Physiological Measurement

IPEM
Institute of Physics and Engineering in Medicine

ACCEPTED MANUSCRIPT

Comparison of impedance cardiography and cardiac magnetic resonance imaging for the evaluation of cardiac function in early-stage breast cancer patients

To cite this article before publication: Erifyli Piastopoulou *et al* 2021 *Physiol. Meas.* in press https://doi.org/10.1088/1361-6579/ac28e5

Manuscript version: Accepted Manuscript

Accepted Manuscript is "the version of the article accepted for publication including all changes made as a result of the peer review process, and which may also include the addition to the article by IOP Publishing of a header, an article ID, a cover sheet and/or an 'Accepted Manuscript' watermark, but excluding any other editing, typesetting or other changes made by IOP Publishing and/or its licensors"

This Accepted Manuscript is **© 2021 Institute of Physics and Engineering in Medicine**.

During the embargo period (the 12 month period from the publication of the Version of Record of this article), the Accepted Manuscript is fully protected by copyright and cannot be reused or reposted elsewhere.
As the Version of Record of this article is going to be / has been published on a subscription basis, this Accepted Manuscript is available for reuse under a CC BY-NC-ND 3.0 licence after the 12 month embargo period.

After the embargo period, everyone is permitted to use copy and redistribute this article for non-commercial purposes only, provided that they adhere to all the terms of the licence https://creativecommons.org/licences/by-nc-nd/3.0

Although reasonable endeavours have been taken to obtain all necessary permissions from third parties to include their copyrighted content within this article, their full citation and copyright line may not be present in this Accepted Manuscript version. Before using any content from this article, please refer to the Version of Record on IOPscience once published for full citation and copyright details, as permissions will likely be required. All third party content is fully copyright protected, unless specifically stated otherwise in the figure caption in the Version of Record.

View the article online for updates and enhancements.

This content was downloaded from IP address 90.254.65.216 on 28/09/2021 at 09:41

IOP Publishing Journal Title
Journal **XX** (XXXX) XXXXXX https://doi.org/XXXX/XXXX

Comparison of impedance cardiography and cardiac magnetic resonance imaging for the evaluation of cardiac function in early-stage breast cancer patients

Erifyli Piastopoulou[1], Parvaiz Ali[1], Gianfilippo Bertelli[3], Martyn Heatley[2], Maung Moe[4], Chandramohan Murugesan[1], Gareth Stratton[1] and Michael Lewis[1]

[1] ASTEM Research Centre, School of Sport and Exercise Sciences, College of Engineering, Swansea University, UK
[2] Singleton Hospital, Swansea Bay University Health Board, UK
[3] Sussex Cancer Centre, Brighton and Sussex University Hospitals NHS Trust, Brighton, UK
[4] Southampton General Hospital, University Hospital Southampton NHS Foundation Trust, UK

E-mail: e.piastopoulou@hotmail.com

Received xxxxxx
Accepted for publication xxxxxx
Published xxxxxx

Abstract

Objective: Breast cancer treatment can negatively impact cardiac function in some breast cancer patients. Current methods (MUGA, echocardiography) used in clinical practice to detect abnormal cardiac changes as a result of treatment suffer from important limitations. Use of alternative techniques that would offer safe, inexpensive and non-invasive cardiac function assessment in this population would be highly advantageous. The aim of this study was to examine the agreement between impedance cardiography (ICG) and cardiac magnetic resonance imaging (CMR) in quantifying stroke volume (SV), cardiac output (CO) and end-diastolic volume (EDV) in this population. *Approach:* Sixteen breast cancer patients underwent ICG and CMR assessments at three time-points (before treatment, immediately after chemotherapy, and four months after chemotherapy). Bland-Altman analysis was used to quantify the accuracy and precision of ICG (relative to CMR) in estimating absolute values of SV, CO and EDV. Four methods (concordance rate, polar concordance rate, clinical concordance rate and trend interchangeability rate) were also used to assess ICG performance in tracking changes in these variables. *Main results:* Bland-Altman analysis showed that the accuracy of ICG relative to CMR was -3.1 ml (SV), 0.2 $L \cdot min^{-1}$ (CO) and -26.0 ml (EDV) and precision was 13.2 ml (SV), 1.1 $L \cdot min^{-1}$ (CO) and 20.1 ml (EDV), respectively. Trending ability assessment showed that 1) the concordance rate was 87% (SV), 73% (CO) and 73% (EDV), 2) the polar concordance rate was 67% (SV), 53% (CO) and 33% (EDV), 3) the clinical concordance rate was 33% (SV), 40% (CO) and 20% (EDV) and 4) the trend interchangeability rate was 29% (SV), 43% (CO) and 17% (EDV), respectively. *Significance:* Our findings show that, although ICG showed good accuracy for absolute SV and CO measurements and for CO and EDV changes, precision was poor for all variables in terms of both absolute measurements and trend tracking performance. This suggests that ICG cannot be used interchangeably with CMR in breast cancer patients.

Keywords: cardiotoxicity, impedance cardiography, cardiac magnetic resonance imaging, Bland-Altman analysis, trending ability, method-comparison study, breast cancer

Accepted Manuscript

 © xxxx IOP Publishing Ltd

1. Introduction

Latest figures show that 80% of women diagnosed with breast cancer now survive for ten or more years (Cancer Research UK, 2020), reflecting substantial improvements in diagnosis and treatment over the last four decades. However, despite this remarkable success, some patients suffer other potentially serious morbidities during and following their anticancer therapy. The most serious of these is cardiotoxicity (Altena et al., 2009; Bovelli et al., 2010), a multi-faceted condition characterised by myocyte damage (Mihalcea et al., 2017). Cardiotoxocity typically progresses from a subclinical form (without signs or symptoms of cardiac deterioration) to a clinical form (in the form of cardiomyopathy or heart failure) following left ventricular remodelling caused by exposure to cancer treatment (Manrique et al., 2017). A reduction in LVEF (left ventricular ejection fraction) to below a specific cut-off value (typically 50%) and/or a relative change in LVEF of a certain magnitude compared to pre-treatment values (typically 10%) is used to denote the development of cardiotoxicity (Chung et al., 2018). Techniques offering the ability to detect the early signs of cardiac deterioration during and following treatment would be extremely valuable in clinical practice.

Cardiac magnetic resonance imaging (CMR) is the gold standard method for assessing left and right ventricular cardiac structure and function, but it is not routinely used in breast cancer patients because of its expense and restricted accessibility (Plana et al., 2014). Echocardiography is the recommended modality for cardiac evaluation in breast cancer (e.g. in the guidelines issued by the Joint Task Force of the American Society of Echocardiography and the European Association of Cardiovascular Imaging (Plana et al., 2014), the Canadian Cardiovascular Society (Virani et al., 2016), the European Society of Cardiology (Zamorano et al., 2016) and the American Society of Clinical Oncology (Armenian et al., 2016)), but it is less accurate than CMR (Bellenger et al., 2000; Armstrong et al., 2012) and its quality is operator-dependent (Zamorano et al., 2016). MUGA (multi-gated acquisition), a nuclear medicine imaging technique (also known as radionuclide ventriculography), is an alternative method that can be used to monitor cardiac function, which can be useful when echocardiographic images are of poor quality (Plana et al., 2014), but it is used less frequently (Kolla et al., 2017) as it exposes patients to ionising radiation (Bloom et al., 2016). These are significant disadvantages because they either prevent the use of multiple cardiac assessments (in the case of MUGA) or limit our confidence in their results (in the case of echocardiography). The use of frequent cardiac assessments in this population would be desirable not only during treatment but also after its completion, as cardiotoxicity can manifest in various forms many years after treatment cessation. An inexpensive, safe and operator-independent method that could be used for frequent assessments of cardiac function would be advantageous in the clinical management of cancer patients: i) it would facilitate improved insight into the dynamic evolution of cardiotoxicity, ii) it would help to identify patients who would benefit from the early administration of cardiac-sparing pharmacological agents or alterations to treatment regimens, and ultimately iii) it would help to reduce the incidence of cardiotoxic morbidities.

One such method might be thoracic electrical bioimpedance (or impedance cardiography, ICG), a technique that non-invasively estimates cardiac volume changes on a beat-to-beat basis (Fortin et al., 2006). ICG has been compared with CMR in paediatric patients (with and without cardiac disease; Taylor et al., 2012), in healthy adult participants (Borzage et al., 2017), in adult patients referred for a CMR scan (Villacorta Junior et al., 2012), in adult and paediatric patients with congenital heart disease (Ebrahim et al., 2016) and in adult patients suspected of suffering from pulmonary arterial hypertension (Panagiotou et al., 2018). The findings from these studies have been inconsistent, but most authors have shown a poor agreement between the two modalities in terms of two of their commonly derived variables: stroke volume (SV) and cardiac output (CO). Although ICG and CMR can also provide end-diastolic volume (EDV) measurements, no studies to date have examined the agreement between ICG and CMR in estimating this important measure of ventricular pre-load. Furthermore, only a few studies have used ICG to assess cancer patients, with suggestions that ICG might be useful in detecting early signs of cardiac deterioration (Massidda et al., 1997), for guiding treatment decisions relating to the minimization of cardiotoxicity risk (Kasznicki and Drzewoski, 1993), and for exploring exercise performance (Jones et al., 2007). Notably however, none of these studies evaluated the performance of ICG by comparing results with a reference technique.

A growing number of researchers have also emphasized the importance of evaluating the ability of a method to track changes (i.e. its 'trending ability') in addition to comparing absolute values acquired at independent time-points (e.g. using Bland-Altman analysis) (Critchley, 2017; Squara et al., 2009). This is because clinicians use monitoring devices for two purposes (Odor et al., 2017): making an accurate diagnosis at each single examination and guiding therapeutic decisions by detecting significant health changes between examinations. Importantly, even if a method demonstrates reduced accuracy in the measurement of absolute values, it might still prove to be advantageous for the management of patients if it can reliably characterise changes (Critchley et al., 2010).

Accepted Manuscript

The aim of the study was to examine the agreement between ICG and CMR in assessing cardiac function in early-stage breast cancer patients. The specific objectives of the study were: 1) to quantify the accuracy[1] and precision[2] of ICG (relative to CMR) in estimating absolute values of SV, CO and EDV, and 2) to compare ICG and CMR in terms of their abilities to dynamically track changes in these variables (the so-called 'trending ability') during and following breast cancer treatment. An important requirement for achieving the second objective was to also review and compare the statistical techniques that are appropriate for comparing the trending abilities of these two modalities.

2. Methods

2.1 Study design

This was a prospective, longitudinal, observational cohort study with serial measurements of the study variables during and following breast cancer treatment. The study was approved by the Health Research Authority and Health and Care Research Wales (Research Permissions; IRAS project ID 188676) and by the Health and Care Research Wales Research Ethics Service (REC reference 16/WA/0183).

2.2 Participants

Eligible participants were women aged 18 years or over who i) had early invasive breast cancer (stage I-III), ii) were due to start adjuvant or neoadjuvant therapy (either anthracycline-based chemotherapy or chemotherapy + trastuzumab +/- pertuzumab), iii) had not previously received adjuvant or neoadjuvant therapy (either anthracycline-based chemotherapy or chemotherapy + trastuzumab +/- pertuzumab), iv) had WHO performance status <=2, v) had no major cardiovascular problems or chronic respiratory problems, and vi) had no contraindications for MRI scans.

For the purposes of the study reported here we did not distinguish between participants on the basis of the type of treatment they received (either adjuvant/neoadjuvant anthracycline-based chemotherapy only, or adjuvant/neoadjuvant chemotherapy along with anti-HER2 antibody treatment).

2.3 Assessments

Continuous beat-to-beat ICG was recorded using the the Task Force Haemodynamic Monitor (TFM; CNSystems Medizintechnik GMBH, Austria). The TFM uses an electrical thorax model based on Ohm's law to measure both the static thoracic impedance (of blood and tissue) and changes in the impedance caused by cardiac contraction and blood flow in the aorta. Electrodes placed on thorax (xiphoid) and neck are used to deliver a small, high-frequency (0.4 mA, 40 kHz) electrical current and to measure the corresponding electrical impedance, with a reproducibility of 97% (Fortin et al., 2006). Simultaneously, the TFM's 4-lead ECG records the electrocardiogram. Left ventricular stroke volume is then estimated using the product of the maximum derivative of the impedance signal and the left ventricular ejection time (LVET, the time-interval between opening and closure of the aortic valve). Other cardiac variables can then be derived, including left ventricular end-diastolic volume (pre-load), cardiac output, and indices representing cardiac contractility and blood flow during ventricular ejection.

Cardiac Magnetic Resonance Imaging (CMR) was used to quantify cardiac structure (chamber size, myocardial mass, aorta and pulmonary artery dimensions) and function (chamber volume, regional wall motion, valve function). CMR images were acquired using a 3T Siemens Magnetom Skyra MRI scanner (Siemens Healthcare GmbH, Germany). A Steady State Free Precession (SSFP) cine sequence was used to capture long-axis (four-, three- and two-chamber view) and short-axis images of the entire left and right ventricles. ECG-gating and breath-hold imaging techniques were used to alleviate cardiac and lung motion, respectively. The CMR images were then analysed by an experienced cardiologist using the QMass post-processing software (Medis Medical Imaging Systems, Netherlands) to evaluate regional and global ventricular function. For the

[1] In this context, 'accuracy' describes the 'bias' or mean difference between values acquired using two methods, calculated using Bland-Altman analysis

[2] In this context, 'precision' describes the distribution of 'differences' (of the two values) around their mean difference (quantified by the Limits of Agreement in Bland-Altman analysis)

Accepted Manuscript

purposes of this study only the following CMR-derived variables will be discussed: stroke volume (SV), cardiac output (CO) and end-diastolic volume (EDV).

Each participant attended three assessments: first at baseline (T1; 1-30 days prior to the initiation of chemotherapy), then immediately following the completion of chemotherapy (T2; typically 4 months after baseline), and finally after the completion of chemotherapy (T3; typically 8 months after baseline). ICG and ECG surface electrodes were placed on the neck, thorax and abdomen in accordance with manufacturer guidelines. As part of the wider study, each participant was asked to perform a variety of postural manoeuvres and physical exercise to provoke a range of physiological steady states over a 25 minute period. However, only data from the initial 5 minute period of supine rest are considered in this study (to correspond with the supine state during CMR assessment). Participants were asked to refrain from drinking coffee, tea, alcohol or a heavy meal within two hours prior to the assessment and also to refrain from strenuous physical activity within 24 hours prior to assessment. Participants were also asked to drink 250 ml of water 10-15 minutes prior to assessment in order to ensure adequate hydration. All CMR scans were performed immediately before or after the ICG assessment for each participant to minimize any potential confounding variations of physiological status.

2.4 Time-independent agreement between ICG and CMR

The agreement between ICG and CMR in terms of absolute SV, CO and EDV measurements was assessed using the Bland-Altman method (Bland and Altman, 1999), with a few modifications (Ludbrook, 2010). Our study design involved longitudinal participant assessments, as some individuals underwent either two or three ICG and CMR assessments at different time-points during the study. We considered these assessments to be serial measurements and not 'repeated' measurements (Hapfelmeier et al., 2016) since we expected their values to vary as a result of treatment. Data from all participants were included in our cross-sectional Bland-Altman analysis, even if they did not complete all assessments.

The differences between pairs of measurement (using ICG and CMR) were first calculated for each measured variable (i.e. SV, CO, EDV). If the paired-measurement differences followed a normal distribution, a simple scatter plot resembling a Bland-Altman plot (i.e. difference between ICG and CMR on the y-axis and average of ICG and CMR on the x-axis) but without the bias, 95% limits of agreement (LOA) or the corresponding 95% confidence intervals (CIs)) was then constructed to examine if proportional bias (where the magnitude of differences between the two methods depends on the average value of the two methods) and/or heteroscedasticity (where the scatter of the differences between the two methods depends on the average value of the two methods) were present. Proportional bias was examined using a least-squares linear regression analysis, and heteroscedasticity was assessed visually.

The following variables were then calculated: i) Mean Difference (MD) of paired ICG and CMR measurements (i.e. the 'bias', reflecting the accuracy of agreement), and the presence of fixed bias was examined using a one-sample t-test; ii) Standard deviation (SD) of the paired differences; iii) 95% limits of agreement (LOA, reflecting the precision of agreement) of the MD ($LOA = MD \pm t \times SD \times \sqrt{\left[1 + \left(\frac{1}{n}\right)\right]}$), where t is a true Student's t test value corresponding to two-sided p=0.05 for n-1 degrees of freedom and n is the sample size (Lentner, 1982)[3] ; iv) 95% CIs of the bias ($MD \pm t \times SEM$), where SEM is the standard error of the MD, given by √(SD2/n); v) 95% CIs of the LOA ($MD \pm c \times SD$), where c was determined from Carkeet et al. (2015) for 2.5% confidence and 97.5% confidence; vi) $Percentage\ Error\ (PE) = 100 \times t \times \left(\frac{SD}{mean\ of\ the\ two\ methods}\right)$, where 'mean of the two methods' is the mean value between the reference method (CMR) and the test method (ICG) for each variable of interest (as suggested by Critchley and Critchley (1999) for CO measurements, and a PE of up to 30% was adopted as the threshold point indicating good agreement between ICG and CMR).

2.5 Assessment of the 'trending ability' agreement between ICG and CMR

The ability of the ICG to accurately characterise temporal changes in SV, CO and EDV measurements during and following treatment (i.e. its trending ability) was assessed against CMR using four different statistical techniques: 1) the four-quadrant plot and concordance analysis (Perrino et al., 1994; Perrino et al., 1998), 2) the polar plot method (Critchley et al., 2010; Critchley et al., 2011), 3) the clinical concordance rate and clinical concordance plot (Montenij et al., 2016a), and 4) the

[3] This equation provides more accurate results for small sample sizes (i.e. <60) than the commonly used LOA calculation (Lentner, 1982)

Accepted Manuscript

trend interchangeability method (Fisher et al., 2016). In each technique, the trending ability is assessed based on the changes (Δ values) calculated for consecutive measurements of the variable of interest (here, ΔSV, ΔCO and ΔEDV).

2.5.1 Four-quadrant plot and concordance analysis

Construction of the four-quadrant plot and calculation of concordance rate were performed according to the method described by Perrino et al. (1994, 1998). This plot (depicting CMR changes on the x-axis and ICG changes on the y-axis) presents the relationship between sequential changes in CMR-derived and ICG-derived variables. Data points located at the centre of the plot represent very small changes and have typically been considered to reflect random noise (rather than 'true' differences between consecutive measurements) and so excluded from trending analysis (Critchley et al., 2010; Critchley et al., 2011). However, we retained these in our analysis as in clinical practice these small changes would reflect 'stable' cardiac function and thus provide valuable information. The concordance rate was determined by calculating the number of points included in the two concordant quadrants (i.e. top-right and bottom-left quadrants). A concordance rate of >90% was considered to indicate good trending ability of ICG compared to CMR, as suggested by Critchley et al. (2010).

2.5.2 Polar plot method

Construction of the polar plot and quantification of trending ability were performed in accordance with methods previously described (Critchley et al., 2010; Critchley et al., 2011). Cartesian coordinates from the four-quadrant plot were converted to polar coordinates, i.e. polar angle (θ, in radians) and radius (r). The angle of each data point relative to the horizontal axis of the polar plot (i.e. the polar angle) represents the degree of agreement between the two methods (for each individually-measured change in the variable of interest) and the length of the vector from the origin to each point (i.e. the radius) represents the mean magnitude of the change (mean of ΔICG and ΔCMR). Data points on the polar plot representing changes that are in agreement for ICG and CMR have a vector/angle close to the horizontal axis, whereas those not in agreement will be closer to the vertical axis. We again included all data points in the analysis (including very small changes).

Trending analysis was performed by calculating the mean polar angle (i.e. angular bias, the mean polar angle relative to the horizontal axis of the polar plot), the standard deviation of the mean polar angle, the radial 95% LOA, and the polar concordance rate (i.e. the proportion of points included within the ±30° polar angle limit). (The 95% radial LOA were determined by following four steps: i) Converting all negative changes (i.e. polar angles from 90° to 270°) to positive changes using a 180° rotation, ii) calculation of the percentage of points included in different radial limits (i.e. 5°, 10°, 15° etc.), b) construction of a plot showing the percentage of points on the y-axis and the radial limits on the x-axis, and c) determination of the radial limits at which the percentage of points is 95%. If, however, the mean polar angle was < ±5°, the 95% CIs (calculated as $2 \times SD$) were used instead of the radial 95% LOA, as suggested by Critchley et al. (2011).) The threshold points for good trending ability were defined as a mean polar angle of less than ±5°, a SD for the mean polar angle of less than ±15°, a radial 95% LOA (i.e. radial area including 95% of the data points) of less than ±30° and a polar concordance rate of more than 95%, as recommended by Critchley et al. (2010, 2011).

2.5.3 Clinical concordance method

Calculation of the clinical concordance rate and generation of the corresponding plot (i.e. clinical concordance plot with a superimposed error grid) were performed according to the method introduced by Montenij et al. (2016a). The first step was to categorise the CMR changes into one of the following three categories: 1) a non-significant change, where the CMR change (ΔCMR) is equal to or less than ±5% of the initial CMR value, 2) a moderate increase or decrease, where ΔCMR is between ± 5-15% of the initial CMR value, and 3) a large increase or decrease, where ΔCMR is equal to or greater than ±15% of the initial CMR value. Data points representing ICG changes (ΔICG) were then considered to reflect good trending ability when they pertained to the same zone as the corresponding CMR changes (i.e. consistent positive or negative changes of similar magnitude).

Clinical concordance rate was then calculated as the percentage of ICG changes reflecting good trending ability (i.e. percentage of points falling into the three categories of clinical concordance). The threshold point for good clinical concordance rate was set to ≥90% (Montenij et al., 2016b) for all variables of interest. The clinical concordance plot was constructed by superimposing an 'error grid' on a four-quadrant plot. The error grid zones can be interpreted in terms of the consequences of making treatment decisions (in our study, this refers to the decision to either continue breast cancer treatment or to initiate cardioprotective therapy) based on adoption of the 'new' method (here, ICG-based measurements).

Accepted Manuscript

2.5.4 Trend interchangeability method

Calculation of the trend interchangeability rate and generation of the corresponding plot were performed according to the method introduced by Fischer et al. (2016). Each CMR-derived change (in the variable of interest) was first defined as 'interpretable' or 'uninterpretable', based on the repeatability coefficient of CMR. True (interpretable) changes are defined only for serial CMR measurements with non-overlapping confidence intervals (determined as $CI = CMRValue_i \pm (CMRValue_i \times RC)$, where i represents serial measurements, CMRValue is the CMR-measured variable, and RC is the repeatability coefficient for the CMR-measured variable).

Each interpretable change was then classified as interchangeable, non-interchangeable or 'in the grey zone' according to the repeatability coefficient of CMR, which was used to define a precision interval for the first pair (in the temporal sequence) of measurements made by CMR and ICG. Classification of interpretable changes was performed by using the following equation (1):

$$x = y(1 \pm RC) + (1 + RC)(Vref_1 - Vtest_1) (1)$$

where x represents the reference method (which in our case is CMR), y represents the test method (which in our case is ICG), RC is the repeatability coefficient of CMR, and $Vref_1$ and $Vtest_1$ represent the first pair of measurements made by CMR and ICG for each variable of interest. The repeatability coefficient of CMR for SV, CO and EDV measurements was set to 5%. An interpretable change was then classified as 'interchangeable' only when the second pair of measurements obtained by ICG and CMR fell inside the precision interval of the first pair of measurements. An interpretable change was considered to fall into the 'grey zone' when the precision interval of the first pair of measurements did not include the second pair of measurements but only overlapped one of the confidence intervals of the second pair of measurements. Finally, an interpretable change was classified as 'noninterchangeable' when the precision interval of the first pair of measurements did not include either the second pair of measurements or their confidence intervals. Following the classification of all changes, the trend interchangeability rate was determined by dividing the total number of interchangeable changes by the total number of interpretable changes. The threshold value for an acceptable interchangeability rate between ICG and CMR was set a priori to >90%, as recommended by Fischer et al. (2016), for all variables of interest. The spreadsheet provided by Fischer et al. (2016) was used to classify changes into the four categories.

3. Results

3.1 Participants

Sixteen breast cancer patients agreed to take part in the study, but three did not attend the cardiac and haemodynamic assessments due to claustrophobia or time-constraints in relation to treatment. Participants were of mean (SD) age 55 (9) years and mean (SD) BMI 25.9 (4.3) $kg \cdot m^{-2}$ (ranges: 36-68 years, 19.5-36.1 $kg \cdot m^{-2}$). Fifteen participants had invasive ductal carcinoma of grade 1-3: grade 1 (1), grade 2 (6), grade 3 (8). One participant had invasive lobular carcinoma of grade 2. Seven participants were HER2 positive and received anti-HER2 treatment in addition to chemotherapy.

3.2 Time-independent SV, CO and EDV values

Figure 1 (left) shows the absolute values of SV, CO and EDV measured at the different assessment time-points (T1: before treatment; T2: immediately after chemotherapy or immediately after the 6^{th} cycle of anti-HER2 targeted treatment; T3: four months after chemotherapy or after the 12^{th} cycle of anti-HER2 targeted treatment), presented as CMR/ICG paired values for individual participants. Figure 1 (middle) shows the between-measurement (T2-T1 and T3-T2) changes in SV, CO and EDV (Δ values) obtained using CMR and ICG, presented as CMR/ICG paired values for individual participants. The spreadsheet provided by Weissgerber et al. (2015) was used to produce Figure 1 (left and middle).

A total of 28 absolute SV, CO and EDV measurements were included in the analysis for the determination of the agreement between ICG and CMR using the Bland-Altman technique. Between-modality difference datasets (DiffSV, DiffCO, DiffEDV) were normally distributed and free of proportional bias and heteroscedasticity. Figure 1 (right) presents the Bland-Altman plots constructed for the SV, CO and EDV datasets, respectively. Table 1 shows the Bland and Altman variables calculated for the SV, CO and EDV datasets. The 95% LOA range (i.e. lower LOA-upper LOA) as a proportion of the mean

Accepted Manuscript

CMR value for each variable showed that ICG under-estimated or over-estimated SV, CO and EDV by up to 71%, 94% and 60%, respectively. The one sample t-test showed that the mean value of the normally-distributed DiffSV and DiffCO (3.14 ml and 0.24 L/min, respectively) represented a non-significant bias ($p<0.218$ and $p<0.276$, respectively) between CMR and ICG, but the mean value of the normally-distributed DiffEDV (-26.03 ml) represented a significant bias ($p<0.0005$), with higher values for CMR.

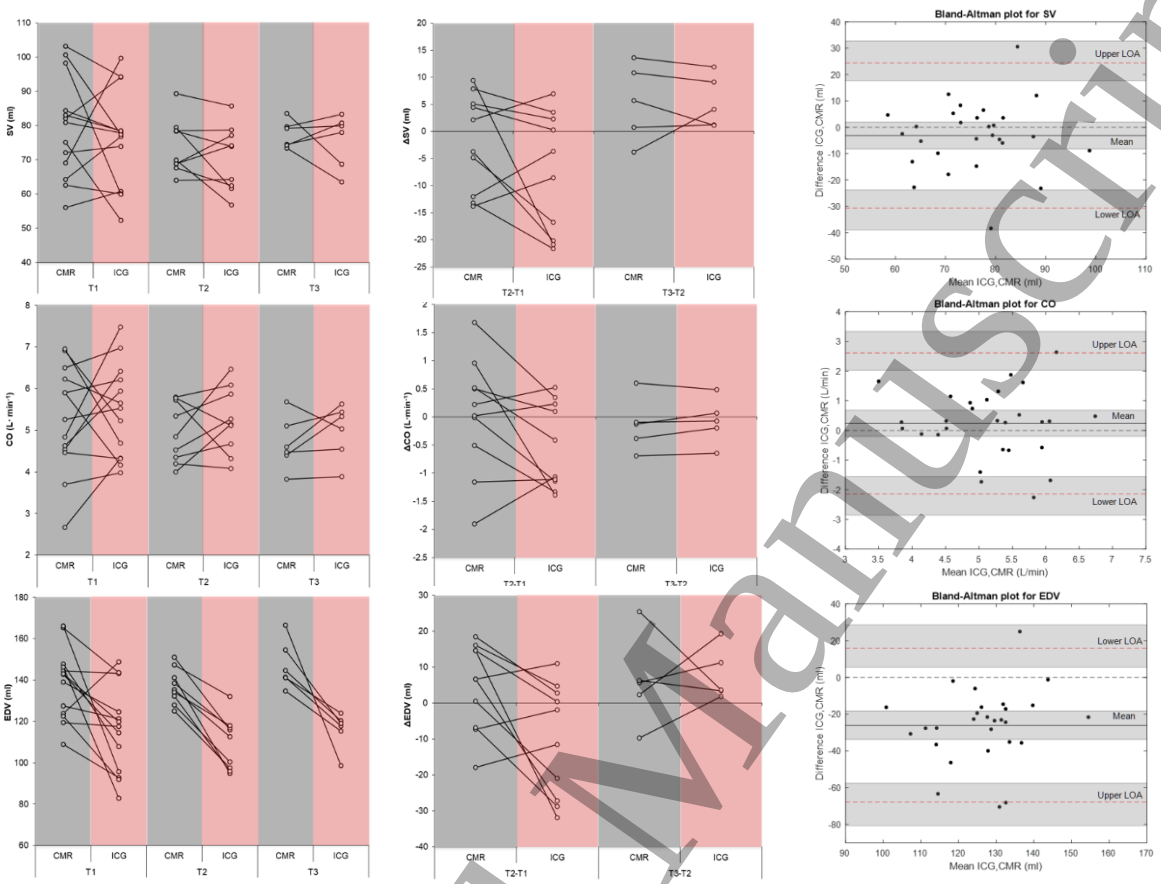


Figure 1. Left: Individual pairs of absolute (a) SV, (b) CO and (c) EDV measurements made by CMR and ICG at the different assessment time-points (T1, T2 and T3). Points included in the grey area represent measurements made by CMR, whereas points included in the red area represent measurements made by ICG; **Middle:** Individual pairs of (a) SV, (b) CO and (c) EDV changes (Δ values) between consecutive measurements (T2-T1 and T3-T2) made by CMR and ICG separately; **Right:** Bland-Altman plots for the (a) SV, (b) CO and (c) EDV datasets. (Black solid line indicates the bias; red dashed lines indicate the 95% LOA and filled grey area indicates the 95% confidence limits for either the bias or the 95%LOA). (Difference values calculated as (ICG-CMR)).

Accepted Manuscript

Table 1. Bland and Altman variables calculated for the SV, CO and EDV datasets.

Bland and Altman variables						
	SV (mL)		CO (L/min)		EDV(mL)	
Bias	-3.1		0.2		-26.0	
SD	13.2		1.1		20.0	
Lower 95% LOA	-30.7		-2.1		-68.0	
Upper 95% LOA	24.4		2.6		15.9	
SEM (bias)	2.5		0.2		3.8	
Percentage Error (%)	36%		45%		33%	
	Lower 95% CI	Upper 95% CI	Lower 95% CI	Upper 95% CI	Lower 95% CI	Upper 95% CI
95% CI for bias	-8.3	2.0	-0.2	0.7	-33.8	-18.2
95% CI for Lower LOA	-23.8	-39.0	-1.5	-2.9	-57.6	-80.7
95% CI for Upper LOA	17.6	32.7	2.0	3.3	5.5	28.6

3.3 Trending analysis of ICG-derived and CMR-derived variables

Fifteen consecutive SV (ΔSV), CO (ΔCO) and EDV (ΔEDV) changes were used in the analysis of the trending ability of ICG relative to CMR. Five participants underwent two ICG and two CMR assessments and five participants underwent three ICG and three CMR assessments (the other six participants either did not undergo a CMR scan (due to claustrophobia or time constraints relative to treatment) or underwent only one CMR scan and one ICG test (due to withdrawal from the study, time constraints relative to treatment, declining to receive chemotherapy or completion of the study).

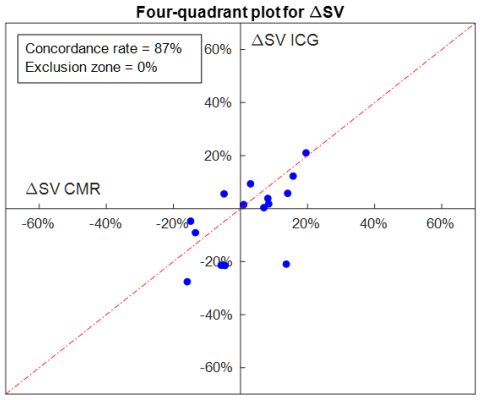
3.3.1 Four-quadrant plot and concordance analysis

Figure 2 shows the four-quadrant plots used to compare the trending ability of ICG against that of CMR (i.e. the similarity of changes (Δ values) in the variables of interest (ΔSV, ΔCO and ΔEDV) for consecutive ICG and CMR measurements). The concordance rate between the ICG and CMR changes was 87% for ΔSV and 73% for both ΔCO and ΔEDV: according to

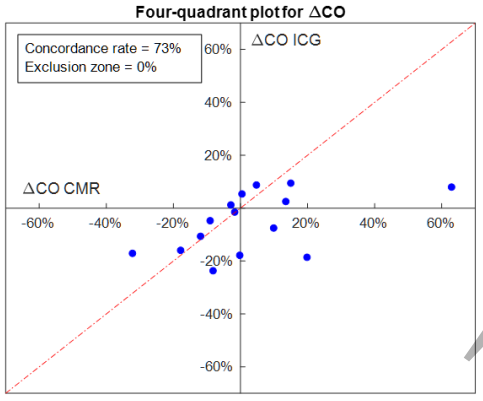
Accepted Manuscript

Critchley et al. (2010), these results fall below the threshold (90%) that indicates good trending ability, suggesting a poor trend-tracking performance of ICG relative to CMR, especially for CO and EDV.

a)



b)



c)

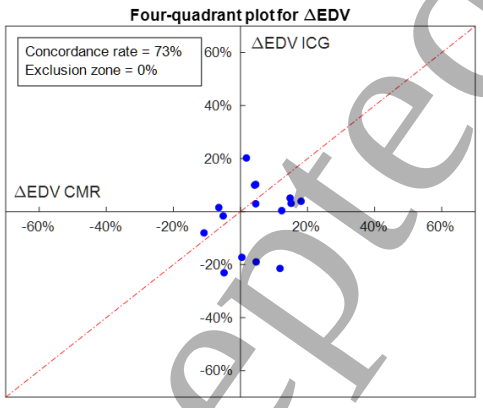


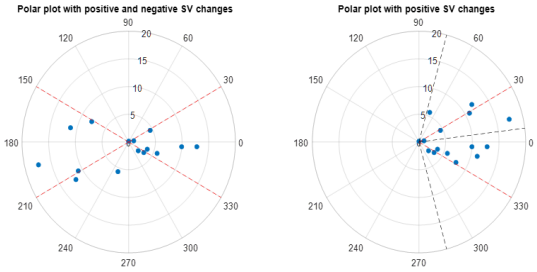
Figure 2. Four-quadrant plots depicting (a) ΔSV, (b) ΔCO and (c) ΔEDV between ICG and CMR. The red dashed line represents the line of identity (y=x), which reflects 100% agreement for changes made by the two methods. All data points were included in the trending analysis (i.e. 0% exclusion zone). A concordance rate of >90% was considered to indicate good trending ability of ICG when compared to CMR (Critchley et al., 2010). ΔSV: sequential SV changes, ΔCO: sequential CO changes, ΔEDV: sequential EDV changes.

Accepted Manuscript

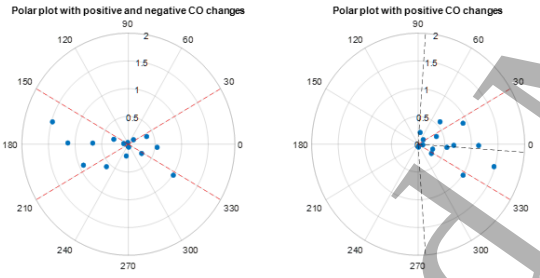
3.3.2 Polar plot method

Figure 3 (a-c) shows the polar plots used to assess concordance for ΔCO, ΔSV and ΔEDV between ICG and CMR. Regarding ΔSV, the polar plot analysis demonstrated a mean polar angle of 7.2° with a standard deviation of 37.5° and 95% radial LOA of ±75°, indicating a relatively poor trending ability of ICG compared to CMR in terms of accurately characterising SV changes. Similar results were found for the trending ability of ICG-derived ΔCO and ΔEDV, with mean polar angles of -4.2° and 1.1°, associated standard deviations of 42.9° and 40.7°, and 95% CIs of ±85.8° and ± 81.5°, respectively (95%CIs were calculated instead of the radial 95% LOA for the ΔCO and ΔEDV, as both mean angles were less than ±5°). The polar concordance rate was 67%, 53% and 33% for ΔSV, ΔCO and ΔEDV (i.e. 67%, 53% and 33% of ΔSV, ΔCO and ΔEDV values, respectively, were included within the ±30° threshold value for the radial 95% LOA (or 95%CIs)), indicating relatively poor trending ability of ICG compared to CMR in terms of characterising CO and EDV changes in our participants. Figure 3 (d) shows a plot of the inclusion rate against different radial size limits, which was used to determine the radial 95% LOA for ΔSV.

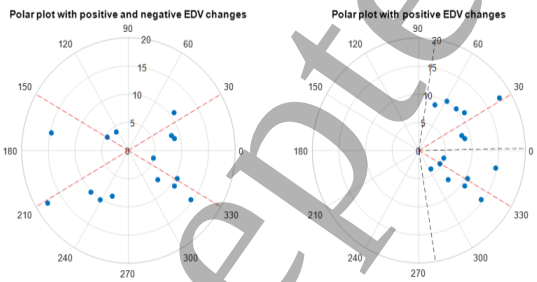
a)



b)



c)



Accepted Manuscript

d)

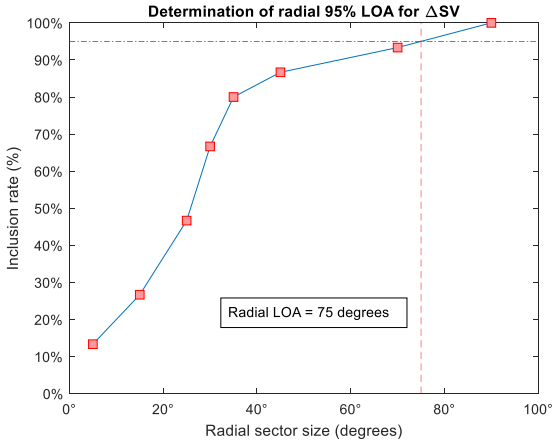


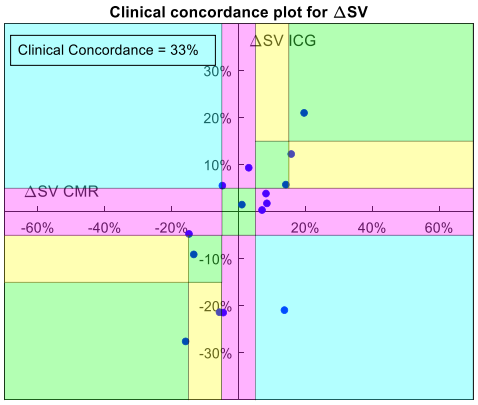
Figure 3 (a-d). Polar plots illustrating the degree of concordance or 'trending ability' between ICG and CMR: (a) ΔSV, (b) ΔCO and (c) ΔEDV. The left-hand polar plots show negative and positive changes for the variable of interest, whereas the right-hand polar plots show absolute values of the changes. (Conversion of all negative changes (points with an angle between 90° and 270°) to absolute (positive) changes was necessary for the calculation of mean polar angle and its SD.) The middle black dotted line (close to 0°) in the right-hand plots represents the mean polar angle, whereas the outer black dotted lines indicate the radial 95% LOA (for ΔSV) or the 95% CIs (for ΔCO and ΔEDV). The red dotted lines in both the left-hand and right-hand polar plots indicate the recommended ±30° cut off values (so points within these lines indicate good trending ability between ICG and CMR); (d) Example of radial sector size determination corresponding to the radial 95% LOA for concordance of ΔSV between ICG and CMR. A perpendicular red dotted line was dropped from where the curve crossed the 95% inclusion rate line (horizontal red dotted line) to indicate the radial sector size that includes 95% of all ΔSV values. (Good trending ability was defined by Critchley et al. (2010, 2011) as a radial 95% LOA of less than ±30°.)

3.3.3 Clinical concordance method

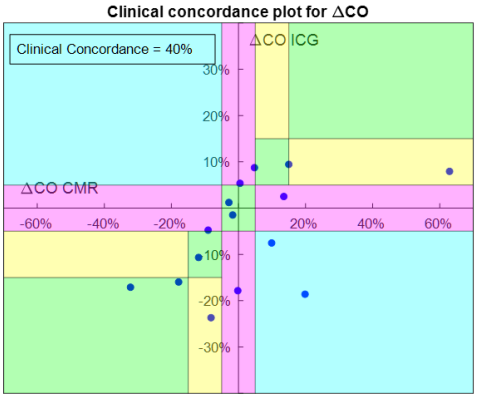
Figure 4 shows the clinical concordance plots for ΔSV, ΔCO and ΔEDV. The clinical concordance rate for SV (proportion of paired values within Zone 1) was 33%, and Zones 2, 3 and 4 included 13%, 47% and 7% of all data points, respectively. Collectively, these findings indicate a poor concordance of the trending abilities of ICG and CMR for SV for our participants. Regarding ΔCO, 40% of the total number of changes fell within Zone 1 (clinical concordance rate), and Zones 2, 3 and 4 included 13%, 33% and 13% of changes, respectively. The clinical concordance rate for ΔEDV was 20%, Zone 2 included no measurements, and Zones 3 and 4 included 73% and 7% of the data points, respectively.. Similar to ΔSV, these findings indicate poor concordance of the trending abilities of ICG and CMR in quantifying serial changes in CO and EDV for our participants.

Accepted Manuscript

a)



b)



c)

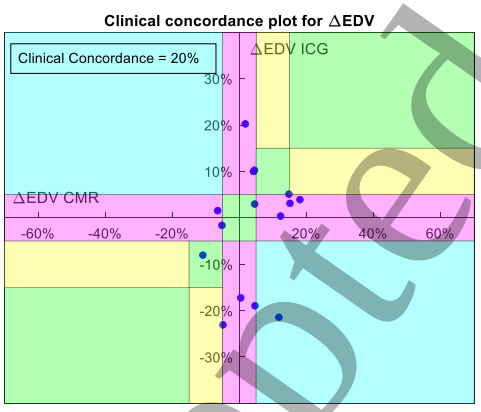


Figure 4. Clinical concordance plots illustrating the clinical concordance between ICG and CMR in quantifying serial (a) ΔSV, (b) ΔCO and (c) ΔEDV. The four different zones of this plot are represented by different colours (Zone 1: green, Zone 2: yellow, Zone 3: magenta and Zone 4: cyan): i) The first zone of the error grid includes changes between consecutive ICG and CMR measurements that are concordant in terms of both direction and magnitude (this zone represents the three different categories used to calculate the concordance rate; for assessments falling in this zone, it would be *correctly* assumed that cardiac function was either significantly affected or not significantly affected by breast cancer treatment, so accurate

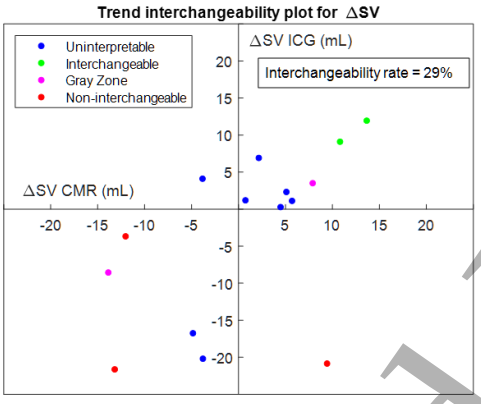
Accepted Manuscript

treatment decisions would be made with ICG; ii) The second zone includes changes that are concordant only regarding direction (either both positive or both negative) but not magnitude; iii) The third zone includes instances of an absence of change with one method (constant value between consecutive measurements) but a change of more than 5% of the initial value with the other method; iv) The fourth zone includes changes that are not concordant in terms of direction, with one method measuring a positive change of more than 5% and the other measuring a negative change of more than 5%. For assessments falling in zones 2-4: it would be *incorrectly* assumed that cardiac function was either stable or had declined, so inaccurate treatment decisions would be made with ICG. The clinical concordance rate is the percentage of points falling into Zone 1 of the error-grid (a clinical concordance rate of >90% was considered to indicate good trending ability of ICG when compared to CMR (Montenij et al., 2016b) for all variables of interest (i.e. SV, CO, EDV)).

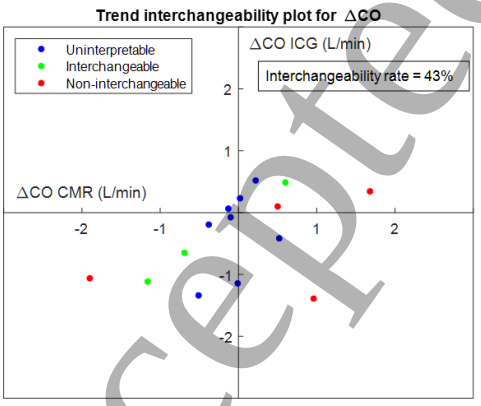
3.3.4 Trend interchangeability method

Figure 5 shows the trend interchangeability plots for assessing agreement (and thus interchangeability) between ICG and CMR in measuring ΔSV, ΔCO and ΔEDV: i) For ΔSV: 7/15 changes (46.7%) were interpretable, of which two were interchangeable (13.3% of the total sample), three were non-interchangeable (20%) and two fell in the grey zone (13.3%); 2) For ΔCO: 7/15 changes (46.7%) were interpretable, three of which were interchangeable (20% of the total sample), four were non-interchangeable (26.7%) and none were in the grey zone; 3) For ΔEDV: 6/15 changes (40%) were interpretable, one of which was interchangeable (6.7% of the total sample), four were non-interchangeable (26.7%) and one fell in the grey zone (6.7%). The trend interchangeability rates for ΔSV, ΔCO and ΔEDV were 29%, 43% and 17%, respectively, reflecting poor interchangeability between ICG and CMR.

a)



b)



Accepted Manuscript

c)

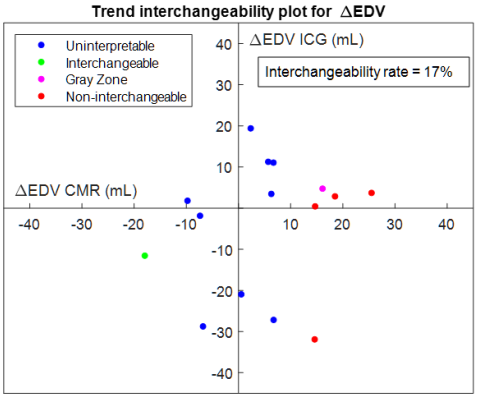


Figure 5. Trend interchangeability plots showing the agreement between ICG and CMR in assessing (a) ΔSV, (b) ΔCO and (c) ΔEDV. Changes are first classified as interpretable or uninterpretable. Changes initially classified as interpretable are then further classified as interchangeable, non-interchangeable and in the grey zone. The acceptable interchangeability rate (i.e. the number of interchangeable points divided by the number of interpretable points) between ICG and CMR was set a priori to >90%, as recommended by Fischer et al. (2016).

4. Discussion

This study sought to investigate whether ICG is a viable method for frequent cardiac assessments in early-stage breast cancer patients, during and following treatment. Serial measurements of the heart's pumping efficiency were compared using ICG and CMR. Our findings indicate that ICG accurately characterised absolute SV and CO (but not EDV) measurements and EDV and CO (but not SV) changes between consecutive measurements. However, the measurement precision for both absolute values and temporal changes in SV, CO and EDV was poor. We therefore conclude that these ICG- and CMR-derived measures are not interchangeable and that ICG cannot be recommended as a replacement for CMR for this purpose.

Bland-Altman analysis revealed that, overall, the agreement between ICG and CMR for absolute SV, CO and EDV measurements was poor: 1) Measurement precision was low, indicated by the wide 95% LOA (-30.66 to 24.38 ml, -2.14 to 2.61 L ·min^{-1} and -67.96 to 15.90 ml, respectively). These showed that ICG either over- or under-estimated SV, CO and EDV by up to 71%, 94% and 60%, respectively, compared to CMR; 2) The accuracy of ICG-derived EDV measurement was poor: a fixed bias (systematic error) of -26.0 ml was observed (with lower values measured by ICG). In contrast, measurement accuracy was good for both SV (-3.1 ml) and CO (0.24 L·min^{-1}); 3) The percentage error was greater than the clinically acceptable 30% threshold value (36%, 45% and 33%, respectively);

This was the first study to compare the performance of ICG and CMR for characterising SV, CO and EDV in breast cancer patients. Five previous studies have examined the performance of ICG in comparison to CMR in other populations: paediatric patients (either with or without cardiac disease) (Taylor et al., 2012), adult healthy participants (Borzage et al., 2017), adult patients referred for a CMR scan (Villacorta Junior et al., 2012), adult and paediatric patients with congenital heart disease (Ebrahim et al., 2016) and adult patients suspected of suffering from pulmonary arterial hypertension (Panagiotou et al., 2018). Our findings of poor precision are in accordance with those reported by Taylor et al. (2012) and Borzage et al. (2017), with ICG showing either an over- or under-estimation of both CI (cardiac index, equal to CO divided by body surface area) and SV of up to 73% and 92%, respectively, compared to CMR. Poor precision in absolute ICG-derived CO measurements (95% LOA: -2.48 to 5.28 L·min^{-1}) was also reported by Panagiotou et al. (2018). However, these authors also reported a poor accuracy (bias: 1.40 L·min^{-1}) for CO, which differs from our results. Furthermore, although our results initially appeared to differ from those of Ebrahim et al. (2016), who reported that absolute ICG-derived and CMR-derived SV had good agreement compared to CMR, it is important to note that those authors based their conclusion only on the mean difference (i.e. bias) between the two techniques. A closer look at the Bland-Altman plot provided in their analysis reveals a wide dispersion of the 95% LOA (approximately -22 to +20 ml) for SV measurements, which is in accordance with our results and indicates poor precision. In contrast, Villacorta Junior et al. (2012) took into account both the bias (approximately 0.3 L·min^{-1}) and the 95% LOA (approximately -1.0 to 1.4 L·min^{-1}) in concluding that ICG-derived absolute CO measurement has acceptable accuracy and precision compared to CMR. Even though the findings of most of these studies (including ours) do

Accepted Manuscript

not support the use of ICG as an alternative to CMR, there is still a need to investigate the performance of different ICG devices in different populations. This is because each ICG device implements its own proprietary algorithms and also uses different types of electrodes to measure haemodynamic data. Our findings extend those of previous studies in that we compared (for the first time in any population) the performances of ICG and CMR in measuring EDV. This is surprising, given the importance of monitoring this variable in the management of health conditions such as heart failure (Yancy et al., 2013). Our study is also one of the first to use ICG to characterise cardiac function in early-stage breast cancer patients *during and following* treatment.

Trending analysis revealed that ICG is also a poor trend-tracker of SV, CO and EDV changes compared to CMR: 1) The four-quadrant plot analysis showed a concordance rate of 87% for serial SV measurements and 73% for both CO and EDV serial measurements (lower than the 90% cut-off value denoting good trending ability); 2) Polar plot analysis showed that radial 95% LOA for SV changes and 95% CIs for CO and EDV changes were considerably greater than the 30% cut-off value denoting good trending ability (with values of ±75°, ±86° and ±81°, respectively) and the angular bias (mean polar angle) was greater than the ±5° cut-off value for SV changes; 3) Error-grid plot analysis demonstrated very low clinical concordance rates of 20%, 40% and 33% for SV, CO and EDV changes (significantly smaller than the 90% threshold value denoting good trending ability). The greatest proportion of changes was observed in Zone 3 for all variables except CO, indicating that, for the most part, one of the two methods (CMR or ICG) provided a much greater temporal change than the other; 4) The trend interchangeability method revealed 29%, 43% and 17% interchangeability rates between ICG and CMR for SV, CO and EDV changes, respectively. These proportions are much lower than the 90% threshold value that denotes good trending ability. Of the five previous studies that compared the performance of ICG against CMR, only Borzage et al. (2017) implemented the polar plot method to investigate its ability to accurately characterise changes in cardiac function (and this was done solely for SV). Although SV angular bias in that study was much larger than in ours (-27° vs ~7°), the radial 95% LOA were almost identical (±73° vs ±75°).

Although our study extended the findings of previous investigations by using statistical techniques to assess the trending ability of ICG in early-stage breast cancer patients, these methods are subject to some important limitations that need to be considered when interpreting the results. First, the authors who introduced the clinical concordance method (Montenij et al., 2016a) did not provide any suggestions for the categorisation of the data points when these fall on the borders of the different zones of the error-grid (as occurred with our data). Second, the use of an exclusion zone in both the four-quadrant plot and polar plot method (where data located around the centre of the plot within a specified limit are excluded from trending analysis) has been based on the supposition that data points falling at the centre of the plot do not reflect real changes, so are considered to not contribute to the trending ability of the method. This is also somewhat true for the trend interchangeability method, where data points are excluded based on the repeatability coefficient of the reference technique. However, in line with Montenij et al. (2016a), we would also argue that a change does not need to be substantial in order to be meaningfully included in the trending ability analysis. If, for example, CMR-derived CO did not show a significant change between serial measurements in breast cancer patients, clinicians would assume that cardiac function has not been affected by anticancer treatment (i.e. this is still a clinically meaningful result). Therefore, very small changes of similar magnitude and direction measured concurrently by ICG and CMR could reflect a good trending ability of ICG compared to CMR. We believe that the clinical concordance method is the most appropriate method for evaluating trending ability for the cardiac variables considered in this study. This method does not exclude any points from the trending ability analysis and is also much simpler to perform than the polar plot and trend interchangeability methods.

4.1.1 Strengths and limitations of the study

This study extends our knowledge of the performance of ICG as a potential method for monitoring cardiac function in early-stage breast cancer patients. ICG has been used previously but it has never been compared with a gold standard technique in this population. The majority of previous studies that investigated the accuracy, precision and trending ability of ICG have used (as their reference method) techniques that are both less accurate and less precise than CMR (such as echocardiography and thermodilution). Moreover, as far as we are aware, no study to date has assessed the performance of ICG in terms of EDV measurement, despite the diagnostic importance of this variable in various diseases. Care was also taken to minimise variability originating from sources other than the treatment; this was achieved by asking participants to avoid physical activity and alcohol, caffeine and tea consumption in the 24 hours prior to ICG and CMR assessments, and standardising pre-assessment water consumption.

Our study's sample size (n=28 for absolute values and n=15 for temporal changes) was relatively small and so the statistical power of our analyses was reduced. Also, we could not perform simultaneous measurements with ICG and CMR, as this was practically unfeasible (the Task Force Monitor utilises a number of electrodes placed on the neck, chest and lower abdomen

Accepted Manuscript

that cannot be left *in situ* during a CMR scan). We were unable to perform sub-group analyses to investigate whether agreement between ICG and CMR was different between patients receiving only chemotherapy and those receiving chemotherapy plus anti-HER2 targeted therapy, or between the different assessment time-points. Performing separate Bland-Altman and trending-ability analyses for these purposes requires a greater number of participants than was available here for these analyses to be meaningful. We also used recommended threshold points (denoting good agreement and good trending ability) that were designed specifically for CO measurements and that had used thermodilution as the reference technique (which is less accurate and precise than CMR for this purpose). Given that repeated measurements with both ICG and CMR at each assessment time-point were outside the scope of this study (being practically unfeasible because of the constraints of treatment plans and other assessments), we also used an arbitrary value for the repeatability coefficient of CMR to perform the trend interchangeability analysis. Although it is known that CMR is a very accurate and reproducible technique, there is a paucity of information in the literature regarding the repeatability of CMR scans in early-stage breast cancer patients (and in other populations), so we made assumptions about the quality of the CMR-derived data. Finally, we evaluated the performance of ICG specifically in early-stage breast cancer patients, which limits the generalisability of our results to patients who have metastatic breast cancer.

4. Conclusion

As far as we are aware, this was the first study to investigate the performance of ICG compared to CMR for characterising both absolute values and temporal changes in SV, CO and EDV measurements in early-stage breast cancer patients. We used a combination of established and novel statistical techniques to provide a comprehensive evaluation of the capabilities of ICG for assessing cardiac function in this population. Specifically, we showed that, accuracy was good for absolute SV and CO values and for EDV and CO changes. However, precision was poor for all variables in terms of both absolute measurements and trend tracking performance. This suggests that ICG is not interchangeable with CMR and so is probably not a useful technique for monitoring cardiac function in early-stage breast cancer patients during or following treatment.

Acknowledgements

The authors wish to thank the breast cancer patients for participating in this study, Mrs Lucy Carr for assisting with the collection of patient information and Dr Olga Roldanreoyo for assisting with the assessments. The funding for this study was provided by Professor Parvaiz Ali and the funding for the CMR scans was provided by Dr Gianfilippo Bertelli.

References

Altena, R., Perik, P. J., Van Veldhuisen, D. J., De Vries, E. G., & Gietema, J. A. (2009). Cardiovascular toxicity caused by cancer treatment: strategies for early detection. *The Lancet Oncology*, 10(4), 391-399.

Armenian, S. H., Lacchetti, C., Barac, A., Carver, J., Constine, L. S., Denduluri, N., Dent, S., Douglas, P.S., Durand, J.B., Ewer, M. & Fabian, C. (2016). Prevention and monitoring of cardiac dysfunction in survivors of adult cancers: American Society of Clinical Oncology Clinical Practice Guideline. *Journal of Clinical Oncology*, 35:8, 893-911.

Armstrong, G. T., Plana, J. C., Zhang, N., Srivastava, D., Green, D. M., Ness, K. K., Donovan, F.D., Metzger, M.L., Arevalo, A., Durand, J.B. & Joshi, V. (2012). Screening adult survivors of childhood cancer for cardiomyopathy: comparison of echocardiography and cardiac magnetic resonance imaging. *Journal of clinical oncology*, *30*(23), 2876.

Bellenger, N. G., Burgess, M. I., Ray, S. G., Lahiri, A., Coats, A. J., Cleland, J. G., & Pennell, D. J. (2000). Comparison of left ventricular ejection fraction and volumes in heart failure by echocardiography, radionuclide ventriculography and cardiovascular magnetic resonance. Are they interchangeable?. *European heart journal*, *21*(16), 1387-1396.

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical methods in medical research*, *8*(2), 135-160.

Bloom, M. W., Hamo, C. E., Cardinale, D., Ky, B., Nohria, A., Baer, L., Skopicki, H., Lenihan, D.J., Gheorghiade, M., Lyon, A.R. & Butler, J. (2016). Cancer Therapy–Related Cardiac Dysfunction and Heart Failure: Part 1: Definitions, Pathophysiology, Risk Factors, and Imaging. *Circulation: Heart Failure*, *9*(1), e002661.

Bovelli, D., Plataniotis, G., Roila, F. E. G. W., & ESMO Guidelines Working Group. (2010). Cardiotoxicity of chemotherapeutic agents and radiotherapy-related heart disease: ESMO Clinical Practice Guidelines. *Annals of Oncology*, 21(suppl_5), v277-v282.

Accepted Manuscript

Cancer Research UK. (2020) Breast Cancer Statistics. Retrieved from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer#heading-Two.

Carkeet, A. (2015). Exact parametric confidence intervals for Bland-Altman limits of agreement. *Optometry and Vision Science*, *92*(3), e71-e80.

Chung R., Ghosh A.K., Banerjee A. (2018). Cardiotoxicity: precision medicine with imprecise definitions. Open Heart, 5:e000774. doi: 10.1136/openhrt-2018-000774.

Critchley, L. A., & Critchley, J. A. (1999). A meta-analysis of studies using bias and precision statistics to compare cardiac output measurement techniques. *Journal of clinical monitoring and computing*, *15*(2), 85-91.

Critchley, L. A., Lee, A., & Ho, A. M. H. (2010). A critical review of the ability of continuous cardiac output monitors to measure trends in cardiac output. *Anesthesia & Analgesia*, *111*(5), 1180-1192.

Critchley, L. A., Yang, X. X., & Lee, A. (2011). Assessment of trending ability of cardiac output monitors by polar plot methodology. *Journal of cardiothoracic and vascular anesthesia*, *25*(3), 536-546.

Critchley, L. (2017). Meta-analyses of bland–altman-style cardiac output validation studies: Good, but do they provide answers to all our questions? *British Journal of Anaesthesia*, *118*(3), 296-297. doi:10.1093/bja/aew442.

Ebrahim, M., Hegde, S., Printz, B., Abcede, M., Proudfoot, J. A., & Davis, C. (2016). Evaluation of impedance cardiography for measurement of stroke volume in congenital heart disease. *Pediatric cardiology*, *37*(8), 1453-1457.

Fischer, M. O., Diouf, M., de Wilde, R. B., Dupont, H., Hanouz, J. L., & Lorne, E. (2016). Evaluation of cardiac output by 5 arterial pulse contour techniques using trend interchangeability method. *Medicine*, *95*(25).

Fortin, J., Habenbacher, W., Heller, A., Hacker, A., Grüllenberger, R., Innerhofer, J., Passath, H., Wagner, C.H., Haitchi, G., Flotzinger, D. & Pacher, R. (2006). Non-invasive beat-to-beat cardiac output monitoring by an improved method of transthoracic bioimpedance measurement. *Computers in Biology and Medicine*, 36(11), 1185-1203.

Hapfelmeier, A., Cecconi, M., & Saugel, B. (2016). Cardiac output method comparison studies: the relation of the precision of agreement and the precision of method. *Journal of clinical monitoring and computing*, *30*(2), 149-155.

Jones, L. W., Haykowsky, M., Pituskin, E. N., Jendzjowsky, N. G., Tomczak, C. R., Haennel, R. G., & Mackey, J. R. (2007). Cardiovascular reserve and risk profile of postmenopausal women after chemoendocrine therapy for hormone receptor–positive operable breast cancer. *The oncologist*, 12(10), 1156-1164.

Kasznicki, J., & Drzewoski, J. (1993). The importance of impedance cardiography in monitoring cardiac function in patients with hematological malignancies. *Acta haematologica Polonica*, *24*(2), 123-130.

Kolla, B. C., Roy, S. S., Duval, S., Weisdorf, D., Valeti, U., & Blaes, A. (2017). Cardiac imaging methods for chemotherapy-related cardiotoxicity screening and related radiation exposure: current practice and trends. *Anticancer research*, *37*(5), 2445-2449.

Lentner, C., Diem, K., & Seldrup, J. (1982). Introduction to statistics. Statistical tables. Mathematical formulae. *Geigy scientific tables*, *2*, 215-218.

Ludbrook, J. (2010). Confidence in Altman–Bland plots: a critical review of the method of differences. *Clinical and Experimental Pharmacology and Physiology*, *37*(2), 143-149.

Manrique, C. R., Park, M., Tiwari, N., Plana, J. C., & Garcia, M. J. (2017). Diagnostic Strategies for Early Recognition of Cancer Therapeutics–Related Cardiac Dysfunction. Clinical Medicine Insights: Cardiology, 11, 1179546817697983.

Massidda, B., Fenu, M. A., Ionta, M. T., Tronci, M., Foddi, M. R., Montaldo, C., & Montaldo, P. L. (1997). Early detection of the anthracycline-induced cardiotoxicity. A non-invasive haemodynamic study. *Anticancer research*, *17*(1B), 663-668.

Mihalcea, D. J., Florescu, M., & Vinereanu, D. (2017). Mechanisms and genetic susceptibil-ity of chemotherapy-induced cardiotoxicity in patients with breast cancer. *American journal of therapeutics*, 24(1), e3-e11.

Montenij, L. J., Buhre, W. F., Jansen, J. R., Kruitwagen, C. L., & De Waal, E. E. (2016a). Methodology of method comparison studies evaluating the validity of cardiac output monitors: a stepwise approach and checklist. *British journal of anaesthesia*, *116*(6), 750-758.

Accepted Manuscript

Montenij, L. J., Sonneveld, J. P., Nierich, A. P., Buhre, W. F., & De Waal, E. E. (2016b). Accuracy, precision, and trending ability of uncalibrated arterial pressure waveform analysis of cardiac output in patients with impaired left ventricular function: a prospective, observational study. *Journal of cardiothoracic and vascular anesthesia*, 30(1), 115-121.

Odor, P. M., Bampoe, S., & Cecconi, M. (2017). Cardiac Output Monitoring: Validation Studies–how Results Should be Presented. *Current anesthesiology reports*, *7*(4), 410-415.

Panagiotou, M., Vogiatzis, I., Jayasekera, G., Louvaris, Z., Mackenzie, A., Mcglinchey, N., Baker, J.S., Church, A.C., Peacock, A.J. & Johnson, M. K. (2018). Validation of impedance cardiography in pulmonary arterial hypertension. *Clinical physiology and functional imaging*, *38*(2), 254-260.

Perrino, J. A., O'connor, T., & Luther, M. (1994). Transtracheal Doppler cardiac output monitoring: comparison to thermodilution during noncardiac surgery. *Anesthesia and analgesia*, *78*(6), 1060-1066.

Perrino, A. C., Harris, S. N., & Luther, M. A. (1998). Intraoperative Determination of Cardiac Output Using Multiplane Transesophageal Echocardiography A Comparison to Thermodilution. *Anesthesiology: The Journal of the American Society of Anesthesiologists*, *89*(2), 350-357.

Plana, J. C., Galderisi, M., Barac, A., Ewer, M. S., Ky, B., Scherrer-Crosbie, M., Ganame, J., Sebag, I.A., Agler, D.A., Badano, L.P. & Banchs, J. (2014). Expert consensus for multimodality imaging evaluation of adult patients during and after cancer therapy: a report from the American Society of Echocardiography and the European Association of Cardiovascular Imaging. *European Heart Journal–Cardiovascular Imaging*, *15*(10), 1063-1093.

Squara, P., Cecconi, M., Rhodes, A., Singer, M., & Chiche, J. D. (2009). Tracking changes in cardiac output: methodological considerations for the validation of monitoring devices. *Intensive Care Medicine*, **35**, 1801–1808.

Taylor, K., Manlhiot, C., McCrindle, B., Grosse-Wortmann, L., & Holtby, H. (2012). Poor accuracy of noninvasive cardiac output monitoring using bioimpedance cardiography [PhysioFlow®] compared to magnetic resonance imaging in pediatric patients. *Anesthesia & Analgesia*, *114*(4), 771-775.

Villacorta Junior H., Villacorta A.S., Amador F., Hadlich M., Albuquerque D.C., Azevedo C.F. (2012). Transthoracic impedance compared to magnetic resonance imaging in the assessment of cardiac output. *Arquivos Brasileiros de Cardiologia*, 99(6):1149–1155.

Virani, S. A., Dent, S., Brezden-Masley, C., Clarke, B., Davis, M. K., Jassal, D. S., D.S., Johnson, C., Lemieux, J., Paterson, I., Sebag, I.A & Simmons, C. (2016). Canadian Cardiovascular Society guidelines for evaluation and management of cardiovascular complications of cancer therapy. *Canadian Journal of Cardiology*, *32*(7), 831-841.

Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond bar and line graphs: time for a new data presentation paradigm. *PLoS biology*, *13*(4).

Yancy, C. W., Jessup, M., Bozkurt, B., Butler, J., Casey, D. E., Drazner, M. H., Fonarow, G.C., Geraci, S.A., Horwich, T., Januzzi, J.L. & Johnson, M. R. (2013). 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. *Journal of the American College of Cardiology*, *62*(16), e147-e239.

Zamorano, J. L., Lancellotti, P., Rodriguez Muñoz, D., Aboyans, V., Asteggiano, R., Galderisi, M., Habib, G., Lenihan, D.J., Lip, G.Y., Lyon, A.R. & Lopez Fernandez, T. (2016). 2016 ESC Position Paper on cancer treatments and cardiovascular toxicity developed under the auspices of the ESC Committee for Practice Guidelines: The Task Force for cancer treatments and cardiovascular toxicity of the European Society of Cardiology (ESC). *European heart journal*, *37*(36), 2768-2801.

Accepted Manuscript