold of 30% recommended by Critchley and Critchley;[4] and (iv) neither device nor technology was interchangeable with bolus TD. The high heterogeneity for bias and SD was not explained by factors tested in the subgroup analysis or by continuous measurements in the meta-regression analysis.

We conducted this meta-analysis to determine whether, in 2016, the accuracy and precision of commercially available completely non-invasive CO technologies had progressed, compared with what was originally proposed in 2010 by Peyton and Chong.[5] The PE for TEB was 37% in 1999,[4] 43% in 2010,[5] and 42% in the present meta-analysis. Also, CO_2r PE was similar to 2010

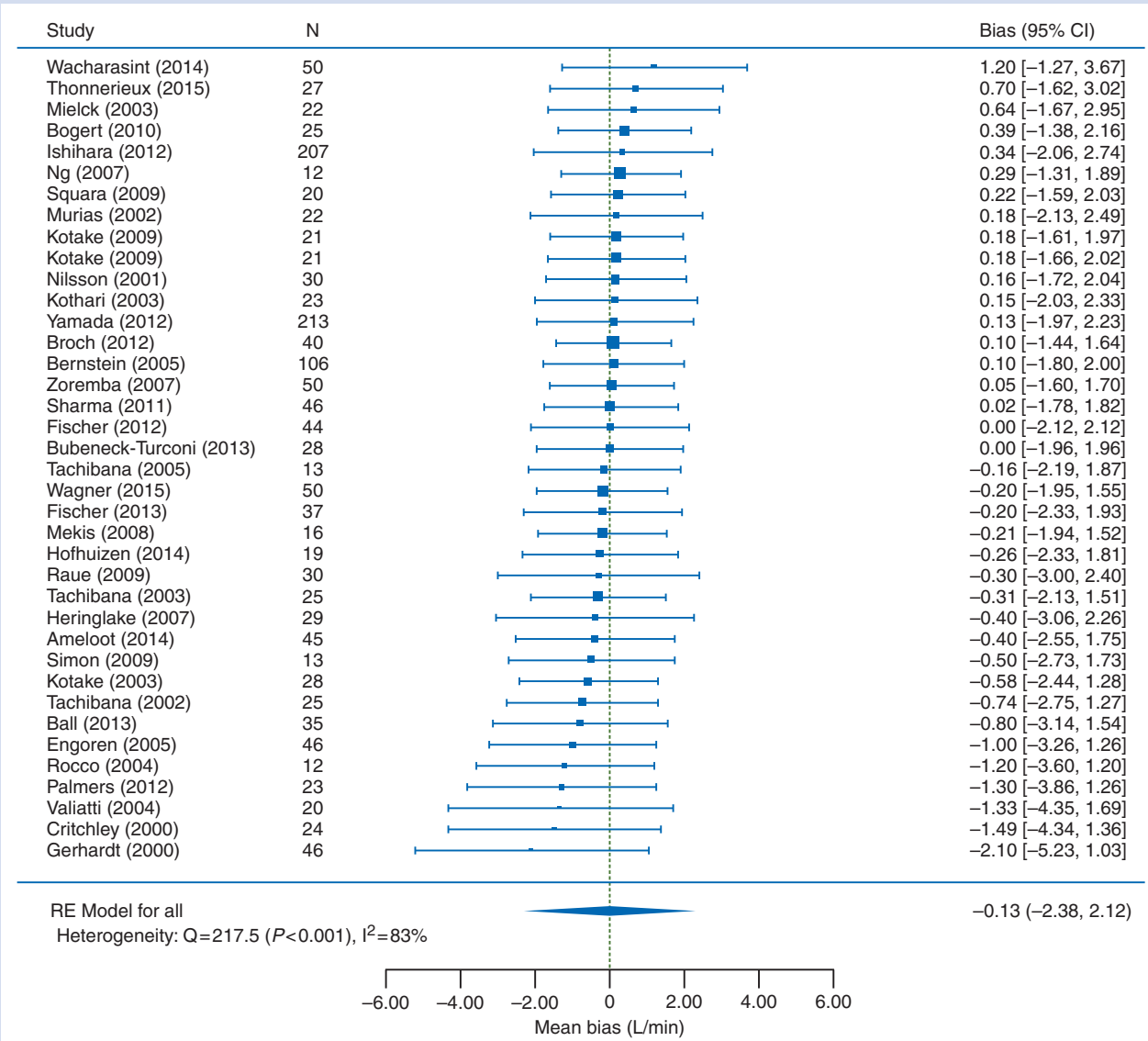

Fig 2 Forest plot showing bias (boxes), 95% limit of agreement (bars), and percentage error for all included studies. The overall random-effects pooled bias and percentage error were −0.13 (2.23) litres min^{-1} and 47%, respectively. Heterogeneity was assessed by I^2 for bias of included studies. Significant inter-study heterogeneity was detected for bias (I^2 = 83.3%, P<0.001).

(44.5%[5] against 40% in this meta-analysis). Therefore, despite advancement of medical technology since the early 2000s, both TEB and CO_2r did not significantly increase their agreement when compared with bolus TD. Additional CO monitoring technologies, such as the PWTT and the niPCA, have also emerged recently. However, with a PE of 45% for niPCA and 62% for PWTT, these newly developed technologies do not bring any consistent reliability. Finally, the PE of the four different non-invasive technologies are all above the recommended threshold of 30%. Moreover, niPCA devices in our meta-analysis (PE of 45%) are as precise as minimally invasive PCA CO devices.[5 51] A recent meta-analysis showed that PE was 44% for the PRAM™ technology and 47% for the third generation of Flo track™. As the PE was similar between minimally invasive and non-invasive devices, the totally non-invasive CO monitors could be an interesting substitute at the bedside. However, further Phase 3 studies (clinical utility/outcome) are clearly mandatory before recommending their widespread use in routine care.[52] Of note, we excluded Doppler-based devices, such as the USCOM and oesophageal Doppler. This was done because the USCOM is not a continuous CO monitoring device and the oesophageal Doppler is relatively intrusive. However, it should be noted that a recent meta-analysis found an equivalent PE for USCOM (43%) when compared with thermodilution.[53]

A major limitation of this study, however, is the population selection bias. As we deliberately chose bolus TD as the reference method, this resulted (mostly for ethical reasons) in the vast majority of studies taking place in a cardio-surgical environment

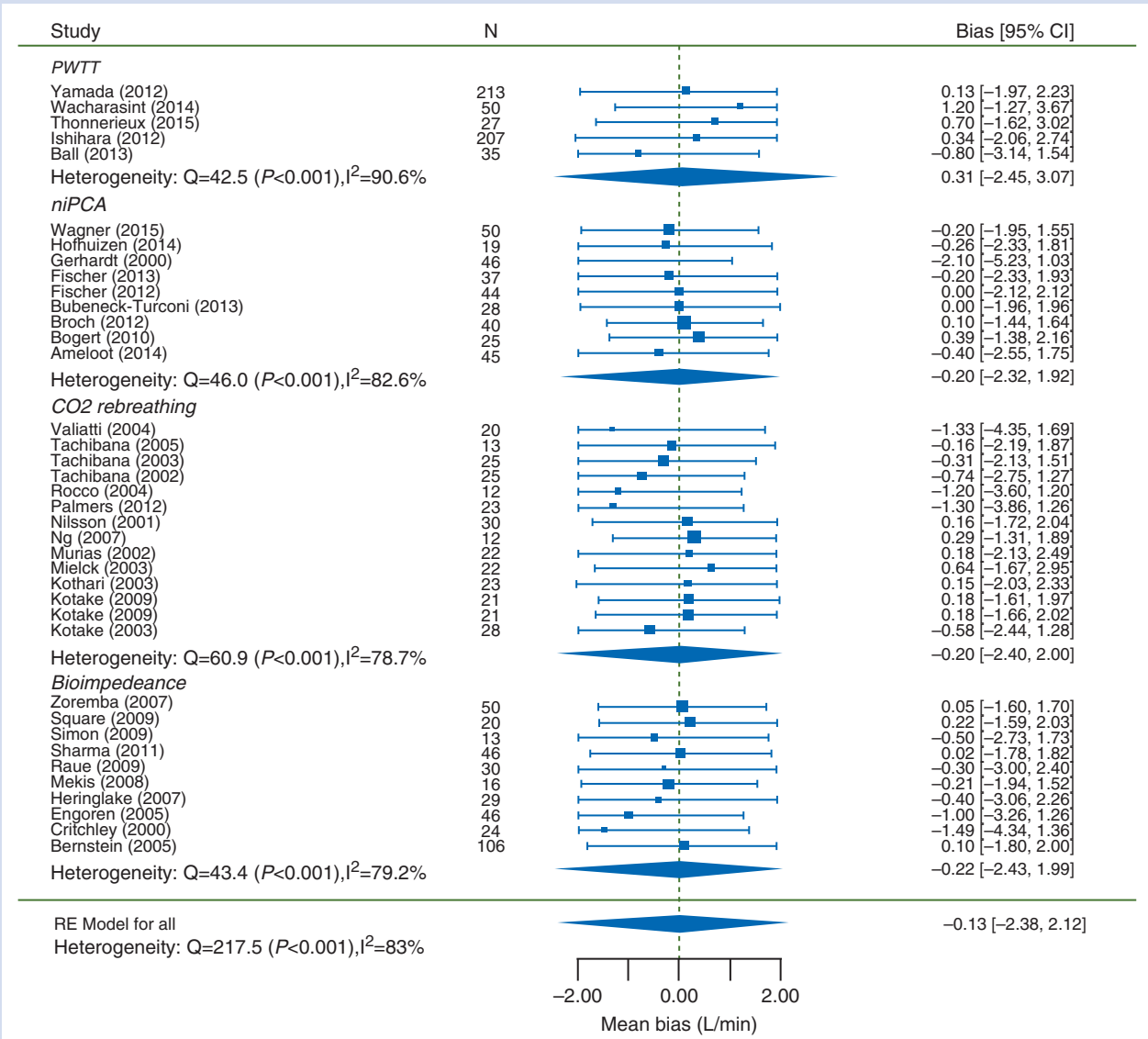

Fig 3 Forest plot showing the results of subgroup analyses by device types.

or during episodes of haemodynamic instability in critically ill patients. This selection bias has two major consequences. Firstly, the studied population is quite disparate from those whom the non-invasive CO device would most benefit.[54] Certainly, patients undergoing moderate- to high-risk general surgical procedures would have the strongest potential to benefit from these technologies. Importantly, the clear discrepancy between the heterogeneity of this high-risk population (small sample size, with various comorbidities and geographical origin) and the homogeneity of the relatively healthy population used to determine the internal algorithms for calculating non-invasive CO might have biased the validity of the initial CO calibration, which thus increases the potential for disagreement when compared with bolus TD. Secondly, peripheral hypoperfusion commonly seen during proceduress with high amounts of haemodynamic instability or in perioperative cardio-surgical care can decrease the sensitivity of peripheral sensors, specifically via decreasing pulse wave amplitude with niPCA and PWTT. Final CO calculations can therefore become erroneous. This might explain why the PE of niPCA and PWTT are greater than that of CO_2r and TEB, as both of the latter technologies do not require peripheral pulse wave measurements to determine CO.

We established our main objective as the PE in order to assess the degree of agreement between the tested and reference method. This variable depends on the LOA (numerator) and the mean CO of both methods (denominator), although no consensus concerning the mean CO (average of both methods or mean of the reference CO) has been clearly defined. As a result of this definition, PE represents the LOA adjusted to the mean CO and therefore represents the random error between both methods. More

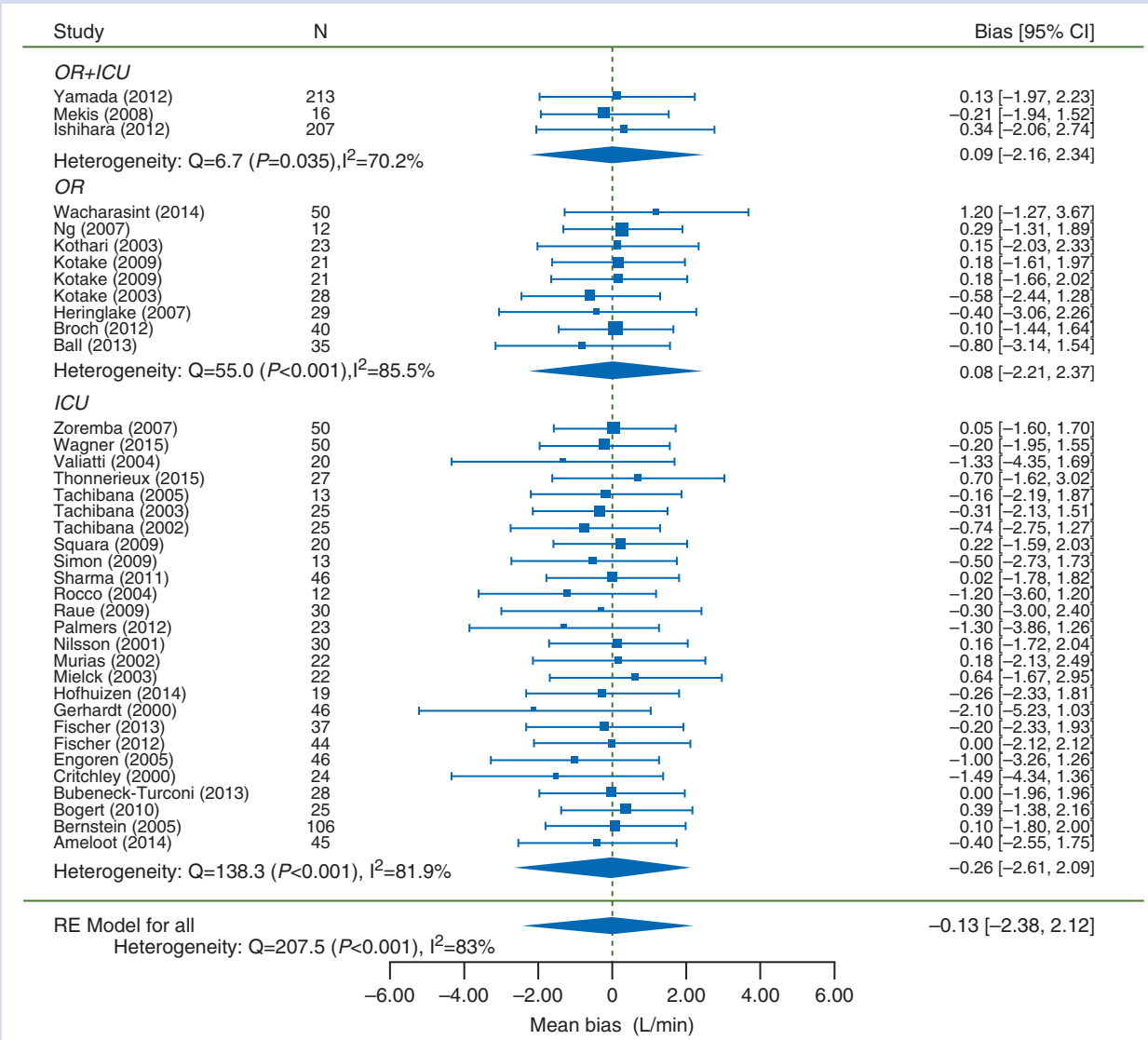

Fig 4 Forest plot showing the results of subgroup analyses by setting (OR, ICU, or OR+ICU). ICU, intensive care unit; OR, operating room.

specifically, it depicts the intrinsic variations of CO that are not linked to true changes of CO, but linked to the environment and the random precision error of both the tested and reference methods. On the contrary, the mean bias depends on the systematic error between both measurements (i.e. the mean constant difference between both methods). When confirming the validity of this approach, one must ensure that the repeatability of bolus TD in the studies is high (precision error <20%). If precision error of the reference method is high, PE will become worthless and acceptability between both techniques will be rejected. This is why Peyton and Chong[5] proposed an increase in the threshold of PE to 45%, allowing for the increased variability noted in bolus TD measurements. All authors in this meta-analysis attempted to reduce the high variability of bolus TD by excluding extreme CO values and recording the average of the three most similar CO measurements. Unfortunately, no observational data presented the dispersion of all TD CO measurements, and only one study assessed the reproducibility of the bolus TD.[42] It could therefore be argued that we should possibly have chosen a different reference technique. However, bolus TD has been considered as the 'gold standard' in other recent meta-analyses.[5 51] Additionally, we do acknowledge that this method is discontinuous, requires manual intervention (therefore increasing the variability of the measure) and is significantly influenced by mechanical ventilation,[55] a caveat which concerns a large proportion of patients in this analysis. Using a reliable continuous CO reference method allows for appropriate interpretation of the PE and therefore should be mandatory for every newly proposed device. In our opinion, ultrasound transit time (USTT) is the most appropriate device tested in *in vitro* studies.[56] It is worth mentioning that when USTT is used as a reference against bolus TD, CO_2r, TEB, or a combination of these, the PE is similar between the tested and reference devices (42%),[57]

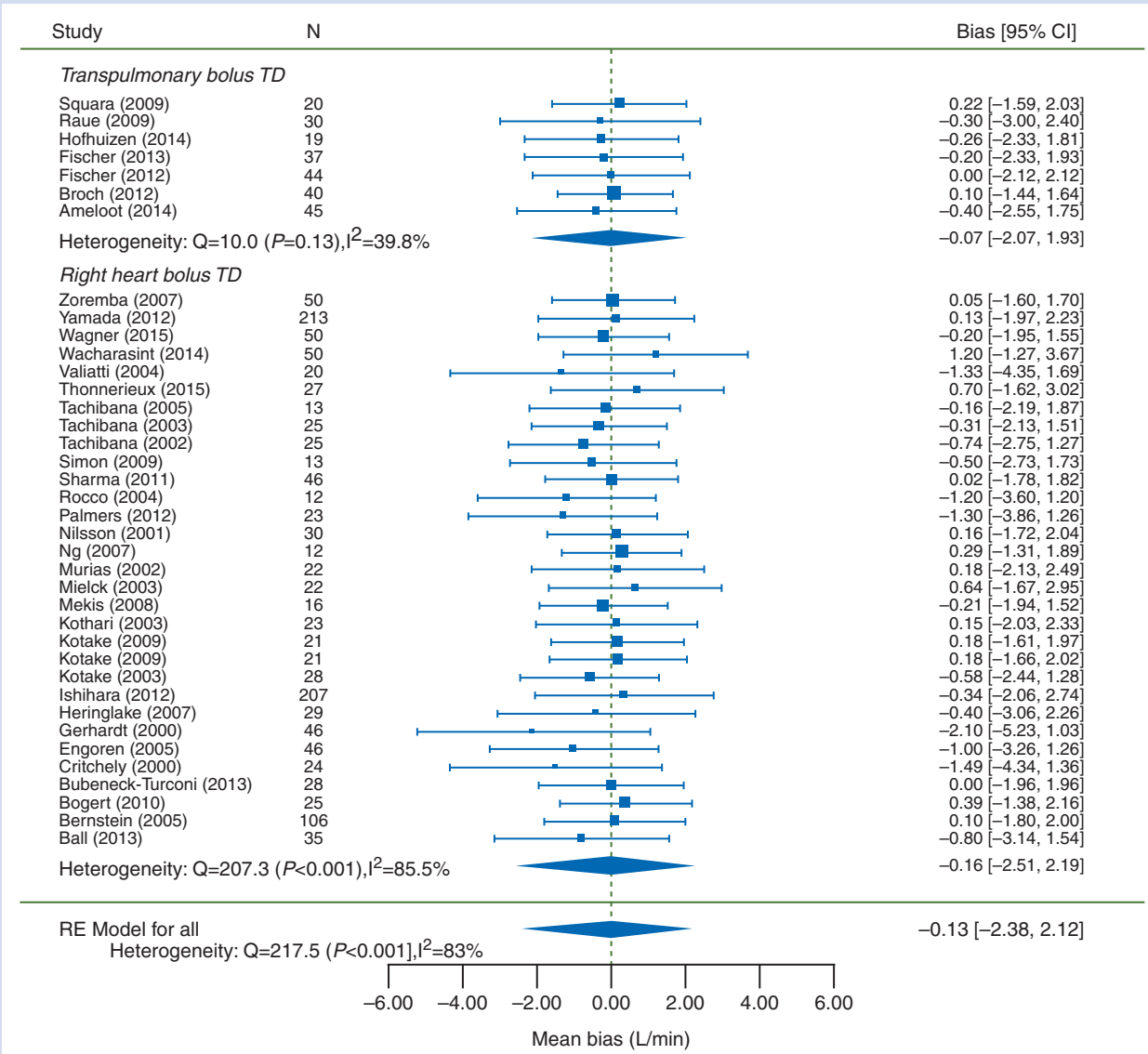

Fig 5 Forest plot showing the results of subgroup analyses by teference method (transpulmonary bolus TD and right heart bolus TD). TD: thermodilution.

further confirming the need for additional experimental data before widespread human use can be endorsed.

Only six studies have previously assessed the repeatability of these devices. [13 15 29 35 42 44] When no repeatability assessment has been conducted on both methods, it is impossible to know whether the disagreement between both techniques is attributable to the tested or the reference method. In the context of this lack of adequate data reporting, GRRAS recommendations insist on the reliability and the repeatability reports when a method-comparison study is designed.[58] The limit of describing repeatability is that there are many ways to express it with similar names (reliability, repeatability, or precision error), which can confuse the reader. Similar to the concept of PE (which is now widely used), standardization for the evaluation of repeatability would allow for better spread of this important concept.[59]

No included study demonstrated that the bias of any device was normally distributed, although recommended by Bland and Altman.[60 61] The method of pooling data in the meta-analysis is also based on a paramount distribution of the differences, further requiring normality.[48] Of note, three studies randomly chosen in this meta-analysis show a non-normally distributed bias. Bland and Altman[60] proposed using a logarithmic transformation of original data when distribution is found to not be normal, especially when the bias is proportional to the mean difference of both methods. Interestingly, no study included here used this approach. Heterogeneity of the results is highly significant. Unfortunately, it is not explained by subgroup sensitivity analysis (tested device, reference device, or setting) or by meta-regression analysis (number of measurements and patients, value of the bias, LOA, or PE). Unknown factors, such as a non-normal distributed bias, are potentially present and might bias our results. The

methodological quality of the studies can therefore be a major factor increasing heterogeneity. Once again, sharing individual data could restrict this phenomenon.

Conclusions

We determined that the overall random-effects pooled bias [95% CI] and percentage error were −0,13 [−2.38 , 2.12] litres min^{-1} and 47%, respectively. Completely non-invasive technologies did not reach an acceptable level of agreement, although minimally invasive technologies present a similar PE. The persistent high heterogeneity after subgroup analysis and meta-regression could be a result of insufficient data reporting and lack of standardization. More rigorous methodology and presentation of method-comparison studies could improve the assessment of consistency and allow physicians to decide better whether non-invasive devices are clinically reliable.

Authors' contributions

Designed the meta-analysis: A.J., O.D., K.S., M.C.
Conducted the search strategy in the different databases: L.S.-K.M.
Contacted authors and several manufacturers of devices for potential additional data: M.E.
Statistics: D.M.B.
Analysed the data: A.J., O.D., K.S., B.A., M.-O.F., L.B., L.V.O., M.C.
Wrote the manuscript: A.J., O.D., B.A., M.-O.F., L.B., L.V.O., M.C.
All authors read and approved the final manuscript.

Supplementary material

Supplementary material is available at *British Journal of Anaesthesia* online.

Acknowledgements

The authors thank Drs Berstein, Broch, Imanaka, Monnet, Murias, Tachibana, Wacharasin, and Palmers for the supply of data shown in their studies.

Declaration of interest

M.C. is a consultant for Edwards Lifesciences (Irvine, CA, USA), Covidien (Boulder, CO, USA), Masimo Corp. (Irvine, CA, USA), ConMed (Irvine, CA, USA), Philips Medical System (Suresnes, France), and Fresenius Kabi (Sevres, France). A.J. is a consultant for Edwards Lifesciences (Irvine, CA, USA). The other authors declare that they have no conflicts of interest concerning this article.

Funding

Fukuda Foundation for Medical Technology (Tokyo, Japan; to K.S.).

References

1. Cannesson M, Pestel G, Ricks C, Hoeft A, Perel A. Hemodynamic monitoring and management in patients undergoing high risk surgery: a survey among North American and European anesthesiologists. *Crit Care* 2011; **15**: R197
2. Thiele RH, Bartels K, Gan TJ. Cardiac output monitoring: a contemporary assessment and review. *Crit Care Med* 2015; **43**: 177–85
3. Squara P, Cecconi M, Rhodes A, Singer M, Chiche JD. Tracking changes in cardiac output: methodological considerations for the validation of monitoring devices. *Intensive Care Med* 2009; **35**: 1801–8
4. Critchley LA, Critchley JA. A meta-analysis of studies using bias and precision statistics to compare cardiac output measurement techniques. *J Clin Monit Comput* 1999; **15**: 85–91
5. Peyton PJ, Chong SW. Minimally invasive measurement of cardiac output during surgery and critical care: a meta-analysis of accuracy and precision. *Anesthesiology* 2010; **113**: 1220–35
6. Saugel B, Cecconi M, Wagner JY, Reuter DA. Noninvasive continuous cardiac output monitoring in perioperative and intensive care medicine. *Br J Anaesth* 2015; **114**: 562–75
7. Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Med* 2009; **6**: e1000100
8. Ameloot K, Van De Vijver K, Broch O, *et al*. Nexfin noninvasive continuous hemodynamic monitoring: validation against continuous pulse contour and intermittent transpulmonary thermodilution derived cardiac output in critically ill patients. *ScientificWorldJournal* 2013; **2013**: 519080
9. Ball TR, Tricinella AP, Kimbrough BA, *et al*. Accuracy of non-invasive estimated continuous cardiac output (esCCO) compared to thermodilution cardiac output: a pilot study in cardiac patients. *J Cardiothorac Vasc Anesth* 2013; **27**: 1128–32
10. Bernstein DP, Lemmens HJ. Stroke volume equation for impedance cardiography. *Med Biol Eng Comput* 2005; **43**: 443–50
11. Bogert LW, Wesseling KH, Schraa O, *et al*. Pulse contour cardiac output derived from non-invasive arterial pressure in cardiovascular disease. *Anaesthesia* 2010; **65**: 1119–25
12. Broch O, Renner J, Gruenewald M, *et al*. A comparison of the Nexfin® and transcardiopulmonary thermodilution to estimate cardiac output during coronary artery surgery. *Anaesthesia* 2012; **67**: 377–83
13. Bubenek-Turconi SI, Craciun M, Miclea I, Perel A. Noninvasive continuous cardiac output by the Nexfin before and after preload-modifying maneuvers: a comparison with intermittent thermodilution cardiac output. *Anesth Analg* 2013; **117**: 366–72
14. Critchley LA, Calcroft RM, Tan PY, Kew J, Critchley JA. The effect of lung injury and excessive lung fluid, on impedance cardiac output measurements, in the critically ill. *Intensive Care Med* 2000; **26**: 679–85
15. Engoren M, Barbee D. Comparison of cardiac output determined by bioimpedance, thermodilution, and the Fick method. *Am J Crit Care* 2005; **14**: 40–5
16. Fischer MO, Avram R, Carjaliu I, *et al*. Non-invasive continuous arterial pressure and cardiac index monitoring with Nexfin after cardiac surgery. *Br J Anaesth* 2012; **109**: 514–21
17. Fischer MO, Coucoravas J, Truong J, *et al*. Assessment of changes in cardiac index and fluid responsiveness: a comparison of Nexfin and transpulmonary thermodilution. *Acta Anaesthesiol Scand* 2013; **57**: 704–12
18. Gerhardt UM, Schöller C, Böcker D, Hohage H. Non-invasive estimation of cardiac output in critical care patients. *J Clin Monit Comput* 2000; **16**: 263–8
19. Heringlake M, Handke U, Hanke T, *et al*. Lack of agreement between thermodilution and electrical velocimetry cardiac output measurements. *Intensive Care Med* 2007; **33**: 2168–72

20. Hofhuizen C, Lansdorp B, van der Hoeven JG, Scheffer GJ, Lemson J. Validation of noninvasive pulse contour cardiac output using finger arterial pressure in cardiac surgery patients requiring fluid therapy. *J Crit Care* 2014; **29**: 161–5
21. Ishihara H, Sugo Y, Tsutsui M, *et al*. The ability of a new continuous cardiac output monitor to measure trends in cardiac output following implementation of a patient information calibration and an automated exclusion algorithm. *J Clin Monit Comput* 2012; **26**: 465–71
22. Kotake Y, Moriyama K, Innami Y, *et al*. Performance of noninvasive partial CO_2 rebreathing cardiac output and continuous thermodilution cardiac output in patients undergoing aortic reconstruction surgery. *Anesthesiology* 2003; **99**: 283–8
23. Kotake Y, Yamada T, Nagata H, *et al*. Improved accuracy of cardiac output estimation by the partial CO_2 rebreathing method. *J Clin Monit Comput* 2009; **23**: 149–55
24. Kothari NAT, Hegde A, Mandke A, Mandke M. Measurement of cardiac output: comparison of four different methods. *Ind J Thorac Cardiovasc Surg* 2003; **19**: 163–8
25. Mekis D, Kamenik M, Starc V, Jeretin S. Cardiac output measurements with electrical velocimetry in patients undergoing CABG surgery: a comparison with intermittent thermodilution. *Eur J Anaesthesiol* 2008; **25**: 237–42
26. Mielck F, Buhre W, Hanekop G, Tirilomis T, Hilgers R, Sonntag H. Comparison of continuous cardiac output measurements in patients after cardiac surgery. *J Cardiothorac Vasc Anesth* 2003; **17**: 211–6
27. Murias GE, Villagra A, Vatua S, *et al*. Evaluation of a noninvasive method for cardiac output measurement in critical care patients. *Intensive Care Med* 2002; **28**: 1470–4
28. Ng JM, Chow MY, Ip-Yam PC, Goh MH, Agasthian T. Evaluation of partial carbon dioxide rebreathing cardiac output measurement during thoracic surgery. *J Cardiothorac Vasc Anesth* 2007; **21**: 655–8
29. Nilsson LB, Eldrup N, Berthelsen PG. Lack of agreement between thermodilution and carbon dioxide-rebreathing cardiac output. *Acta Anaesthesiol Scand* 2001; **45**: 680–5
30. Palmers PJ, Vidts W, Ameloot K, *et al*. Assessment of three minimally invasive continuous cardiac output measurement methods in critically ill patients and a review of the literature. *Anaesthesiol Intensive Ther* 2012; **44**: 188–99
31. Raue W, Swierzy M, Koplin G, Schwenk W. Comparison of electrical velocimetry and transthoracic thermodilution technique for cardiac output assessment in critically ill patients. *Eur J Anaesthesiol* 2009; **26**: 1067–71
32. Rocco M, Spadetta G, Morelli A, *et al*. A comparative evaluation of thermodilution and partial CO_2 rebreathing techniques for cardiac output assessment in critically ill patients during assisted ventilation. *Intensive Care Med* 2004; **30**: 82–7
33. Sharma V, Singh A, Kansara B, Karlekar A. Comparison of transthoracic electrical bioimpedance cardiac output measurement with thermodilution method in post coronary artery bypass graft patients. *Ann Card Anaesth* 2011; **14**: 104–10
34. Simon R, Desebbe O, Henaine R, Bastien O, Lehot JJ, Cannesson M. [Comparison of ICG thoracic bioimpedance cardiac output monitoring system in patients undergoing cardiac surgery with pulmonary artery cardiac output measurements]. *Ann Fr Anesth Reanim* 2009; **28**: 537–41
35. Squara P, Rotcajg D, Denjean D, Estagnasie P, Brusset A. Comparison of monitoring performance of Bioreactance vs. pulse contour during lung recruitment maneuvers. *Crit Care* 2009; **13**: R125
36. Tachibana K, Imanaka H, Miyano H, Takeuchi M, Kumon K, Nishimura M. Effect of ventilatory settings on accuracy of cardiac output measurement using partial CO_2 rebreathing. *Anesthesiology* 2002; **96**: 96–102
37. Tachibana K, Imanaka H, Takeuchi M, Takauchi Y, Miyano H, Nishimura M. Noninvasive cardiac output measurement using partial carbon dioxide rebreathing is less accurate at settings of reduced minute ventilation and when spontaneous breathing is present. *Anesthesiology* 2003; **98**: 830–7
38. Tachibana K, Imanaka H, Takeuchi M, Nishida T, Takauchi Y, Nishimura M. Effects of reduced rebreathing time, in spontaneously breathing patients, on respiratory effort and accuracy in cardiac output measurement when using a partial carbon dioxide rebreathing technique: a prospective observational study. *Crit Care* 2005; **9**: R569–74
39. Thonnerieux M, Alexander B, Binet C, Obadia JF, Bastien O, Desebbe O. The Aability of esCCO and ECOM monitors to measure trends in cardiac output during alveolar recruitment maneuver after cardiac surgery: a comparison with the pulmonary thermodilution method. *Anesth Analg* 2015; **121**: 383–91
40. Valiatti JL, Amaral JL. Comparison between cardiac output values measured by thermodilution and partial carbon dioxide rebreathing in patients with acute lung injury. *Sao Paulo Med J* 2004; **122**: 233–8
41. Wacharasint P, Kunakorn P, Pankongsap P, Preechanukul R. Clinical validation of pulse contour and pulse wave transit time-based continuous cardiac output analyses in Thai patients undergoing cardiac surgery. *J Med Assoc Thai* 2014; **97 Suppl 1**: S55
42. Wagner JY, Sarwari H, Schon G, *et al*. Radial artery applanation tonometry for continuous noninvasive cardiac output measurement: a comparison with intermittent pulmonary artery thermodilution in patients after cardiothoracic surgery. *Crit Care Med* 2015; **43**: 1423–8
43. Yamada T, Tsutsui M, Sugo Y, *et al*. Multicenter study verifying a method of noninvasive continuous cardiac output measurement using pulse wave transit time: a comparison with intermittent bolus thermodilution cardiac output. *Anesth Analg* 2012; **115**: 82–7
44. Zoremba N, Bickenbach J, Krauss B, Rossaint R, Kuhlen R, Schälte G. Comparison of electrical velocimetry and thermodilution techniques for the measurement of cardiac output. *Acta Anaesthesiol Scand* 2007; **51**: 1314–9
45. Kim SH, Lilot M, Murphy LS, *et al*. Accuracy of continuous noninvasive hemoglobin monitoring: a systematic review and meta-analysis. *Anesth Analg* 2014; **119**: 332–46
46. Kim SH, Lilot M, Sidhu KS, *et al*. Accuracy and precision of continuous noninvasive arterial pressure monitoring compared with invasive arterial pressure: a systematic review and meta-analysis. *Anesthesiology* 2014; **120**: 1080–97
47. Suehiro K, Joosten A, Murphy LS, *et al*. Accuracy and precision of minimally-invasive cardiac output monitoring in children: a systematic review and meta-analysis. *J Clin Monit Comput* 2015; **30**: 603–20
48. Williamson PR, Lancaster GA, Craig JV, Smyth RL. Meta-analysis of method comparison studies. *Stat Med* 2002; **21**: 2013–25
49. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Br Med J* 2003; **327**: 557–60
50. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *Br Med J* 1997; **315**: 629–34

51. Schlöglhofer T, Gilly H, Schima H. Semi-invasive measurement of cardiac output based on pulse contour: a review and analysis. *Can J Anaesth* 2014; **61**: 452–79
52. Fischer MO, Lorne E. Perioperative non-invasive haemodynamic monitoring: yes or not yet? *Anaesth Crit Care Pain Med* 2016; **35**: 423–4
53. Chong SW, Peyton PJ. A meta-analysis of the accuracy and precision of the ultrasonic cardiac output monitor (USCOM). *Anaesthesia* 2012; **67**: 1266–71
54. Cannesson M, Manach YL. Noninvasive hemodynamic monitoring: no high heels on the farm; no clogs to the opera. *Anesthesiology* 2012; **117**: 937–9
55. Reuter DA, Huang C, Edrich T, Shernan SK, Eltzschig HK. Cardiac output monitoring using indicator-dilution techniques: basics, limits, and perspectives. *Anesth Analg* 2010; **110**: 799–811
56. Yang XX, Critchley LA, Joynt GM. Determination of the precision error of the pulmonary artery thermodilution catheter using an in vitro continuous flow test rig. *Anesth Analg* 2011; **112**: 70–7
57. Botero M, Kirby D, Lobato EB, Staples ED, Gravenstein N. Measurement of cardiac output before and after cardiopulmonary bypass: comparison among aortic transit-time ultrasound, thermodilution, and noninvasive partial CO_2 rebreathing. *J Cardiothorac Vasc Anesth* 2004; **18**: 563–72
58. Kottner J, Audigé L, Brorson S, *et al.* Guidelines for Reporting Reliability and Agreement Studies (GRRAS) were proposed. *J Clin Epidemiol* 2011; **64**: 96–106
59. Desebbe O, Joosten A, Suehiro K. Next generation of method-comparison studies: standardization of data presentation and clinical application are a goal to reach. *Crit Care Med* 2015; **43**: e468–9
60. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; **1**: 307–10
61. Hanneman SK. Design, analysis, and interpretation of method-comparison studies. *AACN Adv Crit Care* 2008; **19**: 223–34

Handling editor: Paul Myles