Clinical Investigation Report Synopsys

Ligence Heart 2.1 software as a medical device verification and validation

(CIV-ID: CIV-LT-21-02-035769**)**

1 List of Abbreviations and Terms

2D TTE	2-dimensional transthoracic echocardiography
RAMAD	Right Aclinical investigation Minor Axis Dimension (a4ch)
AoA	Aortic Annulus
AoS	Aortic Sinus Diameter
STJ	Sinotubular Junction
AAo	Ascending Aorta Diameter
LAD	Left Atrial Diameter (PLA view)
LAV4A	Left Atrial Volume Apical 4 Chamber
IVSd	Interventricular Septum (diastole)
LVPWd	Left Ventricle Posterior Wall (Diastole)
LVEDD	Left Ventricle End-Diastolic Diameter
LVESD	Left Ventricle End-Systolic Diameter
LVEDV4A	LV End Diastolic Volume Apical 4 Chamber
LVESV4A	LV End Systolic Volume Apical 4 Chamber
LVEDV2A	LV End Diastolic Volume Apical 2 Chamber
LVESV2A	LV End Systolic Volume Apical 2 Chamber
RAA	Right Atrial Area
RVB	Right Ventricular Basal Diameter
RVM	Right Ventricular Middle Diameter
RVOTPD	Right Ventricular Outflow Tract Proximal Diameter (PLA)
RVEDA	Right Ventricle End Diastolic Area
RVESA	Right Ventricle End Systolic Area

2 Study Objectives and Endpoints

2.1 Purpose of the Study

The purpose of this study is to evaluate the safety and performance of the manual and automatic functionalities of medical device software Ligence Heart comparing them to the state-of-the-art alternatives.

2.2 Objectives

2.2.1 Primary Objectives

1. Manual measurement analysis: To calculate the reliability of Ligence Heart manual functions that are used to perform echocardiographic measurements comparing it with the manual measurements performed with other CE marked state-of-the-art medical image viewers.

2. Automatic measurement analysis: To compare Ligence Heart automatic measurements accuracy, variance, and error rate with human physicians in real clinical setting.
3. Time used comparison: To compare Ligence Heart automatic measurement tool execution time with the time it takes a physician to perform the measurements manually.

2.2.2 Safety Objective(s)

To collect safety information, including type and number of AEs, SAEs, and device issues.

2.3 Study Endpoints

2.3.1 Primary Endpoints

The end points respectively to the objective number:

1. Echocardiographic measurements performed manually with Ligence Heart and comparator software. The inter-software reliability between manual measurements performed with Ligence Heart and other CE marked state-of-the-art medical image viewers will be calculated (Part 1 of the investigation);
2. Echocardiographic measurements performed automatically by Ligence Heart software and echocardiographic measurements performed manually by physicians. The accuracy of automatic functions of Ligence Heart will be compared with the accuracy of physicians performing measurements manually (Part 2 of the investigation);
3. Time used to perform echocardiographic measurements manually and time used to perform the measurements automatically.

2.3.2 Safety Endpoints(s)

Type and number of AEs, SAEs, and device issues.

2.4 Summary of Study Design

In this observational quantitative study (Clinical trial register number: CIV-LT-21-02-035769) patients were enrolled retrospectively from a database of 17410 2D TTE studies. Enrolled subjects were studied at the Republican Siauliai Hospital for various indications and had undergone a 2D TTE examination from 1 August 2010 to 1 August 2020. From the database we randomly picked studies using a random number generator and evaluated them based on inclusion/exclusion criteria until 58 patients were selected. The study protocol was approved by the Vilnius Regional Biomedical Research Ethics Committee, and informed consent was waived due to the retrospective nature of the analysis.

3 Statistical Methods

3.1 Statistical Hypothesis

Automatic functionalities of Ligence Heart software perform echocardiography image analysis faster, with non-inferior accuracy compared to a cardiologist.

3.2 Statistical Analysis

Two-way ANOVA was employed in this analysis to prepare for further investigation. Standard error of measurement (SEM) was calculated from the results of ANOVA. SEM and used to get minimal detectable change. Absolute difference between the automated system measurement and mean of the raters' measurements was compared with minimal detectable change as the primary endpoint of the analysis to check if there is a significant difference between physicians and automated system. Same analysis was conducted with measurements of the 4th rater.

4 Investigation Results

4.1 Accuracy of automatic measurements

In order to establish the baseline inter- and intra-observer variability, the original rater group (ORG) consisting of three board certified cardiologists analyzed 58 2D TTE studies, repeating the measurements twice. This generated six values for each measurement in each study. The new rater group (NRG) consisting of a fourth cardiologist (FC) and Ligence Heart performed measurements in the same 58 studies. NRG performance was evaluated by calculating the number of measurements of each type that were in the limits of variation. FC was used as the performance benchmark for Ligence Heart.

There was no significant difference in variation between Ligence Heart and FC ($p > 0.05$) and Ligence Heart had non-inferior accuracy to FC for all automatic measurements **(Table 1)**. The lowest number of studies in agreement with ORG was 93.1% in RAA and 94.55% in RVM for Ligence Heart and FC, respectively. For AoS, STJ, IVSd, LVEDD, LVESD, LVESV4A and RVOTPD both FC and Ligence Heart were in agreement with ORG for 100% of studies. Ligence Heart discarded 2-4 studies in RAMAD, LAV4A, LVEDV4A, LVESV4A, LVEDV2A, LVESV2A, RVEDA and RVESA due to insufficient quality scores of automatic predictions **(Table 1)**. Ligence Heart confidence interval for the number of studies in agreement with ORG intersected FC for all automatic measurements **(Figure 1)** and comparing each measurement between FC and Ligence Heart yielded a median Pearson correlation R 0.74 (IQR 0.59-0.83).

Table 1. Comparison of Ligence Heart and cardiologist agreement with original rater group.

Measurement	Ligence Heart studies in agreement with ORG (95% CI)	FC studies in agreement with ORG (95% CI)	P-value*	N studies	N studies not passing confidence filter in Ligence Heart	Pass
RAMAD	96.3 ± 5.03	98.15 ± 3.59	0.56	54	4	Yes
AoA	98.28 ± 3.35	98.28 ± 3.35	1	58	0	Yes
AoS	100.0 ± 0.0	100.0 ± 0.0		58	0	Yes
STJ	100.0 ± 0.0	100.0 ± 0.0		58	0	Yes
AAo	98.28 ± 3.35	100.0 ± 0.0	0.32	58	0	Yes
LAD	94.83 ± 5.7	100.0 ± 0.0	0.08	58	0	Yes
LAV4A	98.21 ± 3.47	94.64 ± 5.9	0.31	56	2	Yes
IVSd	100.0 ± 0.0	100.0 ± 0.0		58	0	Yes
LVPWd	96.55 ± 4.7	100.0 ± 0.0	0.16	58	0	Yes
LVEDD	100.0 ± 0.0	100.0 ± 0.0		58	0	Yes
LVESD	100.0 ± 0.0	100.0 ± 0.0		58	0	Yes
LVEDV4A	98.21 ± 3.47	100.0 ± 0.0	0.32	56	2	Yes
LVESV4A	100.0 ± 0.0	100.0 ± 0.0		56	2	Yes
LVEDV2A	96.3 ± 5.03	100.0 ± 0.0	0.16	54	4	Yes
LVESV2A	96.3 ± 5.03	100.0 ± 0.0	0.16	54	4	Yes
RAA	94.44 ± 6.11	98.15 ± 3.59	0.31	54	4	Yes

RVB	98.18 ± 3.53	98.18 ± 3.53	1	55	3	Yes
RVM	100.0 ± 0.0	94.55 ± 6.0	0.08	55	3	Yes
RVOTPD	100.0 ± 0.0	100.0 ± 0.0		58	0	Yes
RVEDA	98.18 ± 3.53	100.0 ± 0.0	0.32	55	3	Yes
RVESA	98.18 ± 3.53	100.0 ± 0.0	0.32	55	3	Yes

*ORG - original rater group consisting of three board certified cardiologists; FC – fourth board certified cardiologist. *P values for measurements that fall 100% in agreement in both groups are undefined. The "Pass" column specifies whether measurement accuracy is considered to be non-inferior. In order for the automated measurements to pass, a P value of > 0.05 or undefined is required which means that there is no significant difference in variation between Ligence Heart and FC measurements.*

4.2 Intra-rater variability and time comparison

Comparing measurements for each ORG member between different runs resulted in median R of 0.85 (IQR 0.73 - 0.88), 0.81 (IQR 0.73 - 0.87) and 0.78 (IQR 0.66 - 0.84) for original raters 1, 2 and 3, respectively. Due to Ligence Heart automatic measurements being determined by the input, Ligence Heart had no variation in measurements between different runs and had R of 1.0 for all measurements **(Table 2)**, resulting in significantly lower intra-rater variability ($p < 0.05$).

ORG members on average took 12:58 ± 3:18 minutes to analyze the same measurements in 58 TTE studies while Ligence Heart was significantly faster, taking on average 2:59 ± 1:02 minutes to analyze the same studies ($p < 0.05$).

Table 2. Comparison of correlation coefficients between different runs in original rater group and Ligence Heart.

Label	OR1 Run 1 vs Run 2	OR2 Run 1 vs Run 2	OR3 Run 1 vs Run 2	Ligence Heart Run 1 vs Run 2
RAMAD	0.86	0.86	0.82	1
AoA	0.57	0.73	0.65	1
AoS	0.86	0.83	0.83	1
STJ	0.51	0.77	0.59	1
AAo	0.73	0.46	0.72	1
LAD	0.83	0.83	0.8	1
LAV4A	0.85	0.92	0.84	1
IVSd	0.88	0.62	0.42	1
LVPWd	0.66	0.35	0.41	1
LVEDD	0.92	0.86	0.89	1
LVESD	0.83	0.88	0.78	1
LVEDV4A	0.88	0.81	0.84	1
LVESV4A	0.93	0.91	0.88	1

LVEDV2A	0.93	0.87	0.87	1
LVESV2A	0.96	0.92	0.88	1
RAA	0.9	0.89	0.88	1
RVB	0.75	0.74	0.78	1
RVM	0.73	0.75	0.71	1
RVOTPD	0.88	0.77	0.77	1
RVEDA	0.75	0.68	0.66	1
RVESA	0.7	0.68	0.61	1

Pearson correlation coefficients between different runs for the same rater are shown. Since Ligence Heart performs all measurements automatically and the output is completely determined by input, it has a correlation coefficient of 1 for all measurements. Run 1 and run 2 refers to different repeats of the study by the same rater. OR1 – original rater one; OR2 – original rater two; OR3 – original rater three.

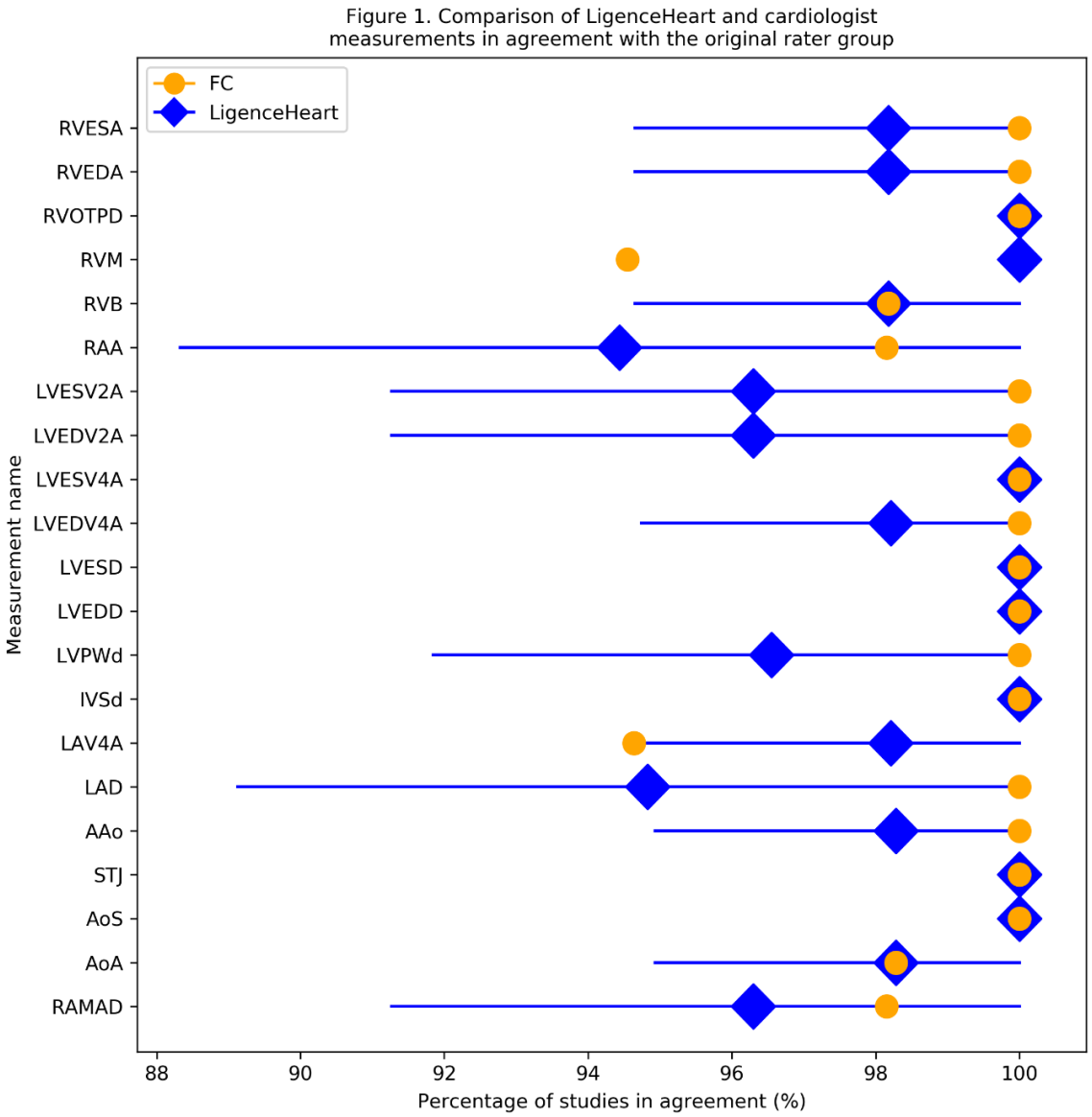

Figure 1. *The X axis represents the number of studies that were measured in agreement with ORG by Ligence Heart or FC. Each Y axis position represents a different measurement. 95% CIs are shown for Ligence Heart only. Ligence Heart 95% CI lower bound is at or above FC performance. ORG - original rater group. FC - fourth cardiologist.*