Automated contouring of non-contrast enhanced echocardiograms result in similar estimates of left ventricular function to manually contoured contrast enhanced images in chemotherapy patients

Ashley Akerman[1], Lori Bernard[2], Tine Deschamps[1], Bernadette Foster[2], Will Hawkes[1], Eila Mirhadi[2], Hania Piotrowska[1], Rizwan Sarwar[3], Louise Tetlow[1], Gary Woodward[1], Harald Becher[2]

Background

- Transthoracic echocardiography (TTE) assessment of the left ventricle (LV) is central in the early detection of cancer therapy-related cardiac dysfunction (**CTRCD**).
- Measurement variability ca be reduced by both contrast enhancement of TTEs, and automated contouring of the LV with **artificial intelligence** (AI).

Methodology

- Single centre retrospective study.
- Patients monitored for CRTCD at Mazankowski Alberta Heart Institute (Edmonton, Canada). TTE at onset of cancer therapy, and after at least 3 months of treatment.
- Manually contoured contrast enhanced images (**MAN-CON**) with IntelliSpace Cardiovascular (Phillips) compared to automated contouring on non-contrast enhanced images (**AUTO-NON**) using EchoGo Core (Ultromics, UK).
- Differences and agreement between methods evaluated by: (i) statistical equivalence (two one-sided t-tests), (ii) error between methods, via root-mean squared error (RMSE; Deming Regression), and (iii) average bias and associated 95% confidence interval (Bland-Altman).
- Reproducibility estimates in a similar cohort informed equivalence bounds (EDV, 20 mL; ESV, 15 mL; EF, 5%; *Thavendiranathen et al., 2013, JACC*), and were used to interpret mean error and bias.

Outcome	Breast Cancer	Lymphoma	Myeloma	Kruskall-Wallis Test
Age (years)	55 (14)	60 (24)	58 (9)	0.052
Height (cm)	162 (7)	175 (16.5)	176 (21)	<0.001
Weight (kg)	70.0 (21.0)	80.0 (23.5)	80.0 (29.3)	0.002
BMI (kg/m2)	25.8 (7.5)	27.4 (5.4)	26.9 (5.7)	0.315
SBP (mm Hg)	122 (19)	112 (31)	126 (20)	0.063
DBP (mm Hg)	73 (11)	72 (14)	79 (11)	0.12

Table 1. *Patient demographics as mean (SD) for the diagnosis of breast cancer, Lymphoma, and Myeloma.*

Table 2. *Tests of differences and agreement between MAN-CON and AUTO-NON for end-diastolic volume (EDV), end-systolic volume (ESV), and ejection fraction (EF).*

Outcome	Diagnosis	Baseline				Follow-up			
		n	*Equivalence*	*Error*	*Bias ± 95%CI*	*n*	*Equivalence*	*Error*	*Bias ± 95%CI*
EDV	All	202	<0.001	20*	9 ± 3*	138	<0.001	21	6 ± 3*
EDV	Breast	151	<0.001	18*	8 ± 3*	112	<0.001	19*	6 ± 4*
EDV	Lymphoma	27	0.028	24	9 ± 11*	18	0.031	31	6 ± 15*
EDV	Myeloma	24	0.083	26	12 ± 11*	8	0.011	14*	4.3 ± 13*
ESV	All	202	<0.001	11*	5 ± 2*	138	<0.001	13*	5 ± 2*
ESV	Breast	151	<0.001	9*	4 ± 1*	112	<0.001	11*	5 ± 2*
ESV	Lymphoma	27	0.02	16	8 ± 6*	18	0.036	23	5 ± 11*
ESV	Myeloma	24	0.001	15*	4 ± 6*	8	0.005	8*	5 ± 7*
EF	All	203	<0.001	6.8	-0.7 ± 0.9*	138	<0.001	6.4	-2.0 ± 1.1*
EF	Breast	152	<0.001	6.5	-0.8 ± 1.1*	112	<0.001	5.8	-1.9 ± 1.2*
EF	Lymphoma	27	<0.001	6.4	-2.5 ± 2.4*	18	0.063	8.2	-2.0 ± 3.9*
EF	Myeloma	24	0.009	7	1.5 ± 2.8*	8	0.17	6.7	-2.6 ± 5.6*

Notes*: * denotes error or bias less than or equal to defined equivalence bounds. Equivalence indicates test statistic outcome from two one-sided t-tests pertaining to defined equivalence bounds (EDV, 20 mL; ESV, 15 mL; EF, 5%). Error relates to RMSE from Deming Regression. Bias relates to mean bias from Bland-Altman test and associated 95% confidence interval (CI).*

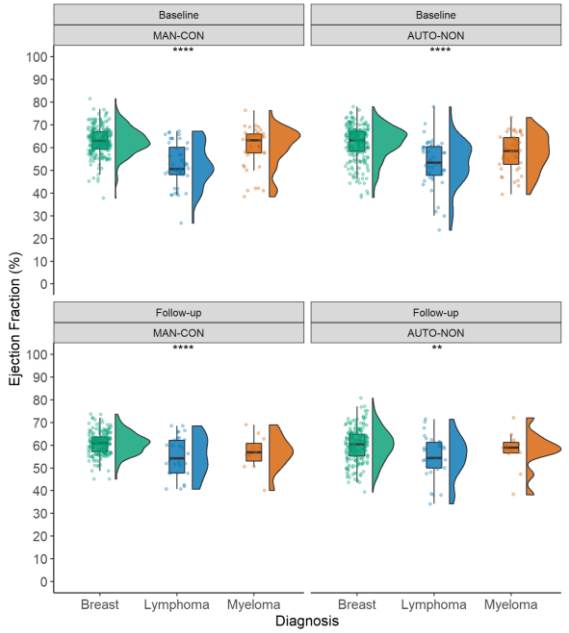

Figure 1. *Ejection fraction (%) determined by manual and automated contouring at baseline and follow up. Breast cancer patients denoted in green, Lymphoma patients in blue, and Myeloma patients in orange. Stars indicates statistical difference between groups via ANOVA.*

Results

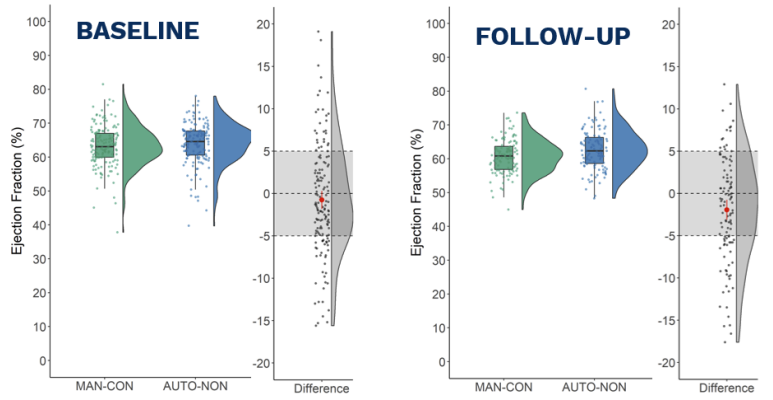

Figure 2. *Difference in estimates of ejection fraction (EF) for Breast cancer patients at baseline (left) and follow up (right) between MAN-CON (green) and AUTO-NON (blue). Regions of equivalence (grey) and mean ± 95% CI (red) highlighted in plot of differences.*

- AUTO-NON differentiated functional cardiac differences between diagnoses, at baseline and during cancer therapy (Figure 1)
- Most differences between MAN-CON and AUTO-NON were **small enough to be considered statistically equivalent** (Table 2 and Figure 2).
- For all outcomes and diagnoses, mean bias (Bland-Altman) was also within acceptable limits (Table 2).

Conclusion

- Despite poorer image quality, Automated contouring of non-contrast enhanced TTE are comparable to manually contoured contrast enhanced images.
- AI contouring of non-contrast enhanced TTEs has the potential to **improve detection and management of CTRCD**.

Affiliations: 1. Ultromics Ltd, Oxford, UK; 2. Mazankowski Alberta Heart Institute, Edmonton, Canada; 3. Experimental Therapeutics, University of Oxford. **Corresponding email**: clinicalresearch@ultromics.com