



# **Clinical Evaluation Report**

in accordance with MEDDEV 2.7/1 revision 4

and in compliance with

Council Directive 93/42/EEC as amended by directive 2007/47/EC

Council Directive 90/385/EEC as amended by directive 2007/47/EC

## ***Clinical Utility and Usability Study- CUUS***

Sponsor: Cardio-Phoenix Inc.  
44 Rosemead Close  
Markham, Ontario, L3R 3Z3  
Canada

**3 Center Version.**



## Table of Content

<b>1 SUMMARY -----</b>	<b>5</b>
<b>2 DEFINITION OF TERMS AND ACRONYMS-----</b>	<b>7</b>
<b>3 SCOPE OF THE CLINICAL EVALUATION -----</b>	<b>8</b>
3.1 DEVICE 1. CARDIO-TRITEST™ (CTT) v6.5 -----	8
3.2 DEVICE 2. CHART PROCESSING ALGORITHM (CPA) 2018A -----	8
3.3 PRIMARY OBJECTIVE -----	9
3.4 SECONDARY OBJECTIVES-----	9
3.5 SPONSOR -----	9
<b>4 IDENTIFICATION OF PERTINENT DATA -----</b>	<b>10</b>
4.1 DATA GENERATED AND HELD BY THE MANUFACTURER-----	10
4.1.1 <i>Identification of device(s)</i> -----	10
4.1.2 <i>Identification of manufacturer</i> -----	10
4.1.3 <i>Device description as part of CE System</i> -----	11
4.1.4 <i>Intended use/purpose</i> -----	14
4.1.5 <i>Status of the device</i> -----	19
4.1.6 <i>Clinical background</i> -----	19
4.2 DATA RETRIEVED FROM LITERATURE -----	21
4.2.1 <i>Applicable standards and guidelines</i> -----	21
4.2.2 <i>Governing directive</i> -----	22
4.2.3 <i>State of the art</i> -----	22
<b>5 APPRAISAL OF PERTINENT DATA -----</b>	<b>26</b>
5.1 ECG FINDINGS -----	26
5.2 PCG FINDINGS -----	30
5.3 MCG FINDINGS -----	30
5.4 HART™ FINDINGS -----	31
5.5 INCLUSION AND EXCLUSION CRITERIA -----	34
5.5.1 <i>Conditions which are processed in CUUS clinical study</i> -----	36
5.5.2 <i>Primary Endpoints</i> -----	38
5.5.3 <i>Secondary Endpoints</i> -----	40
5.6 POTENTIAL RISK HAZARDS/BENEFITS -----	41
<b>6 CLINICAL STUDY PHASES -----</b>	<b>42</b>
6.1 PHASE 1-----	42
6.1.1 <i>Step 1. Patient Attends Primary Care Clinic</i> -----	42
6.1.2 <i>Step 2. Heart Disease Risk Factors Assessment</i> -----	42
6.1.3 <i>Step 3. ECG &amp; CTT (CHART#1) Examinations</i> -----	43
6.1.4 <i>Step 4. General Practitioner (GP) evaluation</i> -----	43
6.1.5 <i>Step 5 - Over-Read</i> -----	43
6.1.6 <i>Step 6 - CHART#2 Examination</i> -----	43
6.1.7 <i>Step 7 - ECHO Examination</i> -----	44
6.2 PHASE 2-----	44



6.2.1	<i>Step 1 – Video Audit</i>	44
6.2.2	<i>Step 2 – Consensus Study</i>	44
6.3	PHASE 3	45
6.3.1	<i>Utility Study Component</i>	45
<b>7</b>	<b>CONDUCT OF THE APPRAISAL</b>	<b>45</b>
7.1	STUDY PARTICIPANTS	45
7.2	HYPOTHESIS	45
7.2.1	<i>Null- Hypothesis</i>	45
7.2.2	<i>Primary Hypothesis</i>	46
7.2.3	<i>CTT Primary Hypothesis</i>	46
7.2.4	<i>Secondary Hypothesis</i>	46
7.3	HYPOTHESIS CALCULATION	47
7.4	THE “SEND/DON’T SEND” DECISION	48
7.5	RANDOMIZATION, BIAS AND BLINDING	48
7.6	SAMPLE SIZE CALCULATION	50
7.7	ANALYSIS POPULATION	52
7.8	STATISTICAL METHODS	56
7.9	EXPECTED RESULTS	57
<b>8</b>	<b>ANALYSIS OF THE CLINICAL DATA</b>	<b>58</b>
8.1	THE CHALLENGE OF SE AND SP MEASUREMENT	58
8.2	PERFORMANCE MEASUREMENTS	60
8.2.1	<i>Referral decision</i>	60
8.2.2	<i>Confusion matrix</i>	61
8.2.3	<i>Performance metrices</i>	61
8.2.4	<i>Area Under Curve</i>	63
8.3	DECISION PERFORMANCE OF CHART OVER ECG	63
8.3.1	<i>GP decision compared to ground truth</i>	64
8.3.2	<i>GP decision compared to ORC and RC decision</i>	68
8.3.3	<i>ORC decision compared to ground truth</i>	72
8.3.4	<i>RC decision compared to ground truth</i>	75
8.4	DIAGNOSIS AND COMPARISON DATA	78
8.4.1	<i>GP and RC Diagnosis</i>	78
8.4.2	<i>Comparison ECG over CHART</i>	83
8.4.3	<i>CTT Primary Hypothesis evaluation</i>	85
8.4.4	<i>Comparison CHART#2 over CHART#2</i>	85
8.4.5	<i>CHART as start point for ECHO</i>	86
8.4.6	<i>Reproducibility of findings</i>	88
8.5	CONSENSUS BASED GROUND TRUTH	94
8.5.1	<i>Consensus Input Statistics</i>	94
8.5.2	<i>Referral Decision</i>	95
8.5.3	<i>ECHO Findings</i>	97
8.5.4	<i>Conclusion</i>	99
8.6	HART FINDINGS CLASSIFICATION PERFORMANCE	100
8.6.1	<i>Performance by ECHO findings</i>	100



8.6.2	<i>Performance by Referral decision</i>	102
8.6.3	<i>Conclusion</i>	104
8.7	USABILITY EVALUATION	104
8.7.1	<i>Group descriptive statistic</i>	105
8.8	CORRECTIONS AND DIAGNOSIS AND DECISIONS	110
8.8.1	<i>Dr. Brankica referred everybody</i>	110
8.8.2	<i>Dr. Daniela- switched from CHART to ECG protocol in the GP evaluation</i>	111
8.8.3	<i>Dr. Tatjana, Cardiologist- too many "No-Action" decisions in consensus</i>	113
8.8.4	<i>Hypothesis Acceptance/Rejection</i>	113
8.9	REQUIREMENT ON SAFETY	118
8.10	ACCEPTABLE BENEFIT/RISK	119
8.11	SUPPORTING DOCUMENTS	119
<b>9</b>	<b>CONCLUSIONS</b>	<b>120</b>
9.1	BENEFIT-RISK DETERMINATION	120
<b>10</b>	<b>DATE OF THE NEXT CLINICAL EVALUATION</b>	<b>122</b>
<b>11</b>	<b>SIGNATURES OF THE RESPONSIBLE EVALUATORS</b>	<b>123</b>
<b>12</b>	<b>QUALIFICATION OF THE RESPONSIBLE EVALUATORS</b>	<b>124</b>
<b>13</b>	<b>REFERENCES</b>	<b>124</b>



## 1 Summary

The Clinical Utility and Usability Study (CUUS) was a pivotal, multicentre (Sombor, Vrsac, and Senta), randomized, blinded study the goal of which was to determine utility of CHART, its usability for its intended use, in a clinical environment, by its intended users, in a study population representative of the target population.

This study was designed with the assistance of and approval by the FDA (the US Food and Drug Agency) as part of Q165xxxx. The study was designs to collect the data to confirm the safety and effectiveness of the CHART system, in particular the IMD- Cardio-TriTest v6.5 (CTT for short) and the CHART Processing Algorithm (CPA) 2018a.

The data collected would confirm the hypothesis that CHART analysis is more effective than ECG only analysis in assisting the GP in determining their referral decision (Send/Don't Send) and the basis for it.

550 patients were recruited into the CUUS. More than 500 patients' clinical results were evaluated to measure the diagnostic and decision support capability of CHART report compared to ECG report. This study population was representative of the target population according to the intended use of CHART. Some 43,0% of the patients are classified as Sent to referral cardiology based on the consensus-based ground truth.

The results confirm that in many ways CHART analysis is more effective than ECG only analysis. False-negative rates are significantly decreased (CHART produced a 15.8% decrease in False-positive) and False-Positive rates decreased by 5% (False-positives in the patient referral decision by GP as compared to ECG-only based decisions).

Furthermore, doctors (GP, ORC and RC) were significantly more sensitive in their referral decision when based on CHART report.

Reproducibility is better than the predicate ECG, and the Usability results shows that the system itself, like their devices separately, are easy to use, user-friendly and that there were no problems or additional risks, not previously understood and mitigated, established while working with the devices.

No adverse events (AE) or effects were reported.

The benefits of this study and its results are:

1. Better able to detect and confirm onset of heart disease earlier (when treatment options are more effective and cost-efficient).
2. No additional risk compared to predicate ECG devices. Device is as safe to use as Predicate ECG devices.
3. Increased effectiveness for a much wider range of disease conditions.
4. Easy to use, little or no additional operator training was required for existing operators of predicate devices.



5. Fits into current workflows for normal Standards of Care (SoC).
6. Is an effective assistant to Primary Care physicians in helping them to better understand their patient's cardiac status.
7. Reduction of FN by 15.8%.
8. Reduction of FP by 5%.
9. Widespread benefit to all patients attending Primary Care, or patient care clinics.
10. Helps Cardiologists with Collaborative triage of patient appointment priorities.
11. Helps Cardiologists identify a start point for Echo examination, saving time, and costs.

This study confirmed safe performance of the CHART system in a clinical setting as demonstrated by the results, and contributed to the benefit/risk assessment. Overall, the probable benefits outweigh the probable risks given the available information concerning the benefits and risks. There is reasonable assurance of the safety and effectiveness for this system for the intended use.



## 2 Definition of Terms and Acronyms

Term	Definition
ML	Machine Learning
AI	Artificial Intelligence
CE	Conformité Européene
CFR	Code of Federal Regulations
EU	European Community
FDA	Food and Drug Administration
CHART	Cardio-HART™
CTT	Cardio-TriTest hardware
CPA	CHART Processing Algorithm
GP	General Practitioner
ORC	Over-Read Cardiologist
RC	Referral Cardiologist
CUUS	Clinical Utility and Usability Study
MT	Medical Test
QMS	Quality Management System
CER	Clinical Evaluation Report
ECG	Electrocardiograph
PCG	Phonocardiograph
MCG	Mechanocardiograph
EMPCG	ECG+MCG+PCG bio-signals
PECG	Predicate ECG
ECHO	Echocardiogram
CI	Clinical Investigation
FP	False Positive
FN	False Negative
AE	Adverse Event
HART findings	These findings detected by CPA algorithm based on Combined Feature Vector including ECG, PCG and MCG medical information
TN	True Negative
TP	True Positive
NS	Not Sure
NSR%	Number of samples or answers
N	Number of samples or answers
SE%	Sensitivity (SE12% the binary case derived from triple)
SP%	Specificity (SP12% the binary case derived from triple)
PPV%	Positive Predictive Value (PPV12% the binary case derived from triple)
NPV%	Negative Predictive Value (NPV12% the binary case derived from triple)
ACC%	Accuracy (ACC12% the binary case derived from triple)
K%	Cohens Kappa
r%	Pearson correlation coefficient
AUC%	Area Under Curve
RCI%	Relative Classifier Information
PR%	Positive rate (by test)
PREV%	Prevalence (by reference)
Don't	Do not refer patient to next level of health care- GP/ORC/RC decision
Send	Refer patient to next level of health care- GP/ORC/RC decision
CI	Confidence interval



### 3 Scope of the clinical evaluation

The Scope of this Clinical Study is to conduct a multi-site clinical evaluation study to confirm the functional and operational aspects between two medical devices competing against each other in clinical practice to determine which provided the most beneficial clinical outcomes for patients. This study is not intended to be a performance study of CHART's findings, albeit device performance was verified and measured.

The Clinical Study is conducted in accordance with the ethical principles that have their origin in the Declaration of Helsinki.<sup>1</sup> These principles protect the rights, safety and well-being of human subjects, which are the most important considerations and shall prevail over interests of science and society. These principles are understood, observed and applied at every step in this Clinical Study.

This study was NOT designed as a performance study to validate CHART's findings. However, for accuracy of findings for each patient, the MT for each patient was verified and validated in accordance with its intended performance by a separate team of cardiologists to ensure an accurate ground truth.

#### 3.1 Device 1. Cardio-TriTest™ (CTT) v6.5

The CTT device is a medical device that captures 3 types of heart generated Bio-signals, including ECG, PCG and MCG signals of an electrical, acoustic and physiological (mechanical) nature emanating from the heart and non-invasively captured on the thoracic wall. Captured bio-signals are outputted as either a printable report of the measured signals or as electronic signals for analysis processing, tested and conformant with ISO 60601-2-25 for ECG.

The CTT device is for use in patient clinical practice, including Primary Care.

#### 3.2 Device 2. CHART Processing Algorithm (CPA) 2018a

CPA is an automated AI based processing system for the analysis of heart bio-signals, including ECG, PCG and MCG. It analyses those bio-signals and outputs diagnostic findings, including ECG, PCG and MCG findings. It also analyses the combined ECG, PCG and MCG signals and outputs them as HART™ findings.

The CPA outputs a report that is indicated for use by clinicians to assist them in diagnosing a patient's current cardiac status.

The CPA is not for use in patient clinical practice.

Furthermore, the extended scope of the Clinical Evaluation is to also confirm the deployability, operational usefulness and cost efficiencies within the healthcare system.

---

<sup>1</sup> <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>



### **3.3 Primary Objective**

Characterize the effectiveness and Utility of CHART system over State-of-the-Art ECG-only analysis in providing diagnostic assistance to Clinician's:

To prove that CHART automated analysis is more effective than ECG-only automated analysis in assisting Clinicians in better understanding a patient's cardiac status in clinical practice.

To show that the bio-signals produced by the CTT device remain reasonably constant and consistent in accordance with the standards, as applicable.

### **3.4 Secondary Objectives**

1. Determine if CHART analysis is better than ECG-only analysis in assisting the Cardiologist in better understanding the basis of the medical justification used for the referral and the prioritization of patients, from primary care to cardiology care;
2. Determine if CHART provides an effective starting point for ECHO examination<sup>2</sup>;
3. Determine whether CHART examinations conducted in Primary Care are consistently reproducible and repeatable in Cardiology Care;
4. Characterize the Usability of CHART in real-world clinical care practice.

This Clinical Evaluation report (CER) is in the support of initial CE-marking. It applies to the Cardio-TriTest v6.5 (CTT) and the CHART Processing Algorithm (CPA) 2018a, as used in the CHART system, that fall within the definition of medical devices under European Medical Devices Directive (MDD) 2007/47/EC- Council Directive 93/42/EEC of 14 June 1993.

### **3.5 Sponsor**

Cardio-Phoenix Inc.  
44 Rosemead Close  
Markham, Ontario, Canada, L3R 3Z3  
Phone: 416-595-0795  
[www.cardiophoenix.com](http://www.cardiophoenix.com)  
[info@cardiophoenix.com](mailto:info@cardiophoenix.com)

#### Serbian Representative:

UVA Research and Development d.o.o.  
Henrika Sjenkevica 14  
24000 Subotica, Serbia

---

<sup>2</sup> This objective came out of previous clinical studies where cardiologists commented that they were using CHART for establishing effective starting points for ECHO examinations



Context- The Clinical Study is designed to determine the functional and operational safety and effectiveness of two medical devices produced by Cardio-Phoenix operating as part of the CHART system in clinical practice.

Patients- This CUUS involved both male and female patients, distributed equally, with at least 3 risk factors for heart disease, attending a Primary care clinic for health reasons, including annual physical examination. Other information about the patients and inclusion/exclusion criteria are detailed in the Chapter 5.5.

Target population- The target sample size was between 400 and 500 patients of which at least 15% should be considered as healthy patients (no risk factors), ranging in age from 20 to 76 or above, grouped as follows: a) 20-40, b) 41-55, c) 56-65, d) 66-75 and e) 76+.

Single MT- Selected Patients were to undergo a single MT session at the Primary Care level, followed up by an MT session in Cardiology care, with no further follow-up, as the Clinical Study is diagnostic in nature.

Risk to Patients- The risk to patients was deemed by risk analysis and experience to be extremely low, no greater than that of a standard ECG examination. The risks to patients are well understood and derived from several prior clinical studies involving over 6.500 patients and 25.000+ tests, with non-unintended or adverse effects, or any effects whatsoever for this same non-invasive procedure.

## 4 Identification of pertinent data

### 4.1 Data generated and held by the manufacturer

#### 4.1.1 Identification of device(s)

Device(s)
<u>Device #1.</u> Name: "Cardio-TriTest" (or "CTT") v6.5", with software- the Cardio-Client v3.0
<u>Device #2.</u> Name: "CPA"- CHART Processing Algorithm ("CHART"- Cardio-HART™) 2018a

#### 4.1.2 Identification of manufacturer

Legal Manufacturer



Cardio-Phoenix Inc.  
44 Rosemead Close  
Markham, Ontario, Canada, L3R 3Z3  
Phone: 416-595-0795  
[www.cardiophoenix.com](http://www.cardiophoenix.com)  
[info@cardiophoenix.com](mailto:info@cardiophoenix.com)

#### **4.1.3 Device description as part of CE System**

During the Clinical Study (CS) the evaluation of each device was based on their combined Utility when used in the healthcare system, and validated with their comparator predicate ECG device and Echocardiograph. The system consisted of 3 main sub-systems as follows:

##### **4.1.3.1 Medical Device #1- The Cardio-TriTest™ or CTT device v6.5**

In Primary Care, (known as “Ambulance” in Serbia), the CTT was used by a trained ECG nurse to collect a patient’s heart bio-signals.

Cardio-TriTest™ or CTT, is a non-invasive, Class IIa (EUROPE, CLASS 2, FDA), hardware medical device, for use in patient clinical care situations, including Primary Care for the recoding and capture of 3 different types of heart bio-signals.

The device is designed to be used by a trained ECG medical assistant or nurse within the context of the current Standards of Care.

##### **4.1.3.2 Medical Device #2- The CHART Processing Algorithm or (CPA) 2018a**

The CPA, is standalone software medical device for the algorithmic analysis and processing of heart bio-signal such as those produced by the CTT device. It is able to diagnose 167 heart parameters and 64 heart findings and output a report for use by clinicians to assist them in better understanding their patient’s cardiac status.

##### **4.1.3.3 The CHART System & Purpose of the CE**

The combination of all modules and devices, is referred to as the CHART system. The Clinical Evaluation, conducted through a pivotal Clinical Study has a main goal to show the safe operation of each of the devices and confirm their performance based on their own validation requirements.

#### **Description in alignment with technical documentation**



### Cardio TriTest v6.5

The main goal of the device is to collect and transfer bio-electric, acoustic sounds and mechanically induced signals (analog and digital) from a subject patient. These signals include: Standard 12-lead ECG, 4 lead PCG and 4 lead MCG signals for a total of 24 signals. The device also collects 24 digital signals from secondary digital sensors.

The Device v6.5 is a piece of hardware that transforms analog ECG, PCG and MCG signals to 24-bit digital data and create Medical Tests for patients with software application. The device is also collecting data from digital sensors (in 16bit resolution). The signals from the digital sensors (DMCG signals) are only recorded at this time for signal comparison purposes only to confirm the viability of the signals. The data collected by the device is sent to the host PC for later processing, analysis, viewing, distribution and storage.

Cardio-Client 3.0 software application is part of the medical device itself to create medical tests. The Cardio-Client 3.0 software application requires - an "off-the-shelf" notebook or desktop PC running Windows 10, 7 or 8 (Windows 10 is preferred), a Cardio Tri-Test medical device, and .NET framework version 4 or newer. The software takes input from the Cardio Tri-Test medical device through a USB (type A) connection and from the user who controls the software. The software automatically checks the recorded medical test for quality, consistency and integrity. If the recording quality is not sufficiently high enough, the software will reject the recorded test and force the user to make another test.

The output of the software is an encrypted file saved to the designated storage. The file is a standard .zip file, containing the patient's data and the recorded signals. Cardio-Client 3.0 generates either a printable report or the file can be transmitted for further processing.

### **The CHART Processing Algorithm or (CPA) 2018a**

The CHART Processing Algorithm (CPA) is a medical device software designed to aid in the clinical assessment of adult cardiac status in patient care clinics, including primary care.

The CPA is a software medical device designed to be installed on the cloud and process and analyze previously captured cardiac bio-signals and algorithmically identify suspected cardiac events and parameters indicative of cardiac status and output them in a written case report. The report serves as an assessment aid in the diagnosis of cardiac status in clinical practice.



The CPA device was developed using Machine Learning technologies, provides algorithmic analysis of the cardiac bio-signals for findings suggestive of pre-specified clinical conditions - specifically, electrical, mechanical, hemodynamic and functional abnormalities as either defined by the reference standards or as learned through Machine Learning. The CPA then outputs those findings in a written case report.

There are two main components of the software device: (1) the CPA which runs as a cloud-based algorithm processing service; and (2) the Report Service which runs as a separate server-based PDF report generator service.

The CPA receives previously captured bio-signals from a valid source device and then algorithmically processes them. After the CPA completes the processing, the results are provided to the Report Service which incorporates them into a case report in the form of a PDF and which is then made available to the originator.

CPA works within existing Standard of Care workflows. After a patient examination has been performed, a copy of the examination is received and processed by CPA. It then performs identification and classification of objects consistent with either the defined reference standards or algorithmically learned to identify clinically relevant cardiac events or dysfunctions and provide findings related to cardiac status. The resulting findings are incorporated into a case report in the form of a PDF.

Indications of cardiac events or parameters above the reference thresholds that are consistent with heart disease (indicative of heart disease) should be referred to specialized levels of cardiac care for further investigation.

CPA results are not intended to be used on a stand-alone basis for clinical decision-making nor is it intended to rule out any particular disease condition or otherwise preclude further clinical assessment of cases.

CPA is intended to be used with 3 type of bio-signals (ECG, PCG and MCG) acquired from a validated device, including the company's own FDA cleared Cardio-TriTest v6.5. (CTT). Refer to the FDA approved label of the Cardio-TriTest v6.5 for relevant contraindications, warnings, and precautions.

Each CTT device is unique and has its own unique identifier (UDI) based on the serial number. Each clinic is provided with their own device and it is registered as part of the CHART system. Each examination includes the UDI, the firmware version, the ID code of the person that conducted the examination. When completed, the Report includes this information and adds the CPA version info (2018a) and date of processing. All examinations can be tracked back through its processing.

#### Reference to technical documentation



Annex A- Technical File Cardio-TriTest v6.5.

Annex B- Technical File CHART Processing Algorithm.

#### **4.1.4 Intended use/purpose**

##### **Intended use/purpose in alignment with Instructions for use (IFU)**

###### **Device Specific Intended Use**

Each device has their own intended use<sup>3</sup> as per their device CE submission according to MDD.<sup>4</sup>

For the Clinical Study, the Intended use is based on the combine intended uses of each device, to ensure full understanding of the clinical operational procedures of each device within the Clinical Study and how they complement each other.

###### **Intended use:**

The CHART system is intended to be used to assist General Practitioners (GP's) to better understand and determine the Cardiac State of their patients.

CHART system is intended to provide objective assessment and diagnostic interpretation support to clinicians in patient care situations to aid in the evaluation of cardiac status and prognosis.

###### **Indications for use:**

1. The CHART system is a post-processing diagnostic assessment and interpretation software algorithm for the qualitative and quantitative analysis of ECG, PCG, MCG (together, "EMPCG") bio-signals previously acquired from a compatible device such as the Cardio-TriTest™ v6.5.

2. The results of the CHART analysis are intended to provide an objective assessment and diagnostic interpretation of resting EMPCG bio-signals for use as an aid in diagnosing commonly recognized cardiac functions and dysfunctions (abnormalities) of a physiological or morphological nature, with mechanical, electrical and structural

<sup>3</sup> Annex C- Intended Use and Indications for Use

<sup>4</sup> European Medical Devices Directive (MDD) 2007/47/EC



characteristics in adult population, age 20+, in patient care situations including primary care settings.

3. CHART analysis is intended to provide screening and diagnostic indications for conditions that might require further confirmation at a higher level of care.

4. The CHART program makes use of the patient's age, gender, race, weight, height, waist diameter, BSA and BMI.

5. CHART is for use by qualified clinicians in conjunction with the patient's clinical history, symptoms, physiology and other diagnostic tests as might be available, as well as the clinician's professional judgment and no treatment or other therapies should be initiated based solely on the indications produced by the CHART Program.

6. For prescription use only.

7. Not intended for use in pediatrics, and for patients in the intensive care units.

8. CPA is not intended to diagnose:

- Infrequent cardiac diseases – excluded as they are rare or require specialist level diagnosis:
  - Pericardial Effusion, Pulmonic Insufficiency or Stenosis, Aortic Root Dilation, Ventricular Septal Defect, Atrial Septal Defect, Thrombus, Paravalvular leak, Ballooning syndrome, Aneurysm, Atherosclerosis, Restrictive Cardiomyopathy, Swinging Heart, Pericardial Cyst, Barlows Mitral Valve Disease;
  - Idioventricular Rhythm, Hypocalcemia, Left posterior fascicular block, Ventricular Pre-Excitation, Ventricular tachycardia, Wide-QRS Tachycardia, Early Repolarization, Second-degree AV block, Accelerated Idioventricular Rhythm, Accelerated junctional rhythm, Junctional rhythm, Complete AV block, AV dissociation, Ectopic atrial tachycardia, Sinoatrial Block, Digitalis toxicity or effect, Chronic pulmonary disease, Pulmonary embolism, Acute pericarditis, Left atrial conduction abnormality, Supraventricular Complex, Fusion Complex, Ventricular escape complex, Junctional escape complex, Ventricular Parasystole;
- Congenital Heart Defects – excluded:
  - Mitral Valve Prolapse (MVP), Congenitally Corrected Transposition, Dextrocardia, Patent ductus arteriosus, etc.
- Known heart interventions – excluded as they should already be known:
  - Artificial Heart Valve, Percutaneous Coronary Intervention (PCI, STENT), Coronary Artery Bypass Graft (CABG), and all other type of cardiac intervention;



- Non-cardiac diseases.

### **Intended Population:**

The CHART system makes use of the patient's age, gender, race, weight, height, waist diameter, BSA and BMI. Furthermore, the BMI has been distributed:

1. Underweight (below 18.5);
2. Normal weight (18.51-24.99);
3. Overweight (25-29.99);
4. Obese (30 and higher).

The clinical reference standards (CRS) used in the development of the algorithm are based on Gender and Age adjusted as per ASE recommendation for BMI, plus BSA, and in accordance with the Standards approved by AHA, CDC, ASE and AAFP.

The cardiac parameters that are measured are based on the available CRS and are universally applicable in nature.

### **Target Patient Population:**

1. Adults, 20 ≤ years of age.
2. Age, grouped:
  - i. 20-40 (at risk only)
  - ii. 41-55
  - iii. 56-65
  - iv. 66-75
  - v. 76+
3. Gender, males and females, approximately evenly distributed (~50/50).
4. Not currently suffering from severe medical condition.
5. Race any
6. Able to provide consent.
7. Risk assessment results- at least 3 identified heart related risks factors (as per the SoC)<sup>5</sup>
8. Healthy/Unhealthy split: 15%/85% based on @risk.

The Cautions, Contraindications and warnings are related to the CTT Cardio-TriTest™ v6.5.

<sup>5</sup> Knowledge of Implementation of Cardiovascular Risk Clinical Practice by General Practitioners and Specialists. Available from: <http://www.revespcardiol.org/en/knowledge-and-implementation-of-cardiovascular/articulo/13092248/>

**Cautions:**

**CAUTION:** Never modify the device. Otherwise there is a danger of electric shock and/or injury.

**CAUTION:** All modifications must be performed by qualified personnel exclusively in services assigned by the manufacturer in authorized centers only.

**CAUTION:** Do not operate the equipment with any cover(s) removed.

**CAUTION:** Do not use any external power supply unit! The device should connect only on the designated PC (via the provided USB3.0 connector) on which the SW application is installed.

**CAUTION:** The connectors on the front enclosure part of the device are only for connecting the sensors shipped together with the device. Do not connect any other type of equipment on these connectors.

**CAUTION:** For connecting the sensors onto the device use only the provided cables which are shipped with the device. Do not use any other type of cables.

**CAUTION:** Before placing sensors on patients, the sensor cables must first be disconnected from the device. Only after cable is re-connected into its color assigned sensor, can the sensor cable then be connected to its color designed entry port in the device (red to red, yellow to yellow, green to green, and blue to blue).

**CAUTION:**

- Place your device on a hard level surface.
- Keep your equipment away from radiators and heat sources.
- Ensure that nothing rests on your equipment's cables and that the cables are not located where they can be stepped on or tripped over.
- Keep the device away from extremely hot or cold temperatures to ensure that it is used within the specified operating range.
- Do not use the device in a wet environment, for example near a bath tub or a sink.
- To help prevent electric shock, plug the PC power cables into properly grounded electrical outlets.
- Do not modify cables or plugs.



- Before you clean the equipment, disconnect the USB3.0 cable from the PC USB3.0 socket. The device during the cleaning procedure should be un-powered.
- Do not spill food or liquids on equipment components, and never operate the equipment in a wet environment.

**CAUTION:** This device to sale by or on the order of a (Licensed healthcare practitioner). It is for prescription only use.

**CAUTION:** To avoid any interference with other devices, the equipment can be used only in presence of other electrical equipment compliant with IEC 60601-1 and IEC 60601-1-2 for EMI and ESD, where as possible.

**WARNING:** MR unsafe!

- Do not expose the device to a magnetic resonance (MR) environment.
- The device may present a risk of projectile injury due to the presence of ferromagnetic materials which can be attracted by the MR magnet core.
- Thermal injury and burns may occur due to the metal components of the device which can heat during MR scanning.
- The device may generate artifacts in the MR image.
- The device may not function properly due to the strong magnetic and radiofrequency fields generated by the MR scanner.

**CAUTION:** Device is not rated for use at high altitudes.

**CAUTION:** Device is rated IP43.

**CAUTION:** Use environment should be restricted to frequencies of 1GHz or less.

**CAUTION:** Use the configuration parameters of IT-network provided by the manufacturer.

**CAUTION:** Connection of the device to an IT-network that includes other equipment could result in previously unidentified risks to patients, operators or third parties. It is upon the responsible organization to identify, analyze, evaluate and control these risks.

**CAUTION:** Successive modification to the IT-network could introduce new risks which require additional analysis. Modifications to the IT-network implicate:

- Modifications in the IT-network configuration;
- Connection of additional items to the IT-network;
- Disconnecting items from the IT-network;
- Update of equipment connected to the IT-network;
- Upgrade of equipment connected to the IT-network.



#### 4.1.5 Status of the device

Current status of device
FDA cleared in December 2018, as <a href="#">K182790</a> CE-Marked Aug 2020 (updated).

#### 4.1.6 Clinical background

The design of the Clinical Study is derived from previous clinical studies and adjusted for confirmation of effectiveness, usability and utility of the system consisting of the two IMD's, as they are intended to be used in real clinical practice settings.

Several prior clinical studies were conducted to support the performance and safety of the IVD's. In particular (oldest to recent):

- 2006-2007- Buda 1 (Hungary)
- 2009-2011 – Buda 2 (Hungary)
- 2011-2013 – Buda 3 (Hungary)  
1254 CTT and Echo Examinations collected for External Independent Validation. (different medical team, different location), each ECHO validated by 2 cardiologist and 1 specialist.
- 2014-2015 – Senta 1 (Serbia)  
1647 CTT and Echo Examinations collected for Training
- 2016 – Senta 2 (Serbia)  
1000 CTT and Echo examinations collected for Validation
- 2017 - Consensus Study. (Serbia)  
2647 Echo Examinations reviewed to determine ground truth

More information of those clinical investigations will be shown in the next paragraphs.

##### 4.1.6.1 Buda 1: Littman PCG Training Database

Buda 1: is Littman PCG Database "LBp1548", with Echo, used for training purposes. It contains 1548 medical sample tests about patients who visited either the Budapest cardiology department or the Beke Client, located in Budapest, between 2006 and 2007. Other than the upper specified parameters, all examination conducted during this trial also includes information about manual phonocardiograph annotations (completed by 2 cardiologist and verified by 1 specialist).

##### 4.1.6.2 Buda 2: EPCG Training Database

Budapest 2: ABp Database "ABp391" is used for training purposes which contains 391 medical test samples on patients who visited the Budapest cardiology department



or Beke Clinic between 2009 and 2010. Other than the upper specified parameters, this trial also includes information about manual annotations related to Echocardiography, electrocardiography and phonocardiography.

#### **4.1.6.3 Buda 3: CTT+ECHO Independent Database**

Budapest 3: 60 second length synchronized 12-lead ECG + 4 sensor PCG + 4 sensor 3D MCG signals (CTT device V4, "B" protocol) and ECHO examination were collected by clinical investigation did in Budapest between year 2011-2013.

~900 CTT records were selected from the 1273 tests by manual and automatic quality evaluation.

This database is exclusively used for independent validation purposes of the HART findings, and external validation of those ECG and PCG findings which are available in this database.

#### **4.1.6.4 Senta 1- CTT+ECHO Training Database**

Each Medical Test consists of three 60 second length recordings consisting of three type of synchronized bio-signals including 12-lead ECG + 4-lead (sensor) PCG + 4 lead (sensor) 3D MCG signals (CTT device V5, "C" protocol) and one ECHO examination, one patient risk questionnaire. These were collected by clinical investigation completed in Senta between year 2014-2015.

The ECHO-measurements were completed by cardiologists Dr. Szabo Erzsebet, Dr. Szabo B. Aniko. The first-level audit of echo-measurements were completed by sonographers Szabo Aniko and Bene Toth Anita based on the recoded echo-videos, while second-level audit was completed by Istvan Kecskes, a signal processing expert.

The ECHO-findings were diagnosed by the cardiologists Dr. Szabo Erzsebet, Dr. Szabo B. Aniko, written in textual format. This forms the primary ECHO-findings, which were replaced by the result of Echo-findings consensus study.

This database is used only for training purpose of ECG-, PCG-, MCG-findings the HART findings.

This database was validated for Ground Truth in the 2017 Consensus Study.

#### **4.1.6.5 Senta 2- CTT+ECHO Validation Database**

The Senta 2 Validation Database Va1000, which contains 1000 medical tests about patients who visited the Senta cardiology in 2016.

Two consecutive Studies were completed in Senta 1 and Senta 2. Both were, pivotal, non-randomized paired diagnostic device clinical studies. The purpose was to collect sufficient data for both training and validation purposes. The study was a paired



study, with each patient getting both an Echo examination and a CTT test. The purpose was to use the Echocardiograph's ability to diagnose heart diseases not otherwise diagnosed by ECG and thereby validate the CTT results. An Echo Examination is the gold standard for non-invasive heart diagnostic assessment; however, it is usually provided beyond primary clinical practice at the higher level of care by a Cardiologist (or Medical Specialist).

- Senta 1, collected ~1640 records that was used only in the Training phase.
- Senta 2, collected ~1000 records that was used only in the Validation phase.

More information about the previous Clinical Investigations you can find in **CHART Database Validation** document attached as an **Annex I-2** in Technical File CHART Processing Algorithm.

#### **4.1.6.6 Consensus Study- Ground Truth ("GT")**

A Consensus Study (CS) was held to review the diagnostic accuracy of the Echocardiograph examinations conducted during the two Clinical Investigations completed in Senta Hospital (Serbia) in 2014-2016.

The goal of the Consensus study is to have minimum 4 independent cardiologists review each of the echo examinations conducted in the *Senta 1* and *Senta 2* CI and thereby establish the Ground Truth of the original diagnosis in accordance with the Clinical Reference Standard (CRS).

The Consensus Study established the Ground Truth of echo findings which will be then used to validate the various findings, in particular the HART findings, diagnosed by CHART Processing Algorithm (CPA)<sup>6</sup> algorithm.

More information about the **Consensus Study** you can find as **Annex H-Consensus Study** document in Technical File CHART Processing Algorithm.

## **4.2 Data retrieved from literature**

### **4.2.1 Applicable standards and guidelines**

Standard	Title
EN ISO 13485:2016	Medical devices — Quality management systems — Requirements for regulatory purposes
EN ISO 14971:2012	Medical devices — Application of risk management to Medical Devices

<sup>6</sup> CHART Processing Algorithm or CPA is an automated AI based system for the analysis of heart bio-signals, including ECG, PCG and MCG.



EN 62304:2006/A1 2015	Medical device software — Software life-cycle processes
EN 62366:2015	Medical devices — Application of usability engineering to medical devices
EN ISO 14155:2011	Clinical investigation of medical devices for human subjects — Good clinical practice (ISO 14155:2011)

- European Medical Devices Directive (MDD) 2007/47/EC
- Council Directive 93/42/EEC as amended by directive 2007/47/EC
- MEDDEV 2.7/1 revision 4

#### 4.2.2 Governing directive

Medical Device according to:
<ul style="list-style-type: none"><li>• European Medical Devices Directive (MDD) 2007/47/EC</li><li>• Council Directive 93/42/EEC as amended by directive MDD 2007/47/EC</li></ul>

#### 4.2.3 State of the art

The State of the Art (or “SOTA”) should be understood from the context of relevant regulatory frameworks, standards and defined Standards of Care as applicable in the use of the SOTA in patient clinical care situations, including Primary Care, the practical implications of which are found in the Standards of Care (SoC).

The current SOTA device used within the existing SoC <sup>7</sup>(GoffJr D.C., Lloyd-Jones D.M. et al., 2014) workflows, is overwhelmingly the Electrocardiograph or “ECG” device.

ECG is the primary device used for determining cardiac status, particularly in Primary Care. Various well established and well-defined standards, regulate it. However, ECG has limitations, which are all too often not understood, in particular by patients, but also by GP's in general (there are exceptions, they are not the rule).

As a result of these limitations, ECG was never adopted as “the” diagnostic device for the widespread screening of the general patient population in clinical practice, although it is often use in that capacity as a result of a lack of a viable alternative.

<sup>7</sup> GoffJr D.C., Lloyd-Jones D.M. et al. ACC/AHA Guideline on the Assessment of Cardiovascular Risk. 2014. Available from: <https://www.ahajournals.org/doi/full/10.1161/01.cir.0000437741.48606.98>



The limitations include the fact that many heart diseases cannot be diagnosed by ECG, over half, as many heart diseases simply do not have an electrical signature – making them undetectable by ECG. Albeit highly accurate, high specificity, for diseases that it can detect, the range of detectable diseases is limited. As a result, ECG is generally skewed to specificity rather than sensitivity, the purpose of which is to reduce the rate of False-positives (FP) in favor of higher diagnostic accuracy. However, this also means it is less useful for the early detection of a wider ranges of common heart disease onset, restricting its use in Primary Care.

In Primary Care higher sensitivity is more desirable as it leads to earlier detection of disease onset, albeit with a higher level of FP. The earlier disease onset can be detected, the earlier preventative and or pharmaceutical treatments options can be accessed by patients, this when they are still at their most effective and at their most cost-efficient.

The current SOTA for ECG means that over half of all heart disease conditions are “missed”, hence missed-diagnosed or False-Negatives (FN). This means that these patients will not be provided with timely access to appropriate heartcare options. It is a serious limitation to the use of the device within the primary SoC context.

It also means the current SOTA is focused on confirming the diagnosis of ECG detectable diseases, i.e. reducing False-Positives (FP) or mis-diagnosis as it is expected that any remaining FP's will be clarified through access to higher levels of heartcare, either by Echo, the most likely first use device in cardiology or other riskier and more costlier procedures.

The current SOTA is therefore best defined as a compromise that uses ECG to accurately diagnose a select number of diseases, that represent less than half of all common heart diseases by prevalence, ignoring the others - some is likely better than none. But not all GP's think this way, where not knowing about the missed diseases can lead to dangerous conclusions about heart health that is unjustified.

The main challenge in Primary Care, is that most patients believe, expect, “all” heart diseases are diagnosable in Primary Care, a patient’s entry point into the healthcare system – it’s what they “believed and expect”.

Few patients understand the limitations and the dangers they represent.

Furthermore, access to cardiology level of care requires a medically justified referral. The ECG is generally the source of this medical justification.

The challenge in Primary Care can be best understood as the result of two gaps. The first gap is related to the Physician’s limited knowledge of the heart. The second is a technological gap related to the limitation of the ECG device itself, (the current SOTA device limitation).

This second gap, results because ECG is blind to the heart’s physiology. In simpler words, the ECG bio-signal captures the electrical not physiological nature of the heart.



Missing is a bio-signal to complement and augment ECG with the physiological functioning of the heart.

Addressing the current SOTA deficiencies requires two complementary innovations.

First, a new bio-signal is needed that is indicative and representative of the heart's physiology. The CTT resolves this issue.

Second, is some way to assist Primary Care physicians to more safely understand the patient's cardiac status and help them determine the best treatment options for the patient.

The two innovations by Cardio-Phoenix Inc., each address and resolve one of the identified short-comings of the current SOTA. These solutions help bridge the two main gaps in Primary Care diagnostic effectiveness and safety. Together the two innovations, found in separate modules of the CHART system, each addresses a different shortcoming of the current SOTA and in so doing, help create the foundation of a new SOTA paradigm that will likely endure for many years to come.

The bio-signal gap, is bridged by the CTT device, which captures 3 bio-signals, ECG, PCG and a novel breakthrough bio-signal. This breakthrough bio-signal is novel because it is both indicative of, and representative of, the heart's physiology. It's literally the missing piece of the cardiac diagnostic puzzle.

The physician's knowledge gap, is bridged by the CPA device, which uses AI as a diagnostic assistant for the Physician during their assessment of a patient's cardiac status. AI assistance means not only is the physician's effectiveness greatly increased, but so is patient safety, helping to reduce FN and FP that lead to better patient outcomes.

Combined, the benefits extend far beyond the immediate patient outcomes to also help reduce costs at both the patient and at the healthcare system levels. A further benefit is access, where better medical justification ensures access to the next level of heartcare, cardiology.

The CUUS shows the tangible benefits of higher diagnostic effectiveness and safety of the devices in their own right and as part of the SoC, but also from a cost efficiency perspective, that makes the system more accessible to all.

### **Other SOTA (State of the art)**

Echocardiography ("Echo") is the only other valid SOTA technology today that is "non-invasive". Although CT is arguably deemed non-invasive as well, it still poses a radiological risk that is all too often downplayed because there is literally, well, nothing else. Echo is commonly used for the diagnosis of heart disease but its use is restricted to only cardiology levels of care.

Echocardiography is not permitted for use in Primary Care for two main reasons: first, it requires a lot training and years of experience to understand the images; and



secondly, it requires extensive knowledge of the functioning heart to understand those images – the knowledge gap!

As such, the SOTA must be understood both within the context of use, the SoC in Primary Care, and within the regulatory framework that determines its use. Sadly, ECG is currently the only legacy device permitted to be used in that context from a usable and regulatory perspective in Primary Care.

Therefore, as ECG is well understood and is framed by regulatory context there is little to no purpose in further expanding the ECG's technical merits within the SOTA. ECG is the SOTA and has been for nearly 100 years. And the limitations are well known to the medical community. Changing of the SOTA means addressing these shortcomings, including the regulatory framework in which it must operate.

### **Technological SOTA**

When considered in a purely technological framework, the SOTA for ECG can be described as stagnant. Although many academic papers are written with respect to ECG, they increasingly focus on specific diseases and disease conditions of an increasingly rare nature that are still bound by the ECG signal. Simply, there is no improvement in the ECG signals in the last 120 or so years, the technology is so mature that it can be said, there is simply nothing left to improve.

Most technological improvements described in the literature and promoted by the competitive giants of ECG, are increasingly unrelated to ECG itself, but rather on how ECG signals are packaged for inclusion into EHR (Electronic Health Records) or distributed. For this, there is much talk of HL7 and Dicom formatting and importing as “essential” value-adds for ECG. This is the equivalent of putting lipstick on a pig - it changes nothing of ECG itself.

One exception is in the space of automated diagnosis of ECG signals. The move to algorithm-based diagnostic support continues to evolve, albeit slowly. The leading algorithms include the Glasgow Algorithm from the University of Glasgow written by Peter MacFarland, considered one of the world's foremost experts on ECG, which ECG algorithms is commonly found in many 3<sup>rd</sup> party ECG devices. Other large medical device manufacturers also have their own, including Veritas by Welch-Allyn, GE, and Siemens. They all basically do the same, finding the same diseases with the same accuracy, splitting hairs as differences. These algorithms are generally arithmetically based and are considered extremely stable – a good thing in medical circles.

Although AI has risen to prominence, the rush to AI based ECG algorithms is more hype than reality, and has mainly been the preserve of the academic world – in a quest to show academic performance, not diagnostic accuracy. Unsurprisingly however these AI based algorithms do not challenge the arithmetically derived algorithms in performance in large part because they are slower than their established counterparts.



Simply, at their very best, they can only match the existing technology. Simply, for the established players, there is no incentive to change, if it works - don't fix it!

Furthermore, to change it would also require the trip through regulatory hell, which as any serious medical device manufacturer will tell you, is no picnic, the regulatory pathway fraught with high costs and risks that do not justify the effort in light of what already exists and really works – wheels have been shown to run even if made square, but round wheels will always dominate. In practical terms as well, AI is much slower than arithmetically derived and adds little, if nothing that is not already there.

Still, the marketing “sirens-song” that is the lure of the promise of AI, has seen the competitive landscape, in particular from the less established players and startups, to venture into AI. But with most diseases already detectable by ECG it is generally considered fool’s gold, with most efforts focused in limited spaces such as in the field of arrhythmia detection, where the ability of AI to automatically distinguish arrhythmia from noise, could be beneficial for automated detection system.

ECG Innovation therefore has generally stagnated as simply there was little if anything new to discover in the ECG bio-signals.

The SOTA will therefore move to the detection of super rare diseases or.... be combined with other inputs that can expand, augment and complement ECG's effectiveness and likely re-invent first-use cardiac devices for use in Primary Care.

## 5 Appraisal of pertinent data

The CUUS is a pivotal clinical study. Medical conditions that are relevant to our devices/system will be shown separately by the findings in the chapters below.

CHART findings are used as a common reference name for all medical findings classified by CHART diagnostic device, including:

- ECG findings
- PCG findings
- MCG findings
- HART™ findings (defined as combined findings).

### 5.1 ECG findings

The required ECG findings are met with the standard which provides sufficient information about the ECG findings and list of diagnostic terms for ECG interpretation.

<b>CHART ECG Findings</b>	<b>Value domain</b>	<b>AHA/ACC/HRS Statement</b>	<b>Comment</b>
<b>1. Rhythm</b>	SR	20 Sinus rhythm	
	ST	21 Sinus tachycardia	



	SB	22 Sinus bradycardia	
	SARR	23 Sinus arrhythmia	
	AFib	50 Atrial fibrillation	
	AFlut	51 Atrial flutter	
	SVR	40 Supraventricular rhythm 55 Supraventricular tachycardia	
	Other Arrhythmias	34 Ectopic atrial rhythm 37 Junctional escape complex(es) 38 Junctional rhythm 39 Accelerated junctional rhythm 52 Ectopic atrial tachycardia 54 Junctional tachycardia 61 Fusion complex(es) 62 Ventricular escape complex(es) 63 Idioventricular rhythm 64 Accelerated idioventricular rhythm 70 Ventricular tachycardia 74 Ventricular fibrillation 76 Wide-QRS Tachycardia	Rare rhythm types in CHART intended population
	No		otherwise
	Yes	180 Atrial-paced complex(es) or rhythm 181 Ventricular-paced complex(es) or rhythm 184 AV dual-paced complex(es) or rhythm	
	No		otherwise
	Yes	60 Ventricular premature complex(es)	
2. Pacemaker	No		otherwise
	Yes	180 Atrial-paced complex(es) or rhythm 181 Ventricular-paced complex(es) or rhythm 184 AV dual-paced complex(es) or rhythm	
3. Premature Ventricular Complex	No		otherwise
	Yes	60 Ventricular premature complex(es)	
4. Premature Atrial Complex	No		otherwise
	Yes	30 Atrial premature complex(es)	
5. Heart Axis Deviation	Normal		otherwise
	LAX	121 Left-axis Deviation	
	RAX	120 Right-axis Deviation	
	IAX	123 Indeterminate Axis	
6. Poor R-wave progression	No		otherwise
	Yes	128 Poor R-wave progression	
7. PR interval	Normal		otherwise
	LongPR	82 Prolonged PR interval	
	ShortPR	80 Short PR interval	
8. B. Branch Block	No		otherwise
	LBBB	104 Left bundle-branch block	



	ILBBB	103 Incomplete Left bundle-branch block	
	RBBB	106 Right bundle-branch block	
	IRBBB	105 Incomplete right bundle-branch block	
	IVCD	107 Intraventricular conduction delay	
	No		otherwise
<b>9. Other Block</b>	LAFB	101 Left anterior fascicular block	
	Other	102 Left posterior fascicular block	
		108 Ventricular preexcitation	
		24 Sinoatrial block, type I	
		25 Sinoatrial block, type II	
		81 AV Conduction Ratio N:D	
		83 Second-degree AV block, Mobitz I	
		84 Second-degree AV block, Mobitz II	
		85 2:1 AV block	
<b>10. Myocardial Infarction (ECG crit.)</b>	No		otherwise
	IMI	161 Inferior MI 162 Posterior MI	
	AMI	160 Anterior MI 166 Extensive Anterior MI	
	LMI	163 Lateral MI Anterolateral MI	
	ASMI	165 Anteroseptal MI	
	SMI	Septal MI	
	UMI	Undefined MI	
<b>11. Ischemia</b>	No		otherwise
	Yes	220 Acute ischemia 226 Ischemia (205 Digitalis effect, 208 Hyperkalemia)	
<b>12. ST deviation</b>	No		otherwise
	STdev	145 ST deviation 501 ST Elevation 502 ST Depression	



	STTdev	146 ST deviation with T-wave change	
<b>13. T-wave Abnormality</b>	No		otherwise
	Yes	147 T-wave Abnormality	
<b>14. QT interval</b>	Normal		otherwise
	Long	148 Prolonged QT interval	
	Short	149 Short QT interval	
<b>15. Ventricular Hypertrophy</b>	No		otherwise
	LVH	142 Left ventricular hypertrophy	
	RVH	143 Right ventricular hypertrophy	
	BVH	144 Biventricular hypertrophy	
<b>16. Atrial Enlargement</b>	No		otherwise
	LAE	140 Left Atrial Enlargement	
	RAE	141 Right Atrial Enlargement	
	BAE	Biatrial Enlargement (140 and 141)	
<b>17. ECG Quality</b>	Good		otherwise
	Poor	12 Missing lead(s) 14 Artifact 15 Poor-quality data	
	Error	10 Extremity electrode reversal 11 Misplaced precordial electrode(s) 4 Uninterpretable ECG	
<b>18. ECG Summary</b>	Normal ECG	1 Normal ECG 2 Otherwise normal ECG 228 Normal Variant	
	Borderline ECG	3 Abnormal ECG 301 Borderline	
	Abnormal ECG	3 Abnormal ECG	
	Uninterpretable	4 Uninterpretable ECG	



## 5.2 PCG findings

PCG findings includes two main categories:

- a) heart sound and murmur existence
- b) systolic time interval classification.

<b>Group</b>	<b>Derived/ classified from</b>	<b>Findings</b>	<b>Value domain</b>
<b>Sound findings</b>	Measurements by time-frequency representation, threshold is adjusted automatically by help of machine learning	1. S1 Intensity (S1int)	Normal/ Increased/ Decreased
		2. S2 Intensity (S2int)	
		3. Ejection Sound (ES)	
		4. Midsystolic Click (MC)	
		5. Opening Snap (OS)	
		6. Third Sound (S3)	
		7. Forth Sound (S4)	Absence/ Presence
		8. Diastolic Murmur (DM)	
		9. Wheeze (WHEE)	
		10. Artifacts (ARTF)	
		11. Systolic Murmur (SM)	Early/Mid/ Late/Holo
<b>Systolic time interval (STI) findings</b>	<i>STI measurements, what derived from PCG segmentation, threshold adjusted by literature and appropriate database</i>	12. S1 Splitting (S1sp) 13. S2 Splitting (S2sp) 14. Electro-Mech. Activation Time (EMAT) or Q-S1 Interval 15. Systolic Performance Index (SPI) 16. Pre-Ejection Period (PEP), 17. Left Ventricular Ejection Time (LVET)	Normal/ Abnormal
<b>Summary</b>	<i>PCG quality check algorithm</i>	18. PCG Signal Quality (PCGq)	Good/ Poor/ Error
	<i>Knowledge based rules applied on findings</i>	19. PCG Summary	Normal/ Abnormal/ Uninterpretable

## 5.3 MCG findings

Cardiac heart events- such as aortic/mitral opening and closure- detection by MCG signal enables to calculate the systolic time intervals (STI). The classification of these intervals forms the MCG findings.



<i>MCG-finding</i>	<i>Value Domain</i>
<b>1. Electromechanical Activation Time (EMAT)</b>	
<b>2. Myocardial Performance Index (MPI)</b>	
<b>3. Systolic Performance Index (SPI)</b>	
<b>4. Pre-ejection Period (PEP)</b>	Normal/ Abnormal
<b>5. Left Ventricular Ejection Time (LVET)</b>	
<b>6. Isovolumetric Contraction Time (IVCT)</b>	
<b>7. Left Ventricular Filling Time (LVFT)</b>	
<b>8. Isovolumetric Relaxation Time (IVRT)</b>	
<b>9. Rapid Ventricular Filling Time (RVFT)</b>	
<b>10. MCG Signal Quality (MCGq)</b>	Good/Poor/Error
<b>11. MCG Summary</b>	Normal/ Abnormal/Uninterpretable

## 5.4 HART™ findings

The CHART™ Algorithm requirements are derived from the intended use, which basically comes from the original idea and further driven as a result of ongoing research over many years.

The original idea was to create a new generation ECG-like heart diagnostic device extending and complementing ECG interpretation with PCG and MCG bio-signal interpretations, as an aid for diagnosing a broader range of cardiac states in adult populations in patient care situations including primary care settings. Such a bio-signal based diagnostic medical device would easily lend itself to fitting with little to no change into existing Clinical Workflow Practices and Standards of Care.

ECG being limited to detecting electrically generated bio-signals characteristics of cardiac conditions cannot by its very nature detect many very common cardiac functions and dysfunctions, for example related to heart valves and structure, which are otherwise possible with Echocardiography. However, Echo is a cardiologist level diagnostic tool that because of the level of expertise required, is not suitable for primary care settings and even less for general patient screening. Whereas, a simple Bio-Signal based medical device would have no such restrictive limitations and would instead easily integrate, with little or no disruption, into existing Clinical Workflows and Standards of Care.

Confirmed by previous clinical studies, described in **section 4.1.6**, the following common heart disease groups can be predicted by the CHART algorithm using the EMPCG bio-signals in addition to the regular, and FDA cleared<sup>8</sup>, ECG findings:

### 1. Heart Functional Abnormalities

<sup>8</sup> FDA has cleared several predicate devices that are based on a use the Glasgow algorithm for ECG Findings. CHART uses this same Glasgow Algorithm in exactly the same way for ECG findings.



- a. Left Ventricular Systolic Dysfunctions
  - b. Left Ventricular Diastolic Dysfunctions
  - c. Left Ventricular Wall Motion Abnormality (ischemia)
2. Heart Size Abnormalities
- a. Dilated Cardiomyopathy
  - b. Left or Right Atrial Enlargement
  - c. Left Ventricular Hypertrophy
  - d. Right Ventricular Enlargement
3. Heart Valve Diseases
- a. Aortic Valve Stenosis and Insufficiency
  - b. Mitral Valve Stenosis and Insufficiency
  - c. Tricuspid Valve Insufficiency
  - d. Pulmonic Hypertension

The HART™ Findings are clinically validated by ECHO-findings. These ECHO-Findings are validated by medical consensus<sup>9</sup> using standard by ASE<sup>10</sup>.

<b>Abbreviation</b>	<b>HART Findings</b>	<b>Validated by ECHO-finding(s)</b>
<b>1. LVH</b>	Left Ventricular Hypertrophy	Left Ventricular Hypertrophy
<b>2. DCM</b>	Dilated Cardiomyopathy	Dilated Cardiomyopathy
<b>3. LAE</b>	Left Atrial Enlargement	Left Atrial Enlargement
<b>4. RAE</b>	Right Atrial Enlargement	Right Atrial Enlargement
<b>5. RVE</b>	Right Ventricular Enlargement	Right Ventricular Enlargement
<b>6. LVSD</b>	LV Systolic Dysfunction	LV Systolic Dysfunction
<b>7. LVDD</b>	LV Diastolic Dysfunction	LV Diastolic Dysfunction
<b>8. WMA</b>	LV Wall Motion Abnormality	LV Wall Motion Abnormality Ischemic Cardiomyopathy
<b>9. AR</b>	Aortic Valve Insufficiency	Aortic Valve Insufficiency
<b>10. AS</b>	Aortic Stenosis	Aortic Stenosis
<b>11. MR</b>	Mitral Valve Insufficiency	Mitral Valve Insufficiency

<sup>9</sup> A Consensus study determined the validity of the Echo Diagnosis taken during the various Comparative Clinical Studies (see Echo Consensus Study for full details).

<sup>10</sup> Recommendations for Cardiac Chamber Quantification by Echocardiography in Adults: An Update from the American Society of Echocardiography and the European Association of Cardiovascular Imaging, 2015



<b>12. MS</b>	Mitral Valve Stenosis	Mitral Valve Stenosis Combined Mitral Defect (Mitral Vitium)
<b>13. TR</b>	Tricuspid Valve Insufficiency	Tricuspid Valve Insufficiency
<b>14. PH</b>	Pulmonary Hypertension	Pulmonary Hypertension



## 5.5 Inclusion and exclusion criteria

Clinical Utility and Usability Study (CUUS) is a multi-centre, randomised, blinded, pivotal study. The study was conducted in Serbia, in three centers: Sombor, Vrsac, and Senta.

The clinical study is approved by the regulatory agency- Medicines and Medical Devices Agency of Serbia (ALIMS) in accordance with the Ministry of Health in Serbia. Ethical Committee approvals from both, hospital and the ambulances are provided.

### **Inclusion criteria:**

- a)** Adults, 20≤ years of age
- b)** Age grouped:
  - i. 20-40
  - ii. 41-55
  - iii. 56-65
  - iv. 66-75
  - v. 76+
- c)** Gender, males, and females, approximately evenly distributed (~50/50)
- d)** BMI, categorized (each category should include at least 12 patients):
  - i. Underweight (below 18.5)
  - ii. Normal weight (18.51 – 24.99)
  - iii. Overweight (25 – 29.99)
  - iv. Obese (30 & higher)
- e)** Not currently suffering from severe medical condition
- f)** Race any
- g)** Able to provide consent
- h)** Risk assessment results- at least 3 identified hearth related risks factors (as per the SoC)

### **Other requirement criteria:**

- a)** Patient population size: ~ 500
- b)** Healthy/Unhealthy split: 15%/85% based on @risk
- c)** Cardiology Clinics: Min 2
- d)** Primary Care Clinics: Min 2
- e)** Number of cardiologists: Min 4
- f)** Number of family physicians (General Practitioner- GP): Min 4

### **Exclusion criteria:**

- a)** Age, Under 20
- b)** Persons already diagnosed for heart disease and undergoing treatment in a cardiology ward



- c) Patients who have suffered and are currently undergoing pharmacological treatments for any heart disease in cardiology ward
- d) Patients who are suffering from any major illness or undergoing treatment for any disease that could influence their heart condition
- e) Lactating and pregnant women
- f) History of heart attack within last 120 days
- g) History of Stroke within last 120 days
- h) Presence of active, uncontrolled infection
- i) Any psychiatric disease/disorder, irreversible cognitive dysfunction or psychological issues likely to impair compliance with study protocol
- j) History of organ transplant
- k) Participation in any other study that can confound the study results or affect the study
- l) Refusal to participate
- m) Any condition that could limit survival to less than 1 week.



### 5.5.1 Conditions which are processed in CUUS clinical study

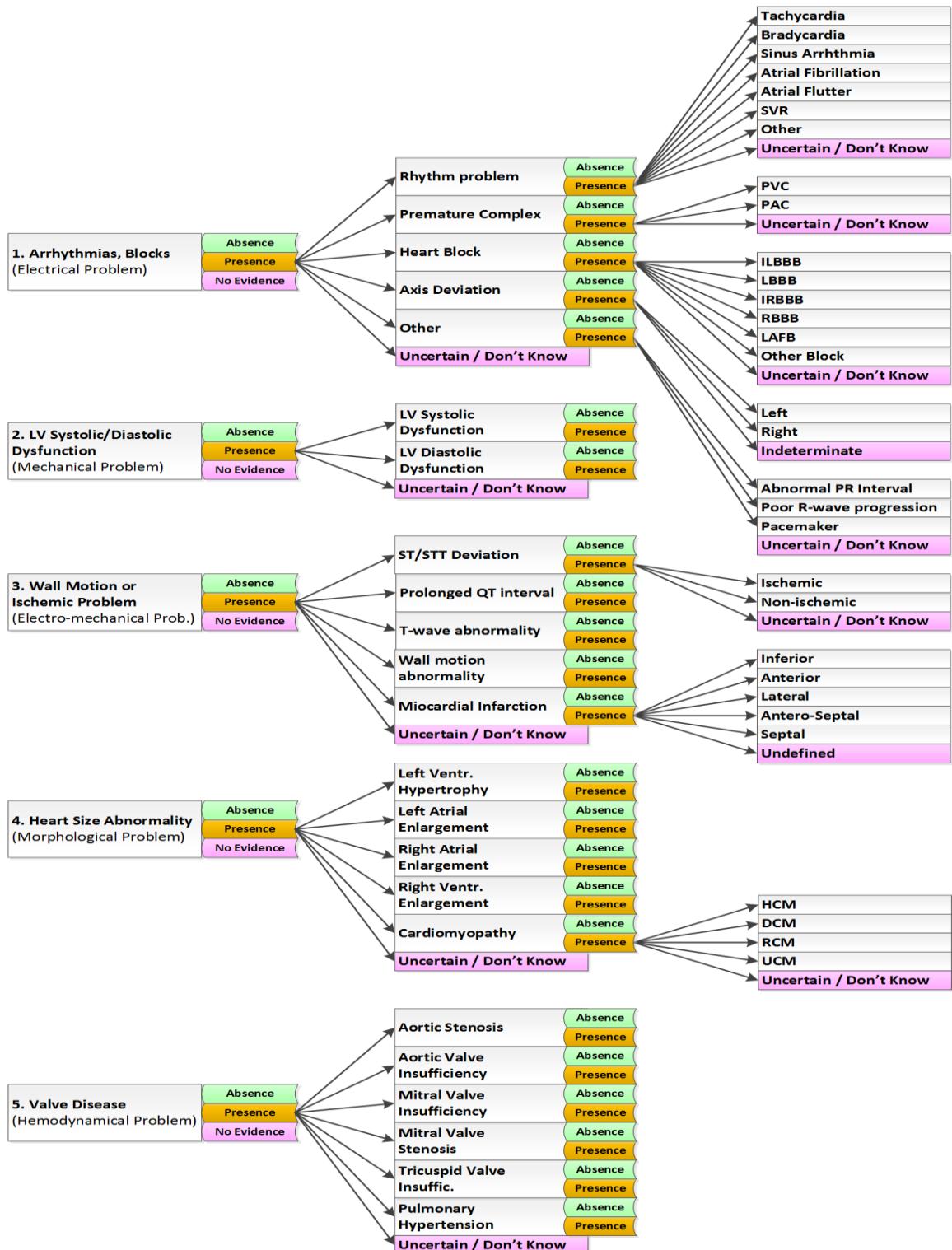


Figure 1- ECG Conditions

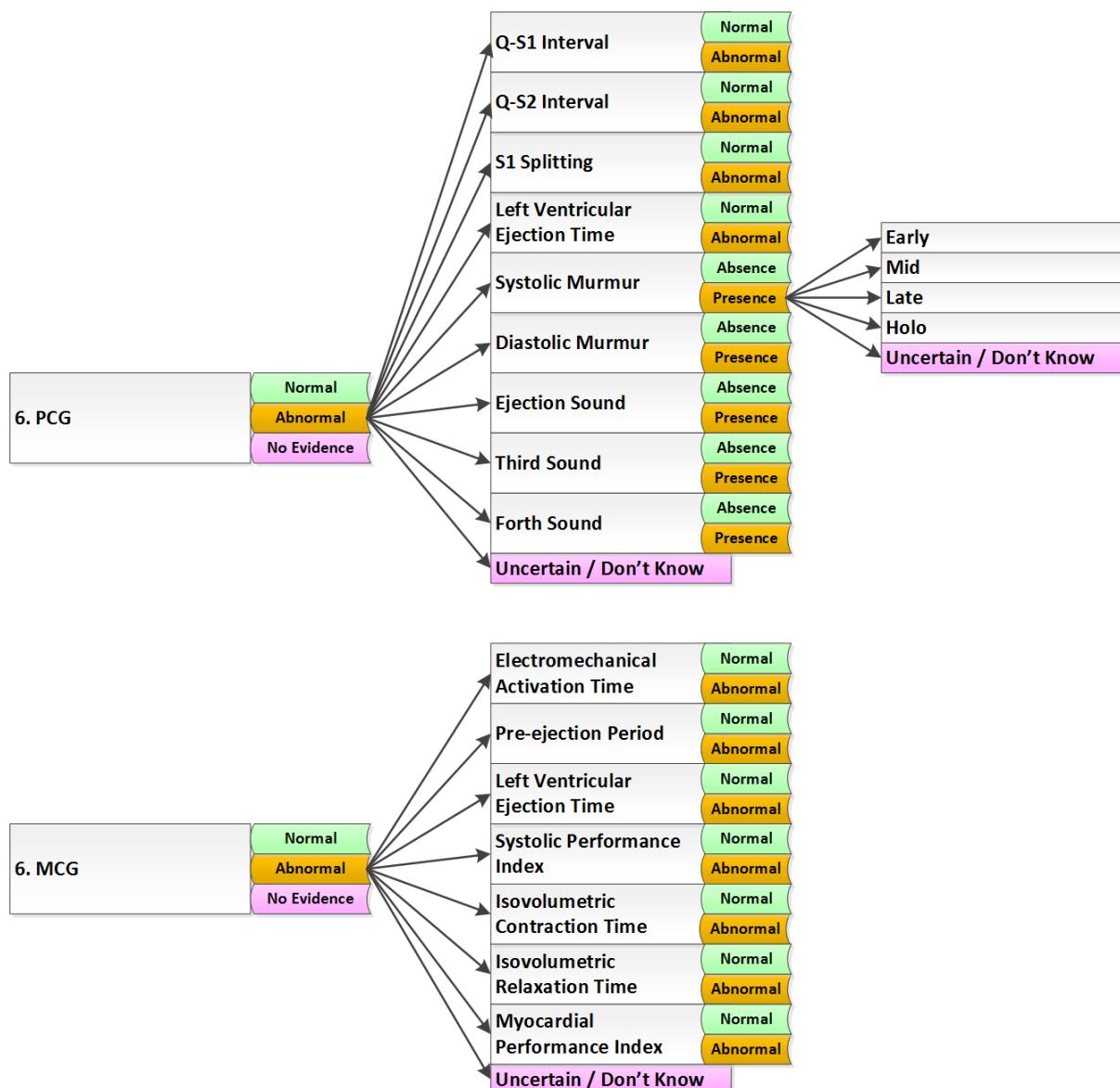


Figure 2- PCG/MCG Conditions

Based on the intended use of the CUUS clinical study, and the medical conditions shown above, the Endpoint established in this clinical study are:



### 5.5.2 Primary Endpoints

Compare CHART analysis to ECG-only analysis to determine which provides Clinician's with better diagnosis decision support, measured as the reduction of false positives, false negatives and 'no-evidence/not sure" rate.

The referral tree is following.

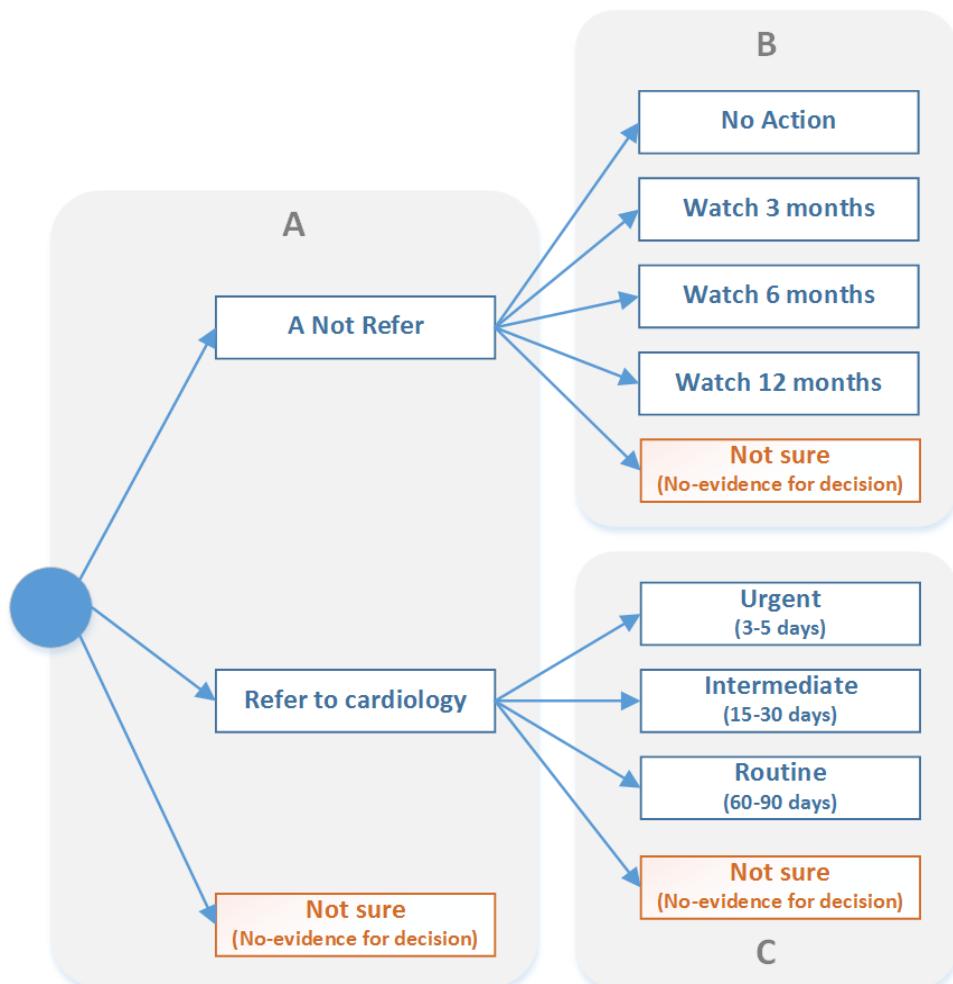


Figure 3- Referral tree



The confusion matrix "A" related to first level decision:

GP decision "A"	Consensus decision (Ground truth)	
	Not refer (negative - N)	Refer to cardiologist (positive - P)
<b>Not refer (negative A)</b>	True Negative (TN <sub>A</sub> )	False Negative (FN <sub>A</sub> )
<b>Refer to cardiologist (positive A)</b>	False Positive (FP <sub>A</sub> )	True Positive (TP <sub>A</sub> )
<b>Not-sure / No-evidence (neutral A)</b>	Neutral Negative (EN <sub>A</sub> )	Neutral Positive (EP <sub>A</sub> )

The confusion matrix "B" related to second level decision:

GP decision "B" after "Not refer"	Consensus decision (Ground truth)	
	Not refer (negative - N)	Refer to cardiologist (positive - P)
<b>No Action</b>		
<b>Watch 3 months</b>	True Negative (TN <sub>B</sub> )	False Negative (FN <sub>B</sub> )
<b>Watch 6 months</b>		
<b>Watch 12 months</b>		
<b>Not-sure / No-evidence (neutral B)</b>	Neutral Negative (EN <sub>B</sub> )	Neutral Positive (EP <sub>B</sub> )



The confusion matrix "C" related to second level decision:

GP decision "C" after "Refer to cardiology"	Consensus decision (Ground truth)	
	Not refer (negative - N)	Refer to cardiologist (positive - P)
<b>Urgent</b>	False Positive (FPC)	
<b>Intermediate</b>		True Positive (TPC)
<b>Routine</b>		
<b>Not-sure / No-evidence (neutral C)</b>	Neutral Negative (ENC)	Neutral Positive (EPc)

The endpoints are the change of following metrics between ECG and CHART based decision:

- 1) False positive rate of first level - A FPR<sub>A</sub> = FPA/(FP<sub>A</sub>+TNA)
- 2) False negative rate of first level - A FN<sub>A</sub> = FNA/(FNA+TP<sub>A</sub>)
- 3) No-evidence rate of first level - A NER<sub>A</sub> = (EN<sub>A</sub>+EP<sub>A</sub>)/(N+P)
- 4) No-evidence rate of second level - B, NER<sub>B</sub> = (EN<sub>A</sub>+EP<sub>A</sub>)/(N+P)
- 5) No-evidence rate of second level - C, NER<sub>C</sub> = (EN<sub>C</sub>+EP<sub>C</sub>)/(N+P)

The change can be expressed as:

$$\Delta P = P_{CHART} - P_{ECG}$$

where P can be FPR<sub>A</sub>, FN<sub>A</sub>, NER<sub>A</sub>, NER<sub>B</sub>, NER<sub>C</sub>

### 5.5.3 Secondary Endpoints

- 1) Compare the Cardiologists answers to the Clinician and the ORC (optional) about prioritization and medical justification between CHART and ECG-only, with 95% confidence interval.
- 2) Confirm the effectiveness of CHART in terms of indicating a start point for ECHO examination by comparing the CHART report and ECHO report and confirm it with the cardiologist statement.
- 3) Compare the CHART#1 and CHART#2 medical tests for repeatability and reproducibility and confirm with summary of Cardiologists comments.
- 4) Confirm from data provided by MA that the usability of the CHART system is easy, understandable and safe, with confidence level 95% internal.



## 5.6 Potential Risk hazards/benefits

The clinical evaluation is expected to address the significance of any clinical risks that remain after design risk mitigation strategies have been employed by the manufacturer.

1) Anticipated clinical benefits

Better decision support to Clinician's in general practice, through better understanding of the patient's cardiac status and ability to detect a broader range of common heart diseases.

Better cardiologist support in specialized clinical practice.

2) Anticipated adverse device effects

There are no Anticipated adverse device effects, as proven during several large previously conducted clinical studies.

3) Residual risks associated with the IMD, as identified in the risk analysis report

See **Annex D-** Risk Assessment for CUUS Clinical Study.

Risk associated with participation in the CI

No known risks are associated with participation in the CI.



## 6 Clinical Study Phases

The Clinical Study protocol is divided into 3 phases. Each phase is then divided into specific steps as required. Each Step represents certain actions that must be taken by specific actors in that Step. Included within each Step, are the various objectives and risks for each.

The users of the IMD's are the persons who are previously trained to be an ECG medical assistant. They were undergone additional CTT training which is typically less than 15 minutes as the IMD's was designed for optimum usability by current users of ECG devices.

### 6.1 Phase 1

#### 6.1.1 Step 1. Patient Attends Primary Care Clinic

- 1) Patient attends Primary Care Clinic (family Clinic) for whatever reason, either for a routine physical or with symptoms of heart ailments of whatever kind.
- 2) Patient is qualified by GP as potential candidate for clinical study based on selection criteria, and inclusion/exclusion criteria which are detailedly described in Chapter 5.5.
- 3) If patient does not qualify or does not wish to participate, then end.
- 4) If they do qualify and also wish to participate then MA starts the welcoming protocol starting with advising them of the clinical study, Helsinki Declaration protocol & informed consent, enters them in CMS (Client Management Software), then goes to next Step.

#### 6.1.2 Step 2. Heart Disease Risk Factors Assessment

General Physician completes an assessment of their patient's heart disease risk factors as per the Standard of Care (SoC).

- a. Is patient at risk?

- i. No, End.

Note: A small number of randomly selected patients, according to Age and Gender criteria, went through entire clinical study as known healthy patients (see Chapter 5.5).

- ii. Yes, go to next Step.



### **6.1.3 Step 3. ECG & CTT (CHART#1) Examinations**

- 1) MA (Medical Assistant) first conducts the Cardio-RisQ questionnaire. (22 questions)- The patient was considered as suitable for Clinical Study if at least 3 risk factors were present.
- 2) MA then conducted the ECG/CTT (CHART) examinations.

### **6.1.4 Step 4. General Practitioner (GP) evaluation**

- 1) Only after both the ECG report and the CHART report were available, they were reviewed by the GP.
- 2) The reports were reviewed by the GP's separately, ECG report was forwarded to one GP, and the CHART report was evaluated by other GP, so the bias is avoided.
- 3) GP was evaluating the assigned report and complete their diagnostic assessment based only on the received report.

### **6.1.5 Step 5 - Over-Read**

- 1) After the GP's completed their initial diagnosis for both ECG and CHART, Cardiologist #1, was provided an online Over-Read of the ECG/CHART findings.
- 2) Cardiologist was also completed this assessment in sequence, ECG first, then CHART and complete own independent diagnosis (no influence from GP's Diagnosis).

### **6.1.6 Step 6 - CHART#2 Examination**

1. MA in Cardiologists office, conducted CHART #2 examination.
2. Cardiologist than do the evaluation follow the next steps:
  - a. Review CHART #2. Provide initial reaction.
  - b. Is CHART #2 diagnosis similar to CHART #1? If changed, what is changed?
  - c. Is the priority still the same?
  - d. Based on CHART #2 only, what is your pre-Echo diagnosis?
  - e. What is your Echo Exam Start Point? Why?
  - f. Based on CHART #2, what are your appropriate Treatment options?



### 6.1.7 Step 7 - ECHO Examination

1. Cardiologists conducted an Echo examination based on Start point derived from CHART.
2. Cardiologist to determine diagnosis based on ECHO findings?
3. Does Echo confirm CHART results?
4. Did CHART provide appropriate start point for examination?
5. Was Echo more productive as a result of CHART?
6. Based on Echo, what are the appropriate treatment options? Did they change from pre-Echo?
7. Overall, did CHART help provide a better patient diagnosis? Why?
8. Upload data.

## 6.2 Phase 2

### Consensus Study – Validation of Ground Truth

#### 6.2.1 Step 1 – Video Audit

- 1) An audit of all the ECHO videos were conducted by Trained sonographer to verify the Echo measurements and consistency of measurement techniques. Any corrections were addressed in cooperation with the attending cardiologist.
- 2) The sonographers compared the measurements entered by the cardiologists with their own and those of the clinical reference standard, as detailed by the ASE (American Society of Echocardiography).

#### 6.2.2 Step 2 – Consensus Study

- 1) Once audited, each Echo examination was then reviewed by no fewer than three (3) Cardiologists, not involved in the original Echo examination, and come to a consensus on the diagnosis, to verify and validate the ground truth of the Echo and of the CHART findings.
- 2) They followed the same study parameters and procedures as was conducted during the previous consensus study which established the ground truth for each echo examination. Additionally, however, after entering their consensus results, the cardiologists also entered their consensus on the individual CHART findings.



## 6.3 Phase 3

### 6.3.1 Utility Study Component

The Clinical Study also validated the usability, operability, deployability and utility of CHART as a functional system.

First, each human interaction with the system was measured and assessed for usability and utility with pre-defined forms.

Second, each system element or sub-system, was assessed for operational efficacy, deployability, supportability and security.

More details about the objectives and risk derived from each step, can be found in the **Annex E- Steps of CUUS**.

## 7 Conduct of the appraisal

The nature of this Clinical Study is that of and IMD conducting non-invasive diagnostic measurements to assist physicians in determining cardiac state. The Clinician is responsible for the patient diagnosis and determining treatment options. As such no medical care is planned to be provided after study is completed.

### 7.1 Study participants

In CUUS Clinical Study the participants were:

- a) 3 Cardiology Clinic- General Hospitals in each center
- b) 6 Primary Care Clinics/Ambulances
- c) 8 Cardiologists
- d) 10 Family Physicians (General Practitioner- GP)
- e) 14 Medical Assistants
- f) 550 patients

### 7.2 Hypothesis

#### 7.2.1 Null- Hypothesis

ECG-only analysis is more effective than CHART analysis in understanding cardiac status in clinical practice.



### 7.2.2 Primary Hypothesis

- 1) "CHART analysis provides better diagnostic decision support in clinical practice compared to ECG-only analysis **leading to better patient diagnosis outcomes.**"

Prove this statement by

- a) measuring the reduction in FP and FN rates, and
- b) reduction in "no-evidence" and "Not Sure" answers,

between ECG-only and CHART analysis, which was verified by consensus decision ground truth, calculated according to a 95% confidence interval.

### 7.2.3 CTT Primary Hypothesis

"CTT bio-signals are valid and consistent in accordance with the accepted standards in clinical practice." This will be proved by:

- 1) ECG – measurements to the IEC60601-2-25 Standard for Diagnostic ECG.
- 2) PCG/MCG – measurements to the Signal Standard for PCG and MCG bio-signal measurement Standard.
- 3) Reproducibility and Repeatability testing of ECG, PCG and MCG between:
  - a. different CHART tests between Clinical Practice and Cardiology (CHART 1 and CHART 2)
  - b. different CHART test between Different Centers.

### 7.2.4 Secondary Hypothesis

- 1) CHART analysis provides better prioritization and referral medical justification than ECG-only analysis

This was accomplished by comparing the prioritization results and the diagnostic findings between the Clinician and the cardiologists, comparing CHART with ECG-only analysis.

The increase of effectiveness was measured by answers of the referral cardiologists using prioritization and medical justification between CHART and ECG-only, with 95% confidence interval.

- 2) CHART can provide an effective starting point for Echo examinations

This was accomplished by Comparing the diagnoses of CHART#2 and ECHO to show that cardiac status is correctly indicated by CHART analysis to show adequate starting point for an ECHO examination. This was confirmed by summing the cardiologists' statements about the effectiveness of CHART analysis to this end.



This effectiveness was measured by summing the cardiologists' statements about the effectiveness of CHART analysis to this end.

3) CHART medical tests are reproducible and repeatable

This was accomplished by comparing the results between CHART#1 and CHART#2, according to 95% confidence interval. This was confirmed by summing the cardiologists' statements related to their comparison between the two tests.

4) MA have effective clinical understanding of CHART usability

This was accomplished by measuring the completion rate by assigning a binary value of '1' if the test participant manages to complete a Medical test and '0' if he/she does not. The equation for this measure is:

$$\text{Effectiveness} = \frac{\text{Number of MT completed successfully}}{\text{Total number of MT undertaken}} \times 100\%$$

Concerning the post-test questionnaire with the questions about the usability and the most positive and negative aspects of the medical device, we confirmed the Usability by summing the answers of the MA and measuring the level of confidence with the 97.5% certainty, applying I-type error.

### 7.3 Hypothesis calculation

The performance evaluation was necessary to accept or reject the study's hypothesis.

The performance evaluation was based on the performance metrics which were recommended by FDA<sup>11</sup>:

- sensitivity (SE) and specificity (SP) at a clinical action point:
  - "SE is defined as the probability that a test is positive for a population of patients with the disease/condition/abnormality"
  - "SP is defined as the probability that the test is negative for a population of normal patients (i.e., patients without the disease/condition/abnormality)"
- receiver operating characteristic (ROC) curve: "ROC based endpoint allows evaluation of the device over a range of operating points"

More calculation for sensitivity and specificity, referral decision and ROC curve can be found in **Annex F- Statistical Plan for CUUS**.

---

<sup>11</sup> Clinical Performance Assessment: Considerations for Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data - Premarket Approval (PMA) and Premarket Notification [510(k)] Submissions, Available from:

<https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/ucm187315.pdf>



## 7.4 The “Send/Don’t send” decision

One possible modality in identifying the benefit of CHART over ECG analysis can be measured in the “Send/Don’t send” decision (referring to diagnosis) where the GP must decide based on the diagnosis assistance they receive from either CHART or ECG whether to send, or don’t send, the patient to the next level of care, or to the Cardiologists level of care. This decision was made based on ECG-only and CHART analysis separately by two different Physicians,<sup>12</sup> for the same patient and evaluated through statistical analysis of comparison to establish the certainty of the decisions and understanding of the diagnosis.

A better understanding of the diagnosis also leads to higher confidence levels in assisting the GP to better understand what to prescribe because better medical justification was provided or missing, depending on whether it is ECG only analysis or CHART analysis.

The main goal of this study was to prove if CHART analysis is more effective than ECG only analysis, is assisting the GP in determining their “Send/Don’t send” decision (diagnosis) and the basis for it.

Secondarily, the “Send/Don’t send” decision has the goal to reduce false positive, false negative values and no-evidence rate, and to then use Consensus decision to prove the Ground Truth. In other words, the CHART analysis report should enable the GP to make better “Send/Don’t send” decision.

When the diagnosis was made, and the decision was to “Send” the patient to the Cardiologist, the GP first provided the basis for their diagnosis, the medical justification, and then determine if as a result they were more certain of their decision. The objective was to, not only determine if their confidence level in their action increased as a result of the additional medical justification provided, and in a secondary effect reduce the “if in doubt, send” syndrome which can paralyze access to higher levels of healthcare by clogging the system with healthy patients, as an example. Furthermore, patients being referred without appropriate medical justification, couldn’t be assigned a proper triage priority, putting the patients at risk of being denied timely access to appropriate treatment options.

## 7.5 Randomization, bias and blinding

To avoid any measurement bias, the MA does all the MT on the patients.

The randomization of the study population is based on a first comer basis, and then qualified through the Risq questionnaire. First-comer subjects required at least 3 risk

---

<sup>12</sup> At GP level, Physicians will be divided by test- one Physician will evaluate ECG-only report and make the referral decision, and the CHART report related to the same patient will be evaluate by other Physician, to prevent bias, and to follow the rules of randomization and blinding.



factors for heart disease to participate in the CI, this is in accordance with the accepted standard of care as established by Medicare in the USA. The subjects visited the medical clinic for their own normal medical reasons, they are not recruited.

The “first comer base” expression was used to explain and defined the patients who are randomly and without invitation going to see their doctor, for various reasons known only to them. The patient wasn’t evaluated nor selected in advance, nor does the GP inform the patient about the ongoing clinical study in order to include him/her according to his/her health history. During the examination, if the GP confirms that the patient has at least 3 risk factors present (as per CIP and protocol), the GP will introduce the patient with the clinical study and asked if he/she is willing to participate. To summarize, the patients are randomly chosen to be participants in the study, according to the risk factors.

The “first-come-first-served basis” means that patients will receive the medical treatment or be dealt with in the order in which they arrived in the doctor’s office, not related in any way with the clinical study.

After the Risq questionnaire, if three risk factors are identified, the subject was enrolled in the study, they assigned to a study sub-group. The randomization of the population then ends. If the study sub-groups did not meet the minimum, the clinical study was not extended, on a first-comer basis, until all sub-group categories are filled.

In a next phase, a Physician evaluated either the ECG-only or the CHART report for that patient, but not both. If one Physician reads the ECG-only report, a different Physician evaluated the CHART report for that patient. This ensured physicians was blinded to the results of the two tests.

Each physician, based on what test they were assigned to review, completed the Send/Don’t send decision and provide their diagnosis. The physician’s role then ends.

The ORC, Over-reading cardiologist, reviewed both the ECG and the CHART reports. He compared them for the purpose of determining if CHART provides more information that helps him/her to better understand the patient cardiac status. There is no randomization at this level. However, several cardiologists conducted these reviews, the patients randomly assigned to them. A bias towards ECG could be set because of the order or reading and familiarity with ECG.

At the Consensus Study level, the Cardiologists were mixed by centers, so the Cardiologists which were doing the initial evaluation of the patient in one center, did the Ground Truth for the other Study Center, and vice versa.

There ensured there is in-center bias between cardiologists.

The patients have their own ID, and the doctors were blinded as they did not know whose reports they were reviewing. Also, the patient would not know which examination determines the outcome of their diagnosis.



Issue of the blinding/masking was resolved using these approaches.

## 7.6 Sample size calculation

We enrolled adult subjects, age 20+, to participate in clinical tests (ECG only with Cardio-TriTest v6.5<sup>13</sup> and CPA, together Cardio-HART™ or CHART) when attending a primary patient care clinic. Subjects are grouped by age, gender, BMI at various locations, so as to reach the expected sample size. Such groupings (categorizing) occurred after subject enrollment and data collection and reviewed weekly to ensure variability.

---

<sup>13</sup> Cardio-TriTest ECG signals are validated according to IEC 60601-2-25:2011 (Medical electrical equipment - Part 2-25: Particular requirements for the basic safety and essential performance of electrocardiographs)

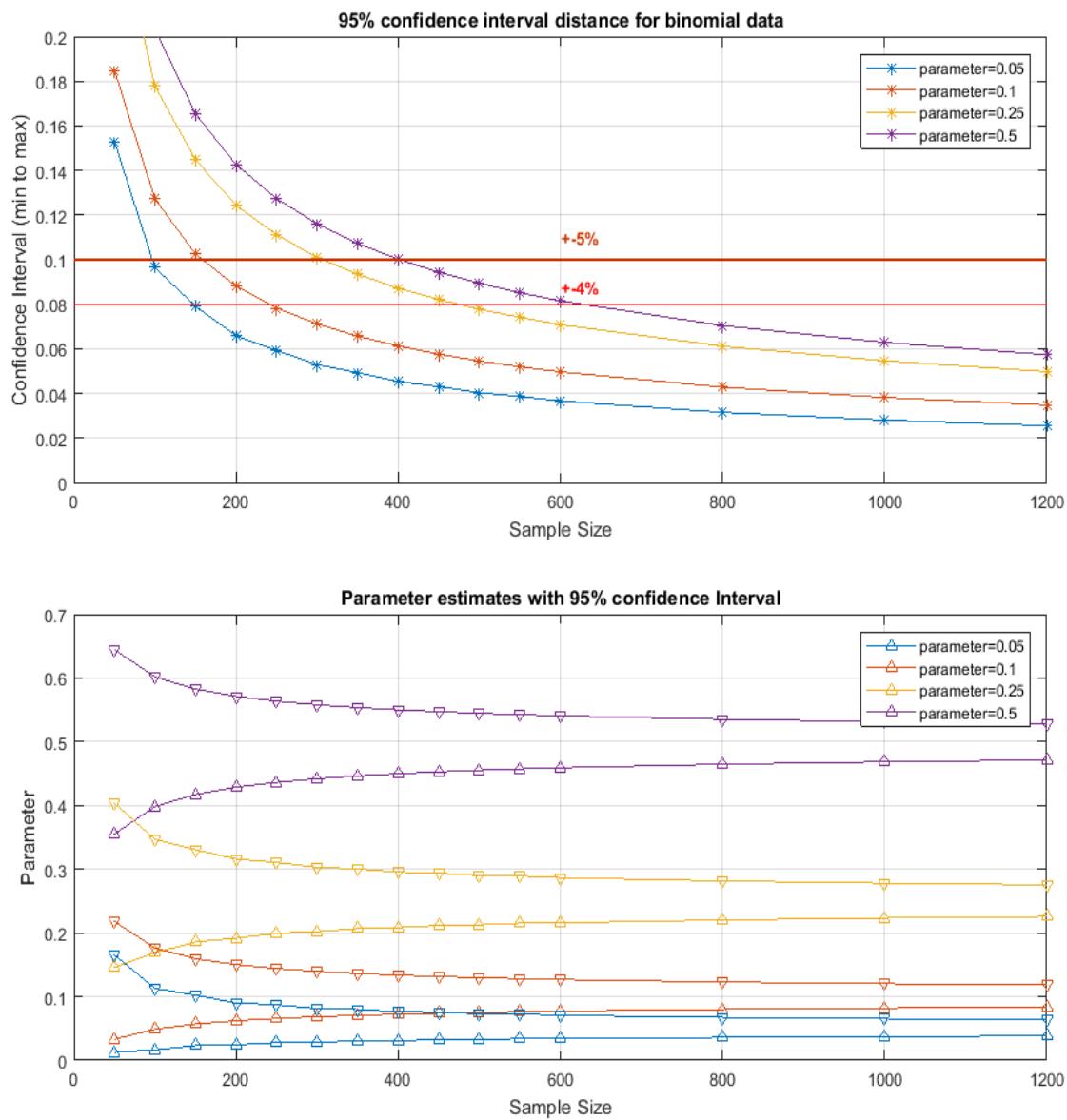


Figure 4- Sample size calculation

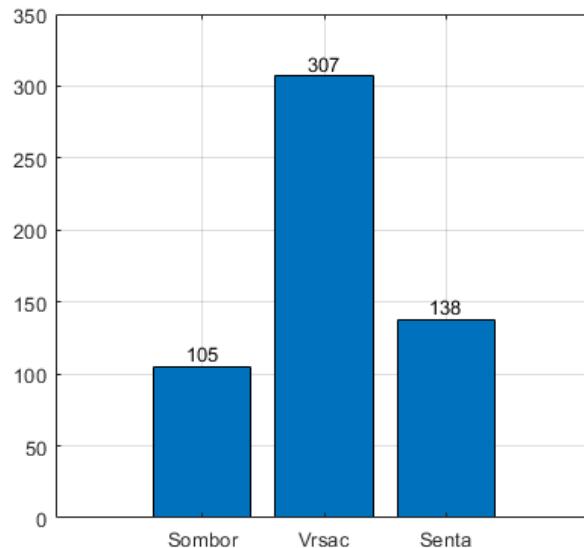


## 7.7 Analysis population

The efficacy analysis was intended to be performed on the population which included all subjects who were randomised and collected during the clinical study- total 421 subjects, in three different centers- Sombor and Vrsac. The most of the subjects are enrolled in Vrsac (56.4% of the whole study population).

**Table 1- Center distribution**

		Frequency	Percent
Valid	Sombor	111	20.2%
	Vrsac	310	56.4%
	Senta	129	23.4%
	Total	550	100%



**Figure 5- Center distribution**

**Table 2- Age distribution**

	Age categories	N	Percent
Valid	20-40	61	11.0
	41-55	130	23.6
	56-65	156	28.3
	66-75	148	26.9
	76+	52	9.5
	Total	550	100.0



Table 3- Descriptive statistic

	N	Minimum	Maximum	Mean	Std. Deviation
Age [years]	550	20	92	59.7	14.3
Valid N (listwise)	550				

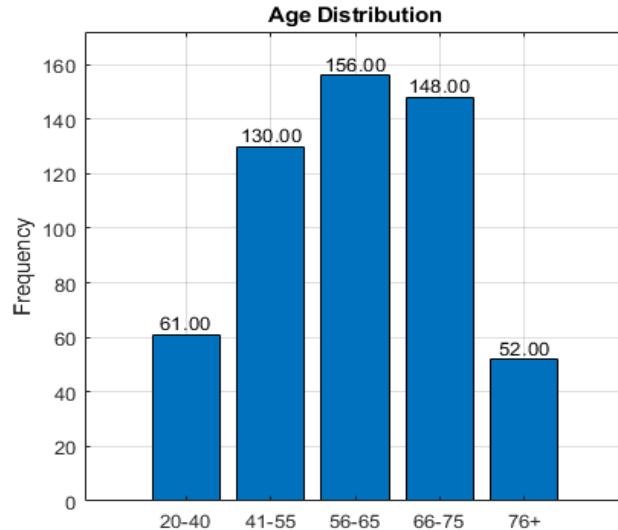


Figure 6- Age distribution

On Table 2 and Graph 6 it is shown the whole sample by Age distribution. We can see that the most of the subjects are 56-65 years (28.3%). The average age in this study is 59 years. 64% of the patients is older than 55 while 9.5% of the patients are older than 76.

Table 4- Gender distribution

		Frequency	Percent
Valid	Male	283	51.5
	Female	267	48.5
	Total	550	100.0

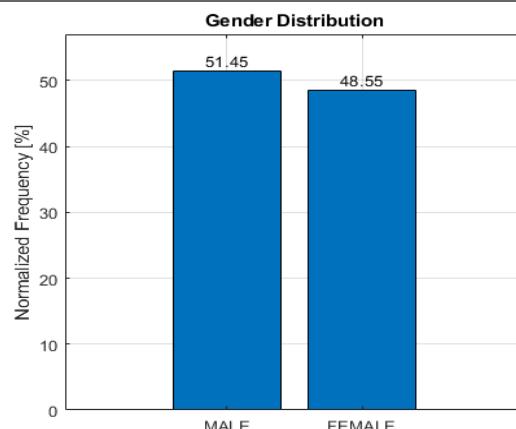


Figure 7- Gender distribution

The gender distribution is approximately equality (Male 51.5%, Female 48.5%).



Table 5- BMI distribution

		Frequency	Percent
Valid	Underweight (below 18.5)	34	6.2
	Normal weight (18.51 – 24.99)	197	35.8
	Overweight 25 – 29.99)	191	34.7
	Obese (30 & higher)	128	23.3
	Total	550	100.0

Table 6- Descriptive statistic

	N	Minimum	Maximum	Mean	Std. Deviation
BMI [kg/m^2]	550	16.11	43.34	26.7	5.18
Valid N (listwise)	550				

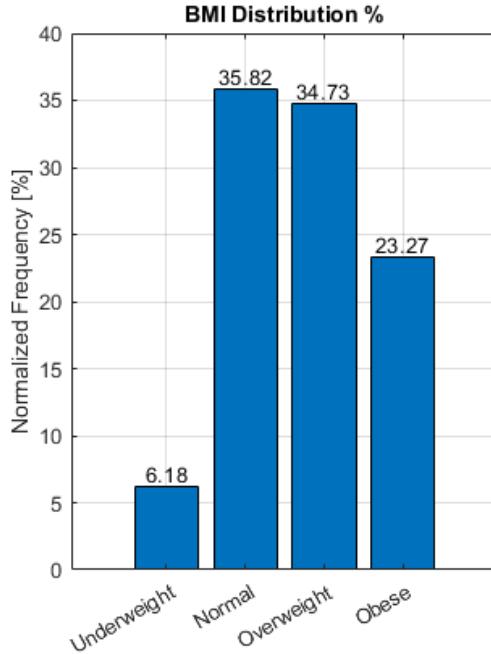
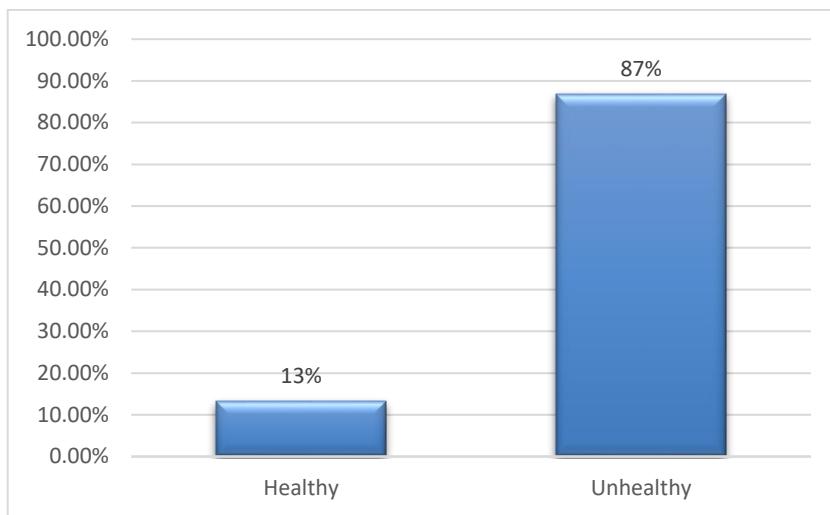


Figure 8- BMI distribution

As per BMI, the average index is 26.7, which appoints to the fact that the enrolled subjects are leaning to obesity. 58% of the subjects have a problem with extra kilograms, which represent one of the risk factors that can lead to the cardiac diseases.

Table 7- Healthy/Unhealthy distribution

		Frequency	Percent
Valid	Healthy	71	13
	Unhealthy	479	87
	Total	550	100



**Figure 9- Healthy/Unhealthy distribution**

We collect 13% of the healthy subjects, and 87% of unhealthy subjects (subjects with at least 3 risk factors established).

During the evaluation of the collected data, we excluded 13 patients from the analysis because the MA at one site did not follow the protocol, and the medical tests were not completed and valid.

Also, some of the patients withdrew in the middle of the study, after their Medical Test but they didn't go to the Cardiologist level, so we missed their CHART#2 reports and ECHO examinations. For that reason, in the RC evaluations we included only 515 valid subjects and 549 subjects for ORC evaluation.

Withdrawal of patients were mainly due to delays in attending the cardiologist level and having to travel some distances to the center.

A confirmatory analysis was performed on the per-protocol population which included all subjects that completed the tests and ECHO, and did not meet any major protocol violation during the study period. The safety population included all subjects who has at least 3 risk factors for cardiac diseases (in 85% of the population).



## 7.8 Statistical methods

All statistical methods were based on the International Conference on Harmonization (ICH) E9 document “Statistical Principles for Clinical Trials”<sup>14</sup> and on E9 (R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials document.<sup>15</sup>

All data was summarized by estimated group. For baseline characteristics and safety outputs, a total overall column was included to summarize all subjects. In summary tables of continuous variables, the minimum and maximum statistics are presented. The arithmetic mean (AM), standard deviation (SD) and standard error (SE) are extracted also, in every descriptive statistic table, if applicable and necessary.

The data was collected based on well-defined inclusion and exclusion criteria. The Gender, Age, BMI, BSA, RisQ Factors and other demographic criteria were categorically distributed, for the purpose of demonstrating the differences between the clinical study population and intended use population, that may impact the device.

For the statistical analysis, the equations presented in Annex F- Statistical Plan for CUUS were performed. All hypothesis testing was carried out at the level of significance 0,05 (every coefficient which is  $\geq 0.05$  will be significant). This means that the objectives were tested at the 5% level of significance (2-sided), or with at least 95% of certainty. P-values were rounded to two decimal places. P-values less than 0.001 were reported as <0.001 in tables. P-values greater than 0.999 was reported as >0.999. Also, for the Usability objective testing, we used Type-I error, with 97,5% confidence interval.

For additional statistics, for presenting the sample of this study etc., we used descriptive statistics measures (Frequencies, Descriptive, Cross-tabulations, Chi-square) and analysis of variance (ANOVA test).

All statistical analyses were performed using SPSS v23 or higher for Windows, MatLab and Weka.

Because of the use of qmsWrapper software custom forms (CRF), each question needed to have an answer (the form is designed in such way that the answers are obligatory to fill), so the missing data were reduced to minimum, basically, they are non-existing.

---

<sup>14</sup> Statistical Principles for Clinical Trial- to establish valid Hypothesis. Available from:  
[http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Efficacy/E9/Step4/E9\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf)

<sup>15</sup> E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials, to determine the suitable statistical analysis we are going to use in our CS. Available from: <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM582738.pdf>



## 7.9 Expected results

We expect that the null-hypothesis will be rejected.

For FP results, compared to ECG only results, CHART will reduce the rate of FP meaningfully.

For FN results, compared to ECG only results, CHART will reduce the rate of FN substantially.

For Secondary Endpoints, the results will generally show that the diagnostic assistance from CHART is superior to ECG only. That CHART will not only match ECG results for ECG detectable diseases, but CHART will also demonstrate its additional performance characteristics for heart diseases.

Success will be measured through various comparisons of the data, starting with comparing the CHART#1 Test done at the primary clinical level, and the CHART#2 Test, done at the cardiologist level of care.

Other usability results will be measured and monitored by the Monitors during on-site visits. The accumulated data, will be analyzed accordingly. However, it is expected to demonstrate that the simplified UI will be more understandable and generate few errors as it “shapes” the test protocol. Furthermore, it is expected to show that deployability and operational ability in clinical practice will be high and that clinical utility will be proven.

Claims on clinical performance and clinical safety:

- 1) Overall reduction of FP
- 2) Overall reduction of FN
- 3) Greater confidence in understanding the diagnosis
- 4) Consistency between initial diagnosis in Primary Care right through to Ground Truth (i.e. what is diagnoses in Primary Care, is more consistent with GT than comparator device)
- 5) At least 10% change in Cardiologist determination of Send/Don't Send decisions
- 6) Little or no AE (Adverse Event) during Clinical Study
- 7) Simple integration into Standard of Care
- 8) Higher confidence in Primary Care Physicians in understanding their patient's cardiac status, over the comparator device.

The Study has been defined in such a way as to confirm the manufacturer's claims for the device.



## 8 Analysis of the clinical data

### 8.1 The Challenge of SE and SP measurement

The use of sensitivity (SE) and specificity (SP) as a measure for comparing the effectiveness between CHART and ECG has to be carefully considered as it can be easily influenced by “out of protocol” (OoP) considerations taken by the individual GP and differences in the application of the Standard of Care (SoC).

In location A, Doctor #1 used ECG, and Doctor #2, used CHART. However, it was soon discovered after the start of the CS, that the local SoC included the practice of referring patients to cardiology on the single purpose of only hypertension – a single risk factor (outside the IDC10 codes). Unknown to the Sponsor, the GPs cannot prescribe hypertension medication to patients and as such the SoC imposed on GP's that they refer all patients to cardiology if they suffered from any form of Hypertension.

Although the CS protocol required 3 risk factors for a referral to cardiology, the local SoC required only one if it was Hypertension.

As a result of this issue, Dr. #1 had “lost” their sense of understanding how ECG could contribute to patient referral, so prevalent was hypertension in the patient population. All patients with Hypertension were subsequently referred to the cardiologist. Dr. #1 also used the Risk assessment information to substitute for lack of ECG based medical justification, leading to a practice of “when in doubt, send”, there was a lot of doubt.

This initially skewed the SE results in favor of ECG, because ECG was not the main basis of the medical justification for the referral by Dr #1. The practice was caught and the Dr.#1 subsequently revised the basis of their referral for those patients where Hypertension, non-IDC10, was the only risk factor. For their patients where Risk assessment information was used in lieu of ECG these examinations were also reviewed and revised with permission of the Dr#1 and the PI.

Although Dr#2 understood the role of hypertension in the SoC, they did not use it as a single risk factor for the basis for the referral, relying instead on CHART to provide the medical justification as per the Clinical Study protocol. There were few Hypertension only referrals, limited to patients with a more chronic form of the condition and valid IDC10 code, which was also considered valid by the cardiologist.

GP's faced with limited ECG knowledge have difficulty understanding what is valid medical justification for their Send/Don't Send decision. This was most clearly demonstrated by the actions of Dr#1 when she switched from the ECG protocol to the CHART protocol. With ECG she referred far more patients to the Cardiologist. After switching to the CHART protocol, she referred dramatically fewer, more in line with her colleague's decisions. Similarly, but not as dramatic, when Dr#2 switched from the CHART protocol to the ECG protocol, her referrals to the Cardiologist also increased,



although not nearly to the same extent of Dr#1. The difference was that Dr#2 had more recent ECG training and some experience.

This was one of the goals of CHART that was “confirmed” as a main benefit in particular for use in Primary Care.

Specifically, when comparing the same patients, Dr#1, using the ECG protocol, and Dr #2 on the CHART protocol, for the same patients, the results indicated that Dr. #1 referred the same patients to the Cardiologist some 46% more often than Dr #2 using CHART. This was much higher than anticipated.

Dr#1's high ECG Send volume of referrals effectively skewed the SE results, making ECG appear to be more sensitive than CHART, when in fact the opposite was expected, as CHART can diagnose a much broader ranger of diseases.

For the each of the same patients, Dr. #2 using CHART, was far more accurate in determining each patient's actual cardiac status, thereby referring far fewer of the same patients to Cardiology. With more accuracy came more confidence in the diagnosis.

When compared to the skewed effect of Dr#1's ECG decisions, Dr#2's CHART results thereby had the effect of making CHART appear to be more SP oriented by contrast.

Further investigation however, revealed that in fact CHART was both more SE and SP oriented, because of its ability to consistently diagnose the patient's actual condition which included a broader range of patient conditions. This was verified by the Ground Truth, and hence the results showing a much lower referral rate to cardiology.

In contrast, when Dr #2 was put on the ECG protocol, two unintended observations were made. First, it was suspected that in some 17 patient cases, the diagnosis could not possibly be from ECG results – they were not ECG diagnosable diseases. It was subsequently discovered, that Dr. #2, having used CHART to diagnosed some 150 patients had simply grown reliant on CHART to provide a high degree of accuracy. So, when Dr. #2 came upon a patient suspected of having a disease that they could not diagnose using ECG, yet the risk factors showed something, they then consulted the patient's CHART report as well, breaking protocol. Their reasoning was simple, the patient was more important to them and if she suspected that CHART could do the job, they did not hesitate. These CHART influenced diagnoses were corrected with Dr. #2's, and the PI's permission. For the remainder of the study, Dr. #2 followed the protocol. Not surprisingly having only access to ECG reports, Dr #2 started referring a higher percentage of patients to the cardiologist than when they were using CHART, albeit not as many as Dr.#1 on ECG, by comparison.

This clearly showed that with ECG, both doctors were less certain of their patient's cardiac status, but ECG also greatly exaggerated the differences between GP's. CHART on the other hand had a diminished effect of exaggeration as both GP's when using CHART had similar rates of referral, with Dr.#1 still being slightly higher, rather than



significantly higher. This was attributed to CHART increasing their confidence levels in their diagnosis.

It should also be noted that when Dr. #1 was in turn put on the CHART protocol their referral to cardiologist rates dropped, meaningfully, and referrals became more consistently similar to the rates of Dr. #2 when on the CHART protocol. When asked why, Dr. #1 responded that it was because CHART gave them more relevant information and as such more confidence in their diagnosis compared to ECG.

It should also be noted, that the CHART diagnosis, whether by Dr. #1 or #2, were much closer to the same diagnosis as concluded by the ORC, the RC, and the Ground Truth team. Compared to ECG, CHART was consistently the same between from the CHART#1 test, to the CHART #2 Test, to the Echo Examination and to the validated Ground Truth. That means from Primary Care to Ground Truth, CHART results showed diagnostic consistency.

## 8.2 Performance measurements

### 8.2.1 Referral decision

The detailed and binary referral decision was organised in an ascending order.

Order	Detailed decision	Abbreviation	Binary decision
1.	No Action	NoA	
2.	Watch 12 months	Wa12	Don't
3.	Watch 6 months	Wa6	
4.	Watch 3 months	Wa3	
5.	Routine	Rout	
6.	Immediate	Imme	
7.	Urgent	Urge	Send
8.	Emergency	Emer	
9.	Not Sure	NSur	



## 8.2.2 Confusion matrix

The confusion matrix of binary decision about patient referral is the following:

		Test Decision (decision on ECG or CHART report)		
		~ Don't	~ Send	~ Not Sure
Reference decision (Ground Truth decision)	# Don't	TN	FP	NS <sup>-</sup> – Negative Not Sure
	# Sent	FN	TP	NS <sup>+</sup> – Positive Not Sure

The number of samples in analysis  $N = TN + TP + FP + FN$

The “Not Sure” decision answers are relatively rare, therefore in the binary decision they are not expressed separately.

The detailed confusion matrixes show “Not sure” kind answers too:

		Test Decision								
		~NoA	~Wa12	~Wa6	~Wa3	~Rout	~Imme	~Urge	~Emer	~NSur
Reference decision	# NoA									
	#Wa12									
	#Wa6									
	#Wa3									
	#Rout									
	#Imme									
	#Urge									
	#Emer									
	#NSur									

## 8.2.3 Performance metrics

Metric	Description	Symbol and Formulae
Sensitivity	True positive rate compared to positive samples	$SE = \frac{TP}{TP + FN}$
Specificity	True negative rate compared to negative samples	$SP = \frac{TN}{TN + FP}$
Overall Accuracy	Rate of correct decision compared to all samples	$ACC = \frac{TN + TP}{N}$
Positive Predictive Value	provide useful insight into how to interpret positive test results	$PPV = \frac{TP}{TP + FP}$
Negative predictive Value	provide useful insight into how to interpret negative test results	$NPV = \frac{TN}{TN + FN}$



<b>Positive Likelihood Ratio</b>	likelihood ratio for positive results	$LR+ = \frac{SE}{1 - SP}$ <sup>16</sup> $LR + * = 10 \cdot LR +$
<b>Negative Likelihood Ratio</b>	likelihood ratio for negative results	$LR- = \frac{1 - SE}{SP}$ $LR - * = 10 \cdot LR -$
<b>Positive Percent Agreement</b>	Positive agreement is the proportion of comparative method positive results in which the test method result is positive <sup>17</sup>	$PPA \equiv SE$
<b>Negative Percent Agreement</b>	Negative agreement is the proportion of comparative method negative results in which the test method result is negative.	$NPA \equiv SP$
<b>Positive rate</b>	Rate of positive decision from test (GP, ORC or RC decision)	$PR = \frac{FP + TP}{N}$
<b>Prevalence</b>	Rate of positive decision from reference (consensus ground truth)	$PREV = \frac{FN + TP}{N}$
<b>Not Sure Rate</b>	The rate of not sure answer compared to all samples	$NSR = \frac{NS^- + NS^+}{N}$
<b>Cohens Kappa</b>	Kappa <sup>18</sup> a more robust measure than simple percent agreement (ACC), since it takes into account the possibility of the agreement occurring by chance	$K = \kappa = \frac{p_o - p_e}{1 - p_e}$ observed agreement $p_o = \frac{TN + TP}{N} = ACC$ hypothetical probability of chance agreement $p_e = p_- + p_+ =$ $= \frac{TN + FP}{N} * \frac{TN + FN}{N} + \frac{FN + TP}{N} * \frac{FN + TP}{N}$
<b>Correlation</b>	Pearson correlation coefficient calculated only for detailed decision to express the similarity in details – when e.g. Watch3 more similar to Watch12 than Urgent	$r(x, y) = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}}$
<b>Relative Classifier Information</b>	Relative Classifier Information (RCI) <sup>19</sup> is similar to Kappa, but proposed to multi-class confusion	

<sup>16</sup> LR+\* and RL-\* are used for better plottable together with percentage performance metrics<sup>17</sup> <https://analyse-it.com/blog/2020/4/diagnostic-accuracy-sensitivity-specificity-versus-agreement-ppa-npa-statistics><sup>18</sup> Chmura Kraemer, Helena, Vyjeyanthi S. Periyakoil, and Art Noda. "Kappa coefficients in medical research." Statistics in medicine 21.14 (2002): 2109-2129.<sup>19</sup> Sindhwani, V., Bhattacharjee, P., & Rakshit, S. (2001). Information theoretic feature crediting in multiclass support vector machines. In First SIAM international conference on data mining (ICDM'01). Chicago, IL, April 5-7.



<b>Area Under Curve</b>	Area under ROC curve. Described in next subsection	<i>AUC</i>
-------------------------	---	------------

### 8.2.4 Area Under Curve

AUC calculated by moving referral threshold on GP's decision toward ground truth binary decision.

The AUC is calculated on the following ROC points:

ROC points	Simulated Positive	Simulated Negative	Reference Positive
<b>1. Emergency</b>	GP decision = Emergency	GP decision < Emergency	Ground truth >= Routine
<b>2. Urgent</b>	GP decision >= Urgent	GP decision < Urgent	Ground truth >= Routine
<b>3. Immediate</b>	GP decision >= Immediate	GP decision < Immediate	Ground truth >= Routine
<b>4. Routine (real point)</b>	GP decision >= Routine	GP decision < Routine	Ground truth >= Routine
<b>5. Watch 3</b>	GP decision >= Watch 3 months	GP decision < Watch 3 months	Ground truth >= Routine
<b>6. Watch 6</b>	GP decision >= Watch 6 months	GP decision < Watch 6 months	Ground truth >= Routine
<b>7. Watch 12</b>	GP decision >= Watch 12 months	GP decision < Watch 12 months	Ground truth >= Routine

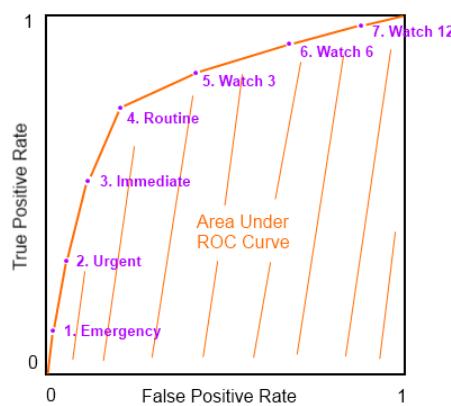


Figure 10- Illustration of ROC analysis and AUC calculation using seven ROC points

## 8.3 Decision Performance of CHART over ECG



In the performance statistics 550 patient are evaluated having ground truth from consensus study from two different cardiology center.

The  $PREV=43,0\%$  of patients are sent to referral cardiology – this is the positive samples from ground truth.

The report and statistics of consensus-based ground truth is written in chapter 8.5.

### **8.3.1 GP decision compared to ground truth**

This section presents the GP decision performance decided on ECG and CHART reports where the reference decision is the consensus-based ground truth.

There are 550 patient having ECG and CHART based decision by GP and ground truth decision.

**Table 8- Confusion matrix**

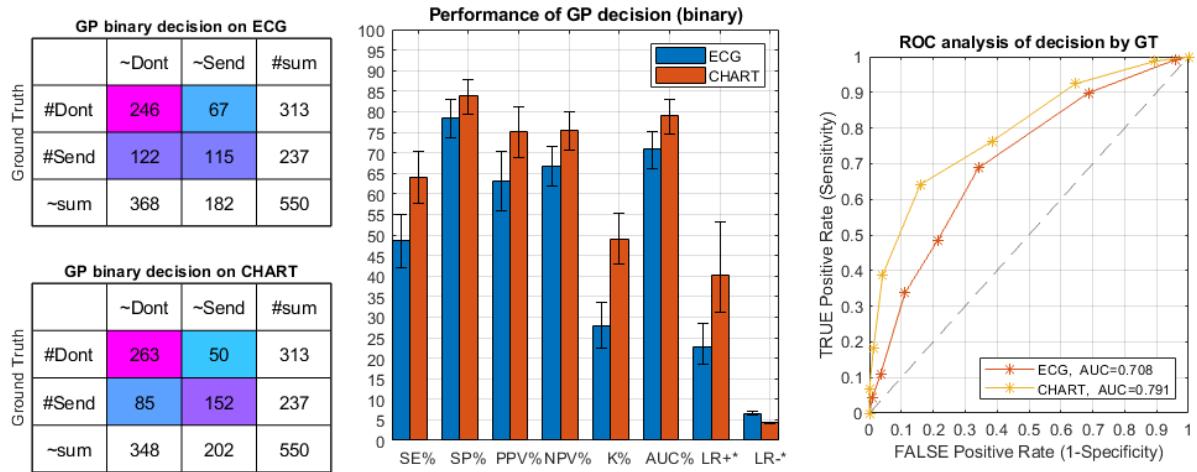
		Ground Truth	GP decision		
			Don't	Send	Summary
ECG report	Don't	246	67	313	
	Send	122	115	237	
		Don't	Send		
CHART report	Don't	263	50	313	
	Send	85	152	237	

**Table 9- Performance**

Metric	Performance			Hypothesis test with 95% CI
	Lower conf. %	Observed Value %	Upper conf. %	CHART compare to ECG
ECG report	SE%	42.00	48.52	55.08
	SP%	73.63	78.59	83.01
	PPV%	55.74	63.19	70.20
	NPV%	61.78	66.85	71.64
	K%	22.52	27.90	33.71
	AUC%	66.02	70.80	75.02
	LR+	1.84	2.27	2.86
	LR-	0.70	0.65	0.62
	PR%	29.17	33.09	37.20



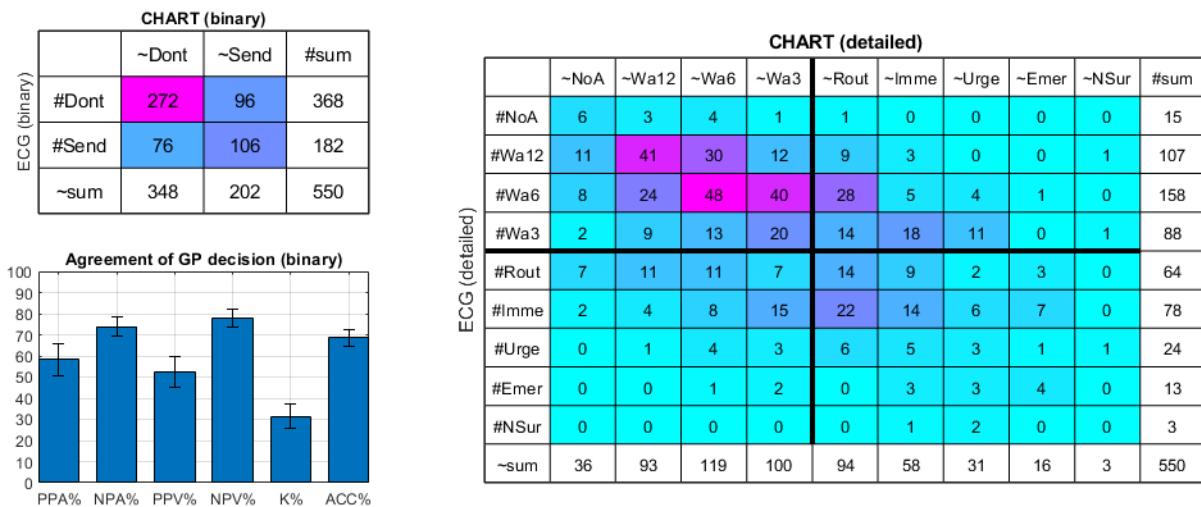
CHART report	SE%	57.67	64.14	70.24	<b>significant increase</b> <sup>20</sup>
	SP%	79.49	84.03	87.91	<b>significant increase</b>
	PPV%	68.70	75.25	81.04	<b>significant increase</b>
	NPV%	70.71	75.57	80.00	<b>significant increase</b>
	K%	42.89	49.04	55.25	<b>significant increase</b>
	AUC%	74.44	79.13	83.03	<b>significant increase</b>
	LR+	3.13	4.01	5.30	<b>significant increase</b>
	LR-	0.45	0.43	0.41	<b>significant decrease</b>
	PR%	32.69	36.73	40.91	not significant increase <sup>21</sup>



**Figure 11- Performance of GP decision by consensus ground truth**

<sup>20</sup> The increase of decrease is decided by comparing the two observed values. The significance is derived from confidence intervals: if the observed value higher than the upper confidence value, then it considered as significant increase

<sup>21</sup> Not significant increase means the increase of observed value, but the observed value is not higher than the upper confidence value, and thus, classified as not significant



**Figure 12- Agreement analysis between GP decision on ECG and CHART reports**

### Conclusions:

- CHART based decisions show higher accuracy compared to ECG, especially its sensitivity have big increase. It was expected based on the performance validation of CPA. CHART decrease the false positive rate from 21% to 16%, and false negative from 50% to 34%. The significant increase of PPV, NPV, Kappa, AUC, and LR+ and decrease of LR- confirm this benefit by CHART report over ECG only report.
- The positive rate or send rate by GP is very similar to the ground truth for both kind of reports (compare PR%). This means that send rate of GP referral decision is the expected and meet with the cardiologist expectation.
- There is significant deviation in GP decision between ECG and CHART (see agreement analysis in Figure 12). From the 550 patients there are 96 patients send by CHART and Don't send by ECG, and there are 76 patients Send by ECG and Don't sent by CHART.
- The most frequent deviation in the detailed decision statistics is the 28 patients classified as "Send - Routine" by CHART and "Don't - Watch for 6 months" by ECG. This confirms the increase of sensitivity.

#### 8.3.1.1 Subgroups performance comparison

In the diagnosis we can distinguish typical ECG diagnosable (e.g. arrhythmia or ischemia) and typical CHART+<sup>22</sup> - diagnosable (e.g. systolic dysfunction or valve disease), see details in Table 13 in section 8.5.1.

In this analysis the GP decision performance four subgroup of patients based on the type of diagnoses diseases:

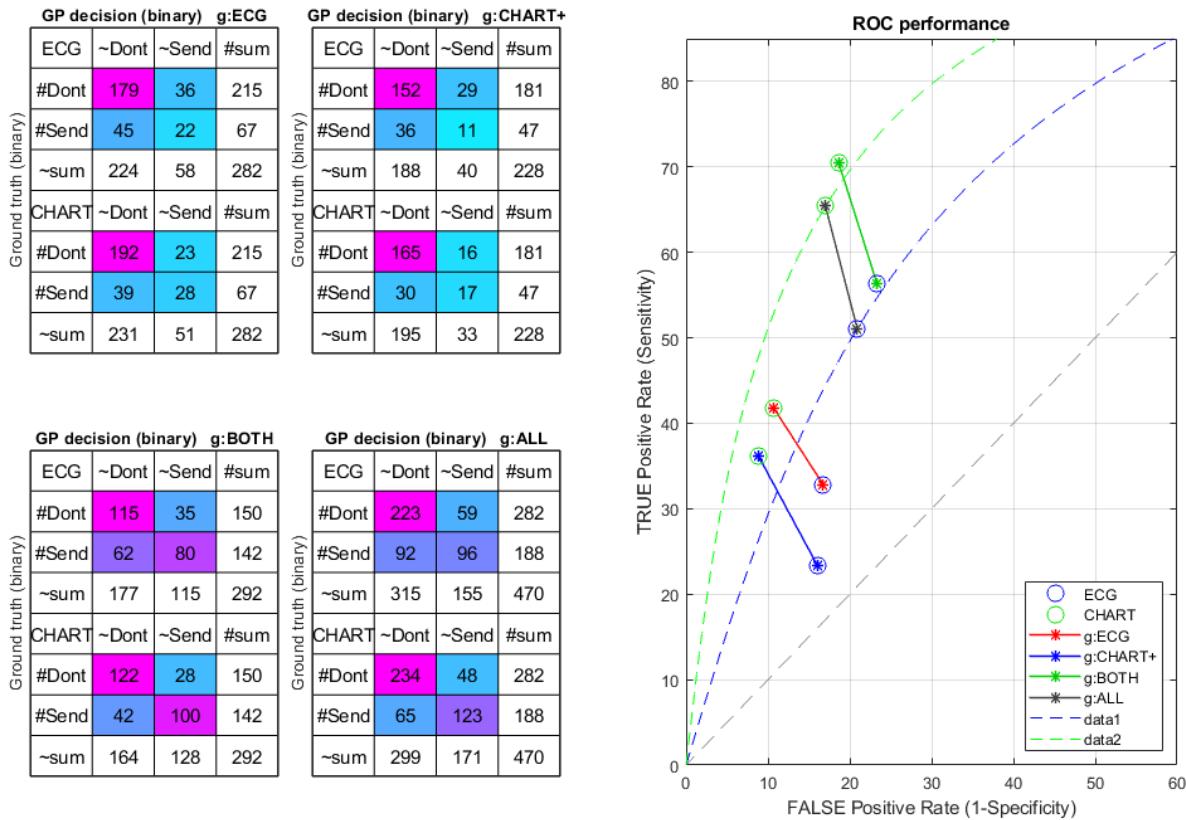
1. g:ECG – positive ECG diagnosable findings, plus the normal set (no cardiac disease)

<sup>22</sup> CHART+ because CHART includes ECG findings as well, but shows the additional non-ECG findings, which denotes the "+" after "CHART"



2. g:CHART+ – positive CHART+ diagnosable findings, defined as non-ECG diagnosable plus the normal set
3. g:BOTH – patients having both ECG based and CHART+ based diagnosable findings in parallel plus the normal set. This excludes patients having only ECG or only CHART+ diagnosable findings (abnormal set from the previous two subgroups)
4. g:ALL – include all the patients, all abnormal, all normal

The confusion matrix and ROC performance analysis are plotted in the Figure 13.



**Figure 13- Performance of GP decision compared to ground truth for subgroups**

#### Conclusion:

- The benefit of CHART over ECG is very consistent through these subgroups. This means the GP decision is more accurate using the CHART report for each subgroup, so benefit is independent from the type of diagnosis, e.g. it is not dependent on whether it is a typical ECG abnormality or a CHART+ abnormality.
- Both the positive decision rate or prevalence are significantly different in these subgroups, in other words the rate of diseases patients is different. The ECG only and CHART+ diagnosis subgroups have smaller positive ground truth and decision compared to BOTH subgroups. This means that when both ECG and CHART+ indicate any abnormality, then it is more likely that the patient should be Sent to cardiology.

CHART provides important additional diagnostic power beside ECG findings, and can more clearly indicate that the patient has some measure of cardiac disease and should



therefore be referred to cardiology. – This conclusion is derived from the previous conclusion point.

### **8.3.2 GP decision compared to ORC and RC decision**

This section compared the GP to ORC/RC decision made on the same patient ECG and CHART reports.

There are 549 patients with GP and ORC decisions and 515 with GP and RC decisions.

**Table 10- Confusion matrixes**

		Reference decision	Test decision - GP		
		ORC decision	Don't	Send	Summary
ORC and GP decision on <b>ECG report</b>		Don't	271	78	349
		Send	96	104	200
		ORC decision	Don't	Send	
ORC and GP decision on <b>CHART report</b>		Don't	279	64	343
		Send	68	138	206
		RC decision	Don't	Send	Summary
RC and GP decision on <b>ECG report</b>		Don't	266	73	326
		Send	78	98	189
		RC decision	Don't	Send	
RC and GP decision on <b>CHART report</b>		Don't	262	64	326
		Send	61	128	189

**Table 11- Performances**

	Metric	Performance			Hypothesis test with 95% CI
		Lower conf. %	Observed Value %	Upper conf. %	CHART compare to ECG
ORC and GP on ECG report	PPA%	44.8	52.0	59.1	
	NPA%	72.9	77.7	81.9	
	PPV%	49.6	57.1	64.4	
	NPV%	69.0	73.8	78.3	
	K%	24.5	30.2	36.2	
	PR%	29.2	33.2	37.3	

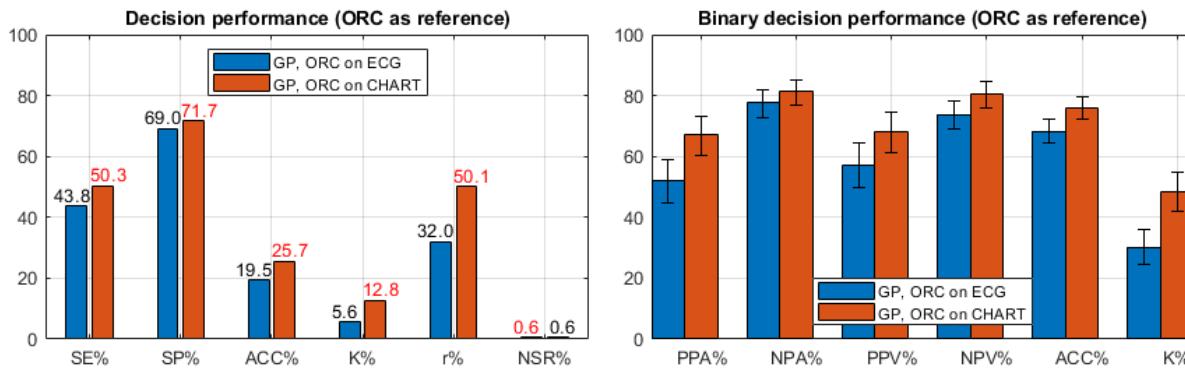


ORC and GP on CHART report	PPA%	60.1	67.0	73.4	<b>significant</b> increase
	NPA%	76.8	81.3	85.3	not significant increase
	PPV%	61.4	68.3	74.7	<b>significant</b> increase
	NPV%	75.8	80.4	84.5	<b>significant</b> increase
	K%	42.2	48.5	54.7	<b>significant</b> increase
	PR%	32.8	36.8	41.0	not significant increase
RC and GP on ECG report	PPA%	48.0	55.7	63.2	
	NPA%	73.7	78.5	82.7	
	PPV%	49.5	57.3	64.8	
	NPV%	72.5	77.3	81.6	
	K%	28.2	34.4	40.9	
	PR%	29.2	33.2	37.5	
RC and GP on CHART report	PPA%	60.6	67.7	74.3	<b>significant</b> increase
	NPA%	75.6	80.4	84.5	not significant increase
	PPV%	59.5	66.7	73.3	<b>significant</b> increase
	NPV%	76.4	81.1	85.2	not significant increase
	K%	41.5	47.9	54.4	<b>significant</b> increase
	PR%	33.1	37.3	41.6	not significant increase

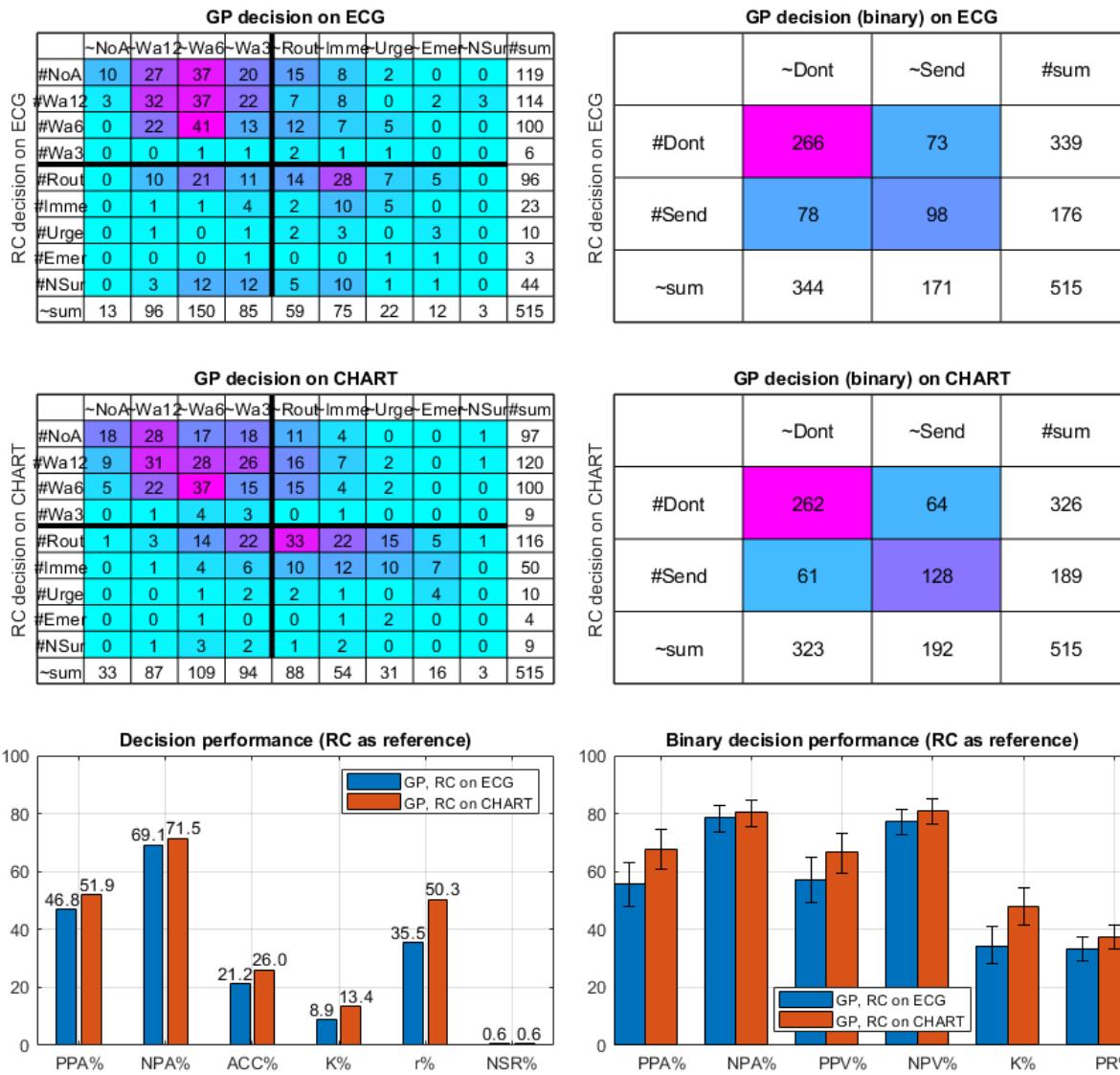


GP decision on ECG											GP decision (binary) on ECG			
ORC decision on ECG	-NoA	-Wa12	-Wa6	-Wa3	-Rout	-Imme	-Urge	-Eme	-NSur	#sum		~Dont	~Send	#sum
#NoA	9	30	29	18	13	7	0	1	1	108				
#Wa12	4	29	33	17	18	9	3	0	0	113				
#Wa6	1	29	47	19	10	9	2	0	0	117				
#Wa3	0	1	4	1	1	3	0	1	0	11				
#Rout	0	6	28	20	9	27	12	5	1	108				
#Imme	0	4	5	10	10	12	6	4	1	52				
#Urge	0	0	0	1	0	2	0	2	0	5				
#Eme	0	0	0	1	0	2	0	0	0	3				
#NSur	1	8	12	0	3	7	1	0	0	32				
~sum	15	107	158	87	64	78	24	13	3	549				

GP decision on CHART											GP decision (binary) on CHART			
ORC decision on CHART	-NoA	-Wa12	-Wa6	-Wa3	-Rout	-Imme	-Urge	-Eme	-NSur	#sum		~Dont	~Send	#sum
#NoA	16	28	16	22	9	6	0	0	1	98				
#Wa12	8	35	33	16	20	2	1	0	0	115				
#Wa6	8	19	42	29	13	5	0	0	1	117				
#Wa3	2	1	1	3	4	2	0	0	0	13				
#Rout	1	6	21	20	28	24	15	7	1	123				
#Imme	0	0	4	6	13	15	12	7	0	57				
#Urge	0	0	0	1	3	1	2	2	0	9				
#Eme	0	0	0	0	0	1	1	0	0	2				
#NSur	1	3	2	3	4	2	0	0	0	15				
~sum	36	92	119	100	94	58	31	16	3	549				



**Figure 14- Performance analysis of GP decision versus ORC decision, detailed comparison on left, binary comparison on right**



**Figure 15- Performance analysis of GP decision versus RC decision, detailed comparison on left, binary comparison on right**

### Conclusions:

- In case of CHART the decision agreement between GP and ORC are significantly higher compared to ECG. In case of CHART the PPA, PPV, NPV, Kappa show significantly higher value compared to ECG.
- This suggests that when using CHART, a GP can benefit significantly. Furthermore, the ORC role is less necessary when GP uses CHART instead of ECG in the decision procedure.
- The agreement comparison between GP and RC is similar to GP and ORC.

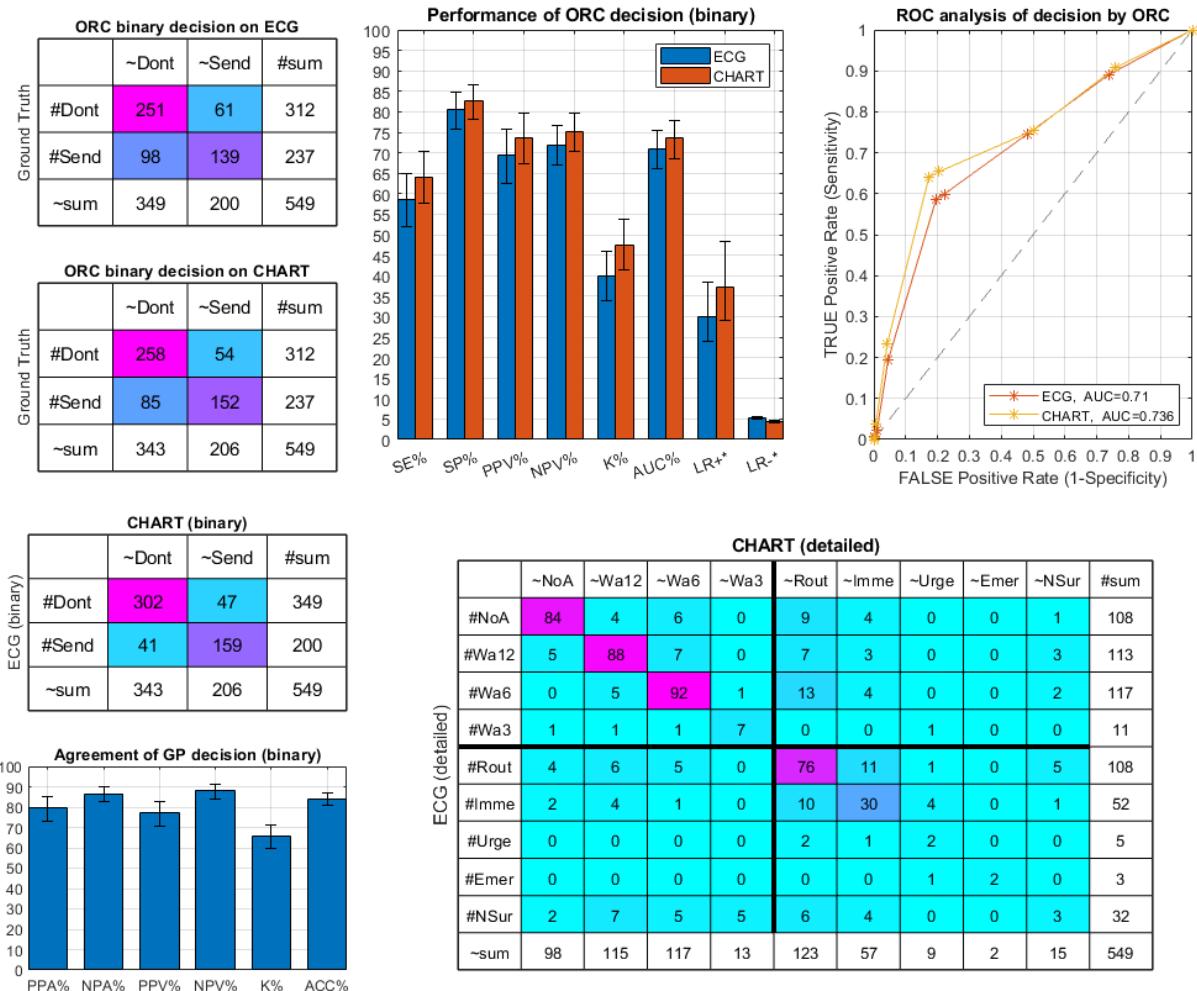


### 8.3.3 ORC decision compared to ground truth

This section presents the ORC decision performance decided on ECG and CHART reports where the reference decision is the consensus-based ground truth.

There are 549 patients having ORC decision and ground truth.

Confusion matrixes and performance are plotted in Figure 16.



**Figure 16- Performance of ORC decision by consensus ground truth (top part), agreement analysis between ORC decision on ECG and CHART reports (bottom part)**



Table 12- Performances

	Metric	Performance			Hypothesis test with 95% CI
		Lower conf. %	Observed Value %	Upper conf. %	CHART compare to ECG
ECG report	SE%	52.1	58.7	65.0	
	SP%	75.6	80.5	84.7	
	PPV%	62.6	69.5	75.8	
	NPV%	66.9	71.9	76.6	
	K%	33.8	39.9	46.0	
	AUC%	66.1	71.0	75.4	
	LR+	2.40	3.00	3.83	
	LR-	0.55	0.51	0.49	
	PR%	32.4	36.4	40.6	
CHART report	SE%	57.7	64.1	70.2	significant increase
	SP%	78.0	82.7	86.7	not significant decrease
	PPV%	67.2	73.8	79.7	significant increase
	NPV%	70.3	75.2	79.7	significant increase
	K%	41.4	47.6	53.8	<b>significant</b> increase
	AUC%	68.6	73.6	77.9	not significant increase
	LR+	2.92	3.71	4.83	not significant increase
	LR-	0.46	0.43	0.41	not significant increase
	PR%	33.5	37.5	41.7	not significant increase

## Conclusions:

- The agreement of ORC decision between ECG and CHART report is smaller than deviation observed at GP level. There are 47 patients send by CHART and don't sent by ECG, and 41 send by ECG and don't sent by CHART.
- The Kappa have significant increase in case of CHART report compared to ECG report. The rest of performance metrics show not significant increases.
- ORC decision show higher agreement on the detailed decision too. This means that in most cases the ORC does not change a decision made from ECG when shifting to CHART report. The reason for this is likely resulting the protocol where the same ORC reads ECG and then right after CHART report, whereas in the case of GP evaluation, one GP did ECG and a different GP did CHART report for the same patient.



- In future, a different ORC should have evaluated CHART compared to ECG (Cardiologist 1, ECG on patient #1 and CHART on Patient #2, whereas Cardiologist 2, did CHART on Patient #1, and ECG on patient #2).

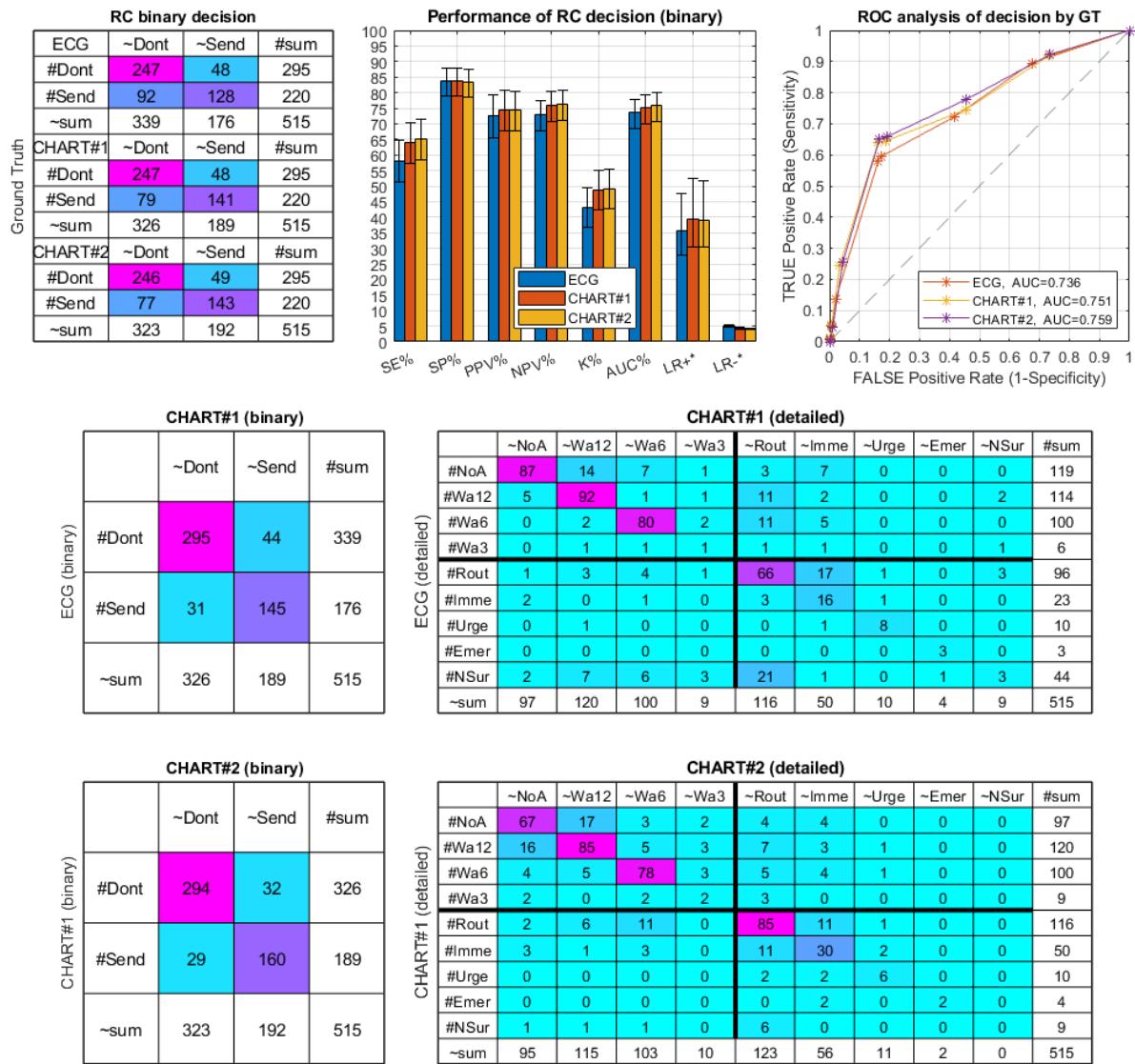


### 8.3.4 RC decision compared to ground truth

This section presents the RC decision performance based on ECG, CHART#1 and CHART#2 reports where the reference decision is the consensus-based ground truth.

There are 515 patients having RC decision and ground truth.

Confusion matrixes and performance are plotted in Figure 17.



**Figure 17- Performance of RC decision by consensus ground truth (top part), confusion analysis between RC decision on ECG-CHART#1 (middle) and CHART#1-CHART#2 reports (bottom part)**



Table 13- Performances

	Metric	Performance			Hypothesis test with 95% CI
		Lower conf. %	Observed Value %	Upper conf. %	CHART#1 compare to ECG CHART#2 compare to CHART#1
ECG report	SE%	51.4	58.2	64.8	
	SP%	79.0	83.7	87.8	
	PPV%	65.5	72.7	79.2	
	NPV%	67.8	72.9	77.5	
	K%	36.8	43.0	49.5	
	AUC%	68.6	73.6	77.9	
	LR+	2.77	3.58	4.75	
	LR-	0.53	0.50	0.48	
	PR%	30.1	34.2	38.5	
CHART#1 report	SE%	57.4	64.1	70.4	significant increase
	SP%	79.0	83.7	87.8	not significant decrease
	PPV%	67.8	74.6	80.6	not significant decrease
	NPV%	70.7	75.8	80.3	not significant decrease
	K%	42.4	48.7	55.2	significant decrease
	AUC%	69.9	75.1	79.5	not significant decrease
	LR+	3.05	3.94	5.23	not significant decrease
	LR-	0.45	0.43	0.41	not significant decrease
	PR%	32.5	36.7	41.0	not significant decrease
CHART#2 report	SE%	58.3	65.0	71.3	<b>significant</b> increase
	SP%	78.6	83.4	87.5	not significant increase
	PPV%	67.7	74.5	80.5	not significant increase
	NPV%	71.1	76.2	80.7	not significant increase
	K%	42.8	49.2	55.6	not significant decrease
	AUC%	70.9	75.9	80.2	not significant increase
	LR+*	3.04	3.91	5.18	not significant increase
	LR-*	0.45	0.42	0.40	not significant increase
	PR%	33.1	37.3	41.6	not significant increase

**Conclusions:**

- RC decision based on CHART report produce slightly better accuracy compared to ECG. CHART#1 or #2 has borderline significant increase of sensitivity with similar specificity.
- One possible reason of RC was not benefit from CHART as GP and ORC is the incorrect interpretation of CHART findings: might be RC interpret some mild CHART+ abnormality as serious problem which causes similar false positive referral decisions as at ECG.
- CHART#1 and CHART#2 shows the same decision performance without any significant changes.



## 8.4 Diagnosis and Comparison Data

### 8.4.1 GP and RC Diagnosis

GP#1 diagnoses the ECG report, and GP#2 the CHART#1 report, whereas the RC diagnoses both ECG (first) then the CHART#2 report – altogether a set of three complete diagnosis. There are 400 patients having GP and RC diagnostic forms.

In the diagnosis questionnaires there are seven main diagnostic points:

*Table 14- Seven diagnostic points used in GP and RC diagnosis forms*

Diagnostic points	Description	Supported by	
		ECG	CHART+
<b>Arrhythmia, Blocks</b>	Rhythm Problem, Premature Complex Beats, Heart Block, Axis Deviation Other: Pacemaker, etc.	Yes Primarily on ECG	No
<b>LV Dysfunction</b>	LV Systolic and Diastolic Dysfunction	No	Yes HART finding classify
<b>Ischemia</b>	ST/STT deviation QT interval T-wave abnormality Wall motion abnormality Myocardial Infarction	Yes Primarily on ECG	No Additional WMA <sup>23</sup>
<b>Size Abnormality</b>	LH Hypertrophy LA Enlargement RA Enlargement RV Enlargement Cardiomyopathy	Yes ECG indicate these abnormalities	No Classified by HART findings with different criteria
<b>Valve Disease</b>	Aortic Stenosis and Insufficiency Mitral Stenosis and Insufficiency Tricuspid Insufficiency	No	Yes HART findings

<sup>23</sup> over time we expect CHART to diagnose ischemia better than ECG as in this study and in training the only ischemia CHART learned, was learned mainly from ECG... however being a physiological bio-signal it is expected to get more accurate.



	Pulmonary Insufficiency Pulmonary Hypertension		
<b>PCG-Murmur Presence</b>	Q-S1, Q-S2 intervals S1 splitting LVET, Systolic and diastolic murmur Ejection sound Third and Fourth sound	No	Yes Primarily on CHART report / PCG
<b>MCG-STI Abnormality</b>	EMAT, PEP, LVET interval ICVT, IVRT intervals SPI and MPI indexes	No	Yes Primarily on CHART report / MCG

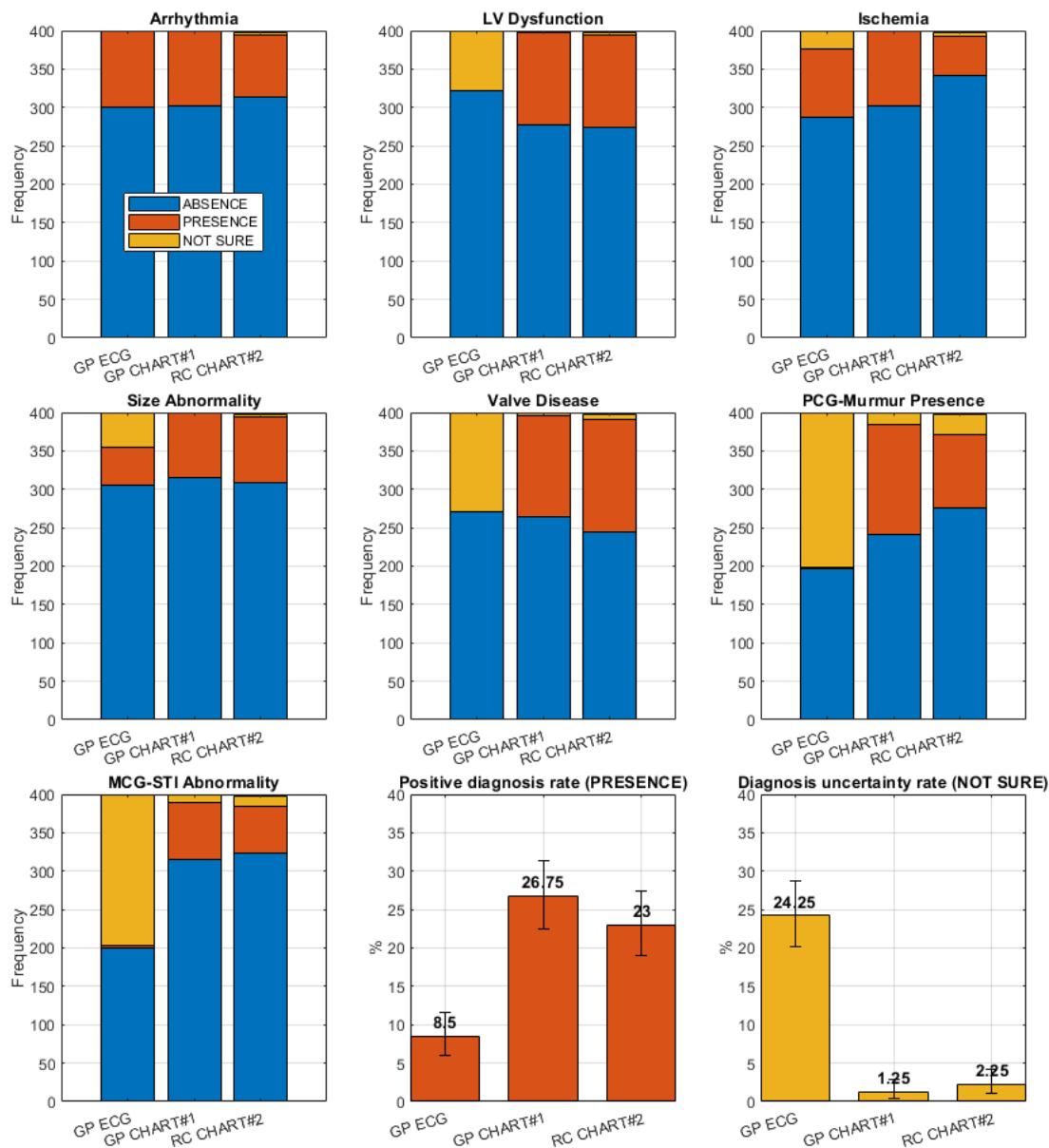


Figure 18- GP and RC diagnosis summary



Figure 18 summarize the three types of diagnosis, where the three color represent the rate of “absence”, “presence” and “not-sure” answers.

#### **Point based conclusions:**

- The diagnostic of Arrhythmia-Blocks shows difference between GP and RC but not between ECG and CHART report.
- The diagnosis of LV dysfunction shows differences between ECG and CHART. The ECG based diagnosis is full of Uncertain, while CHART based diagnosis is mostly Certain.
- The diagnosis of Ischemia shows smaller difference between GP and RC and between ECG and CHART as well. RC diagnosis includes much less positive Ischemia, which indicates that the RC uses stricter thresholds than GP.
- The diagnostic of Size abnormality shows bigger difference between ECG and CHART. The ECG based diagnosis shows half of them as Uncertain, while CHART based diagnosis is mostly rather Certain.
- The diagnosis of Valve disease, PCG-murmur and MCG-STI shows differences between ECG and CHART. The ECG based diagnosis are full of Uncertain, while CHART based diagnosis is mostly Certain.

#### **Overall Conclusions:**

- The GP's positive (number of presence) diagnosis rate is significantly increased, from 8.5% to 26.7% by CHART#1, while the uncertainty rate (number of not sure) is significantly decreased, from 24% to 1.7%.
- The RC's positive diagnosis rate on CHART#2 is smaller than GP's positive diagnosis on CHART#1, yet more similar when compared to ECG. RC generally diagnose less positive (23%) compared to GP (26%).
- The RC's uncertainty rate on CHART#2 is similarly small as GP on CHART#1.
- The point-based conclusions are expected, because:
  - The two typical ECG-based diagnosis - Arrhythmia and Ischemia – show similarity between ECG and CHART, but with much less Uncertain.
  - The Size abnormality is mixed, one half solved by ECG, but the certainty of the other half is increased by CHART (decreased Uncertainty).
  - The rest of diagnostic points are more supported fully by CHART, which dispels all the uncertainties observed with use of ECG.

##### ***8.4.1.1 Subgroup analysis***

We can categorize the patients based on RC diagnosis into four subgroups: (The criteria are similar to subgroup analysis in section 8.3.1.1)

- g:None – patient has no abnormality according to the seven diagnostic points
- g:ECG – patient has primarily ECG based abnormality according to Table 14
- g:CHART+ – patient has primarily CHART+ based abnormality according to Table 14
- g:Both – patient has both ECG and CHART+ based abnormality according to Table 14



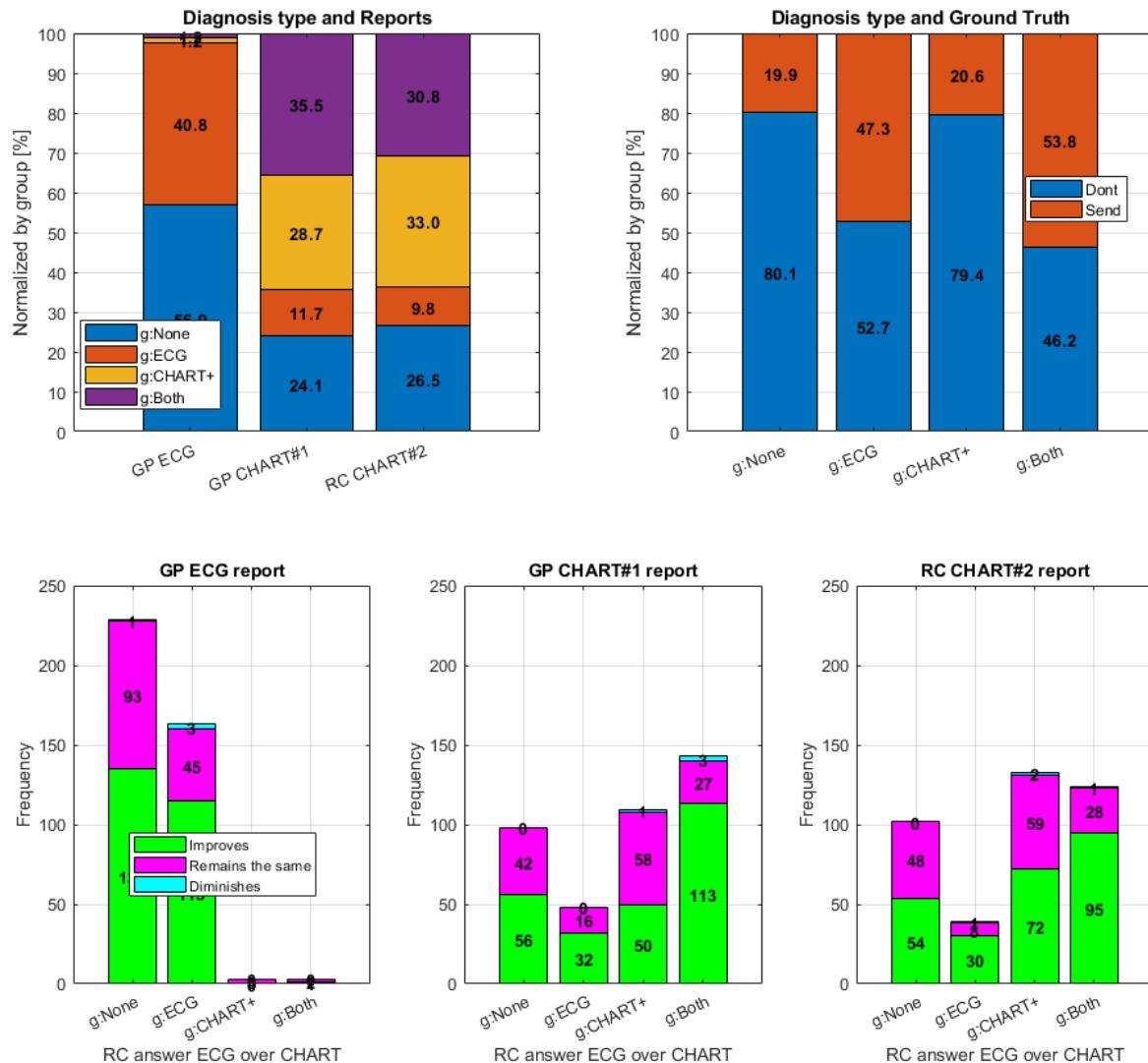
The distribution results between the three types of diagnosis, the four subgroups, the binary ground truth decision, and the RC answer on ECG over CHART (the question #1 in Table 15) are demonstrated in Figure 19.

### **Conclusions:**

- GP and RC diagnosis using the CHART reports are similar based on the four subgroups.
- The ECG based diagnosis miss the CHART+ for Both subgroups, and this is expected.
- The ECG over CHART evaluation (Improve, Remains, Diminishes) shows consistent distribution in the four subgroups. This means the benefit of using CHART (the two-thirds improvement) is not dependent on whether an ECG or CHART+ diagnosable abnormality is present.
- The ECG abnormality more often indicate a “Send” referral decision (50%) compared to CHART+ abnormality (20%)<sup>24</sup>. However, in case of co-presence, i.e. g:Both subgroups, the “Send” decision reached 57%.

---

<sup>24</sup> Note: This is the suspected as the reason for the higher ECG specificity in ORC and RC comparison, in sections 8.3.3 and 8.3.4.



**Figure 19- GP and RC diagnosis in function of four type diagnosis (left top), with mix of ground truth decision (right top), and with mix of RC answer over CHART benefit (bottom graphs)**



#### 8.4.2 Comparison ECG over CHART

ORC and RC were asked to compare ECG and CHART reports. The main questions and the answer distribution expressed in percentage in the following table. The lower and upper value are also listed according to 95% CI.

There are 550 ORC and 522 RC<sup>25</sup> record from this questionnaire form.

**Table 15- ORC and RC statistic on ECG over CHART Questionnaire**

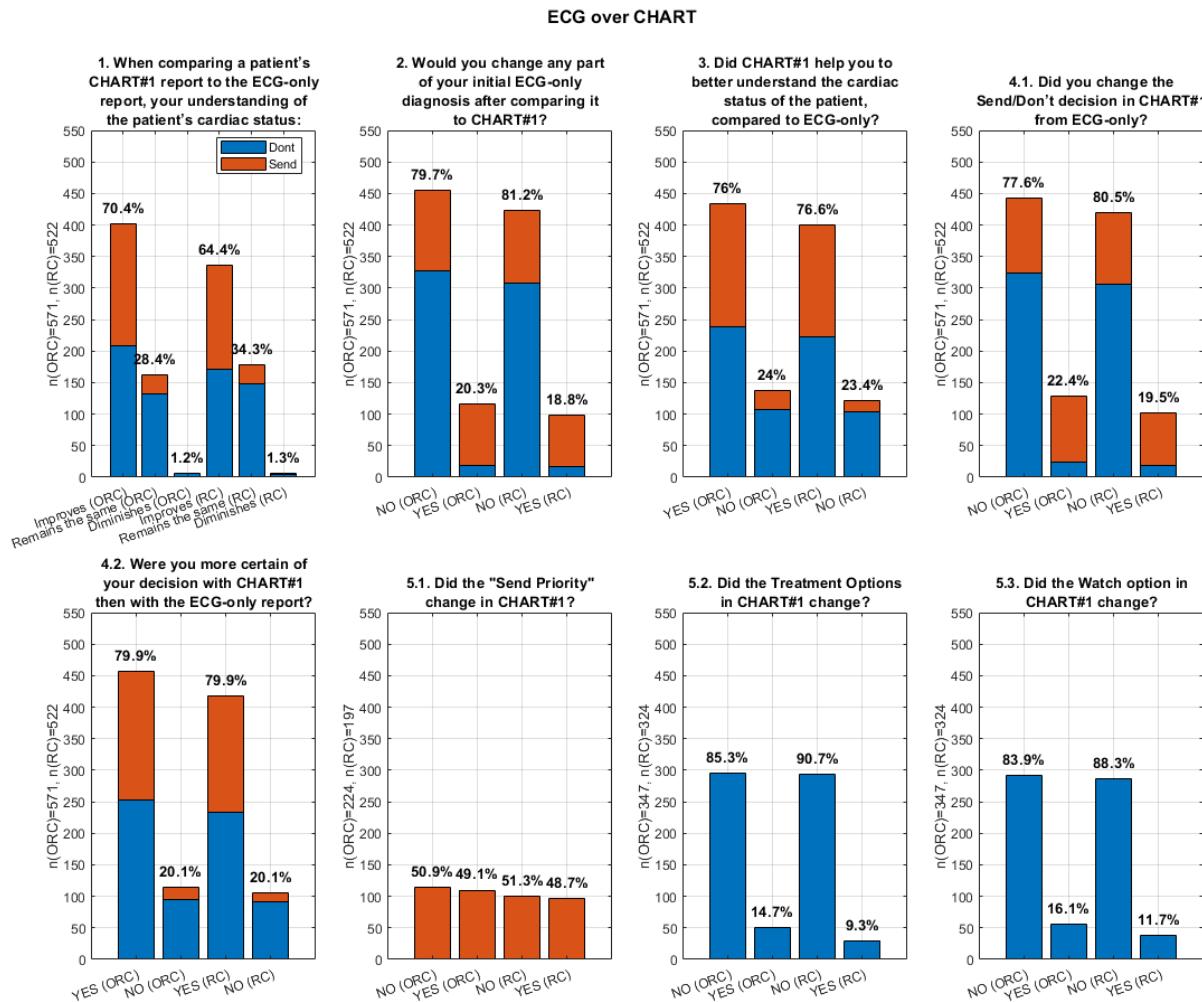
Question and answers	Lower conf. %	Observed Value %	Upper conf. %	Question and answers	Lower conf. %	Observed Value %	Upper conf. %
<b>#1. When comparing a patient's CHART#1 report to the ECG-only report, your understanding of the patient's cardiac status:</b>				<b>#2. Would you change any part of your initial ECG-only diagnosis after comparing it to CHART#1?</b>			
Improves (ORC)	66.5	70.4	74.1	NO (ORC)	76.1	79.7	82.9
Remains the same (ORC)	24.7	28.4	32.3	YES (ORC)	17.1	20.3	23.9
Diminishes (ORC)	0.5	1.2	2.5	NO (RC)	77.6	81.2	84.5
Improves (RC)	60.1	64.4	68.5	YES (RC)	15.5	18.8	22.4
Remains the same (RC)	30.2	34.3	38.5				
Diminishes (RC)	0.5	1.3	2.7				
<b>#3. Did CHART#1 help you to better understand the cardiac status of the patient, compared to ECG-only?</b>				<b>#4.1. Did you change the Send/Don't decision in CHART#1 from ECG-only?</b>			
YES (ORC)	72.3	76.0	79.5	NO (ORC)	73.9	77.6	80.9
NO (ORC)	20.5	24.0	27.7	YES (ORC)	19.1	22.4	26.1
YES (RC)	72.8	76.6	80.2	NO (RC)	76.8	80.5	83.8
NO (RC)	19.8	23.4	27.2	YES (RC)	16.2	19.5	23.2
<b>#4.2. Were you more certain of your decision with CHART#1 then with the ECG-only report?</b>				<b>#5.1. Did the "Send Priority" change in CHART#1?</b>			
YES (ORC)	76.3	79.9	83.1	NO (ORC)	44.1	50.9	57.6
NO (ORC)	16.9	20.1	23.7	YES (ORC)	42.4	49.1	55.9
YES (RC)	76.2	79.9	83.2	NO (RC)	44.1	51.3	58.4
NO (RC)	16.8	20.1	23.8	YES (RC)	41.6	48.7	55.9
<b>#5.2. Did the Treatment Options in CHART#1 change?</b>				<b>#5.3. Did the Watch option in CHART#1 change?</b>			
NO (ORC)	81.1	85.3	88.9	NO (ORC)	79.6	83.9	87.6
YES (ORC)	11.1	14.7	18.9	YES (ORC)	12.4	16.1	20.4

<sup>25</sup> The RC form evaluation in the study was not finished by all the participated cardiologists.



NO (RC)	87.0	90.7	93.7	NO (RC)	84.3	88.3	91.6
YES (RC)	6.3	9.3	13.0	YES (RC)	8.4	11.7	15.7

This is graphically presented in Figure 20.



**Figure 20- Results of ECG over CHART related questions assigned to ORC and RC**

## **Conclusions:**

- The results of ORC and RC are very similar in all of the questions.
- Question #1: CHART improves the understanding of patient cardiac status in 64-70%, in sync with the better understand (Question #3) and more certain question (Question #4.2).
- Question #1: CHART diminishes the understanding of patient cardiac status only in 1.2%.
- Question #4.1: ORC and RC change their ECG-based referral based on the CHART report in 19-22%. This is sync with the 19% change in the initial diagnosis (Question #2).
- Question #5.1: In case of consistent "Send" decision CHART change the referral priority in 49%.



- Question #5.3: In case of consistent “Don’t” decision CHART change the watch option priority in 11-16%, in sync with the change of treatment option (Question #5.2).

#### 8.4.3 CTT Primary Hypothesis evaluation

For the information and evaluation of the CTT Primary Hypothesis, please see [Bio-Signal Verification and Validation by CUUS](#) separate document.

#### 8.4.4 Comparison CHART#2 over CHART#1

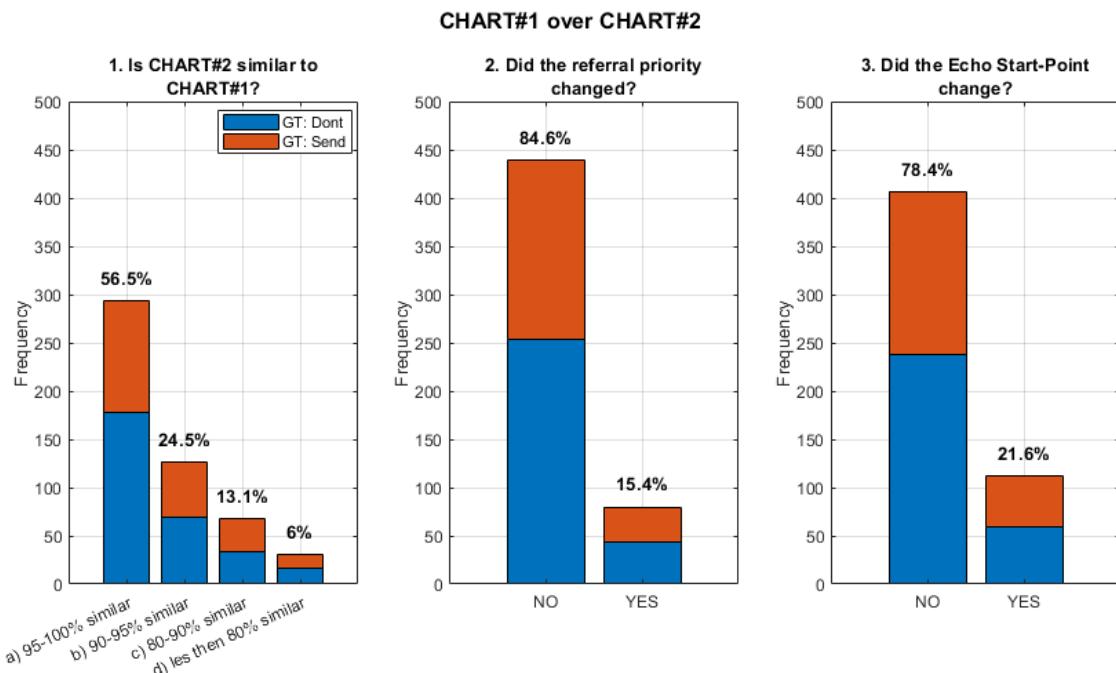
RC were asked to compare CHART#1 and CHART#2 reports. The main questions and the answer distribution expressed in percentage in the following table. The lower and upper value are also listed according to 95% CI.

There are 520 RC record from this questionnaire form.

*Table 16- RC statistics on CHART#1 over CHART#2 questionnaire*

Question and answers	Lower conf. %	Observed Value %	Upper conf. %
#1. Is CHART#2 similar to CHART#1?			
a) 95-100% similar	52.1	56.5	60.8
b) 90-95% similar	20.8	24.5	28.4
c) 80-90% similar	10.3	13.1	16.3
d) less than 80% similar	4.1	6.0	8.4
a) + b)	72.9	80.9	89.2
#2. Did the referral priority changed?			
NO	81.2	84.6	87.6
YES	12.4	15.4	18.8
#3. Did the Echo Start-Point change?			
NO	74.6	78.4	81.9
YES	18.1	21.6	25.4

This is graphically presented in Figure 21.



**Figure 21- Results of CHART#2 over CHART#2 related questions assigned to RC**

### **Conclusions:**

- The 81% of CHART#1 are similar to CHART#2 in minimum 90%, because the rest 19% are different in minimum of 10%. This confirmed by the 21% change of echo start-point (Question #3). This ratio is expected based on the preliminary reproducibility tests and the effect of time delay between the two tests.
- Less frequently changed the referral priority (15%) by RC compared to change of in Echo start-point (21%). This is expected, because starting point change does not necessarily mean the priority changed.

#### ***8.4.5 CHART as start point for ECHO***

RC were asked to evaluate the CHART report to show how it was useful in providing an appropriate start point to ECHO examination. The main question and the answer distribution expressed in percentage in the following table. The lower and upper value are also listed according to 95% CI.

There are 391 RC record from this questionnaire form.

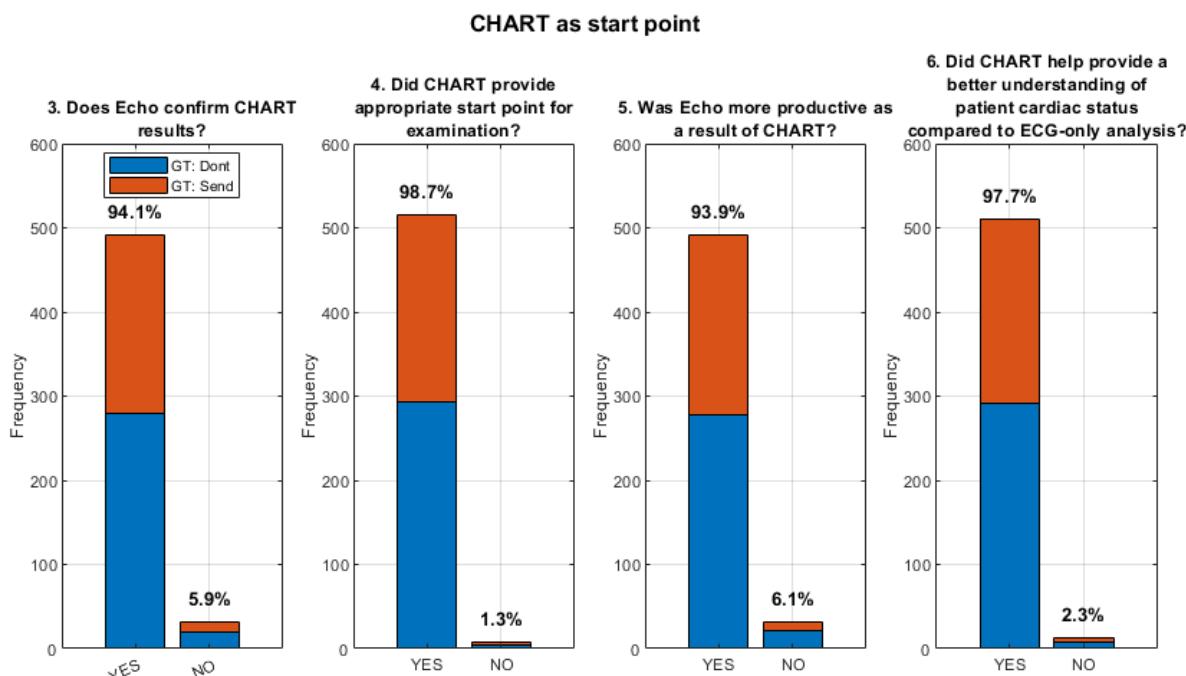
**Table 17- RC statistic on CHART as start point for ECHO**

Question and answers	Lower conf. %	Observed Value %	Upper conf. %
#3. Does Echo confirm CHART results?			
YES	91.7	94.1	95.9
NO	4.1	5.9	8.3
#4. Did CHART provide appropriate start point for examination?			



	YES	97.3	98.7	99.5
	NO	0.5	1.3	2.7
#5. Was Echo more productive as a result of CHART?				
	YES	91.5	93.9	95.8
	NO	4.2	6.1	8.5
#6. Did CHART help provide a better understanding of patient cardiac status compared to ECG-only analysis?				
	YES	96.0	97.7	98.8
	NO	1.2	2.3	4.0

This is graphically presented in Figure 22.



**Figure 22- Results of CHART as start point for ECHO related questions assigned to RC**

These results are reinforced by the HART findings performance analysis, see chapter 8.6.

### **Conclusions:**

- The 94% of patients the CHART diagnosis was confirmed by the ECHO examination (question #3).
- RC found CHART as useful to determining an appropriate start point in 98% (question #4).
- CHART provides a better understanding of patient cardiac status in 97.7% compared to ECG (question #6).

The ECHO examination was more productive as CHART results in 94%. This was expected, because ECHO provide more details and resolution, cause-effect explanations



and additional visual information compared to HART findings. This confirm that CHART report cannot be used instead of ECHO, CHART only indicate the problems and is a good start point for ECHO examination,

#### 8.4.6 Reproducibility of findings

The reproducibility is evaluated between the first and second CHART reports, generated from CTTs made on two different days. The average time delay between the two records is 10 days, for details see next subsection.

##### 8.4.6.1 Time delay statistics between CHART#2 and CHART#2

The original requirement was for a minimum of 3 days delay between the first and second CTT, i.e. CHART#1 and CHART#2 reports.

Except 39 patients the minimum 3 days is fulfilled, the average delay is 9.5 days, the maximum is 80<sup>26</sup> days, the minimum is 4 hours. The distribution of time delay expressed in days can be seen on Figure 23, colored by centers.

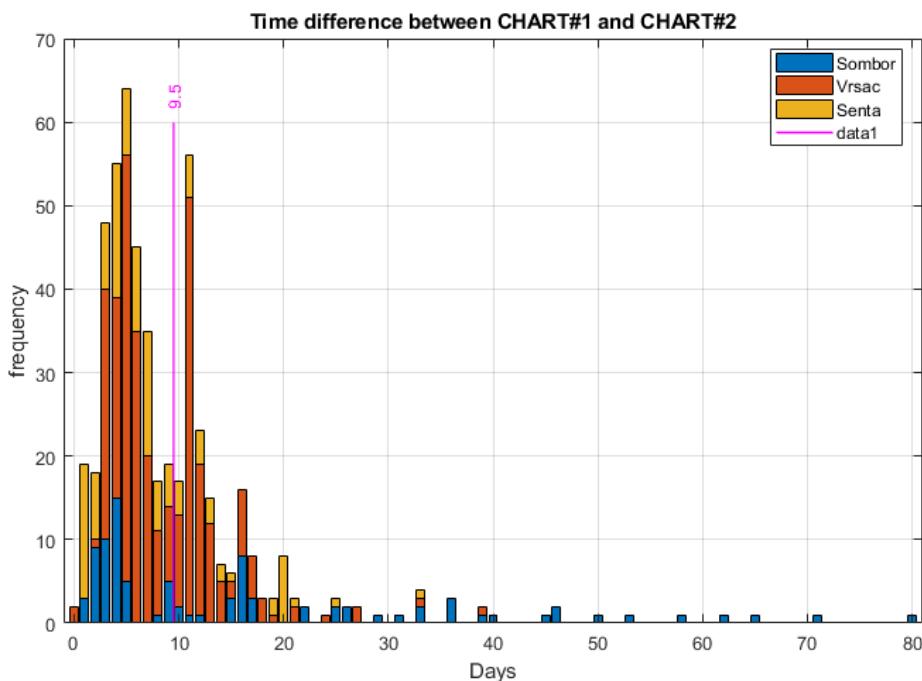


Figure 23- Time delay distribution between two CTT-CHART#2 and CHART#2 reports

##### 8.4.6.2 ECG findings

The reproducibility is analyzed separately for each finding. The following performance metrics are calculated:

<sup>26</sup> 80 Days represents patients that were tested at start of trial but never returned. They were later convinced to return to complete the protocol. This true for most of 30+ Days patients.



- PPA% and NPA% - the positive and negative percent agreement represent the ratio between agreed positive or negative cases compared to the all positive and negative cases (estimated by CHART#1)
- ACC% – accuracy is the overall agreements between CHART#1 and CHART#2 compared to all patients (the number of patients having same finding class in both reports). This is highly skewed by prevalence, and not the best metric to compare reproducibility between two findings.
- K% – the kappa statistics between the CHART#1 and CHART#2. This represents the accuracy between CHART#1 and CHART#2 independent from finding prevalence, thus, it is a better metric to compare reproducibility between two findings
- r% – correlation coefficient, similar to K%, but it takes into account the distance between non-agreed classes. For example, the distance between Normal and Mild is smaller than between Normal and Abnormal at HART findings.
- PREV% – express the positive rate in CHART#1 (considered as reference)
- PR% – positive rate in CHART#2 (considered as test)

**Table 18- List of metrics for CPA-ECG and Predicate ECG (PECG) findings**

ECG Finding	PPA%		NPA%		ACC%		K%		r%		PR CHART#1 (PREV%)		PR CHART#2 (PR%)	
	CPA	PECG	CPA	PECG	CPA	PECG								
Rhythm	30.1	29.9	97.6	95.5	87.7	81.8	39.2	22.8	43.8	27.3	10.6	12.7	11.2	13.2
Pacemaker		33.3	99.6	99.7	99.6	99.3	0.0	33.0		33.0	0.0	0.5	0.4	0.5
PVC	56.4	20.0	94.4	98.4	91.5	97.8	45.4	12.3	45.8	13.1	7.5	0.9	9.4	1.7
PAC	49.6	27.8	88.1	98.8	79.6	96.6	38.9	31.6	39.0	32.3	22.2	3.1	20.2	2.1
PR interval	68.7	45.4	96.5	98.4	91.3	94.9	68.1	62.8	68.2	61.4	15.6	7.9	15.2	6.9
QT interval	65.2	62.1	97.7	98.9	94.8	96.1	66.1	58.9	66.3	59.0	8.9	5.0	7.9	5.2
Axis	69.6	86.5	98.7	99.1	96.0	94.9	78.1	70.8	77.2	70.9	10.2	10.0	10.0	9.3
BBB	67.9	70.4	99.1	99.0	95.8	97.3	76.6	80.4	81.3	84.1	9.6	7.4	9.6	7.4
Other Block	37.5	72.7	99.1	99.3	97.3	97.8	70.7	70.0	70.2	69.8	4.6	3.8	5.0	4.0
MI	35.7	30.9	98.7	98.4	91.9	90.9	56.3	44.3	57.6	48.2	9.6	8.9	10.2	8.4
Ischemia	75.9	57.5	94.1	95.0	92.1	89.9	63.8	55.1	64.2	55.3	11.2	13.8	13.7	12.2
ST deviation	35.9	57.5	97.2	97.5	93.3	93.5	61.6	59.1	61.6	62.0	8.9	9.3	10.4	8.1
Twave abnormal	68.3	43.3	91.3	89.2	87.7	82.1	56.3	32.3	56.4	32.3	15.8	15.5	18.1	15.8
Atrial Enlargement	9.1	8.3	99.7	99.9	97.7	99.0	24.0	24.7	25.9	28.4	2.1	1.0	1.0	0.3
Hypertrophy	35.7	56.5	99.7	99.5	98.5	97.1	78.2	59.0	78.0	58.9	4.1	4.0	3.3	3.4
PRWP	50.0	48.4	96.0	97.3	91.9	94.7	47.9	46.4	47.9	46.4	8.9	5.3	8.1	5.2
ECG quality	0.0	9.1	99.4	97.8	98.7	91.2	-0.5	14.5	-0.6	14.4	0.4	5.7	1.0	5.2
ECG Summary	81.4	72.6	79.1	86.3	68.4	70.5	51.9	52.3	65.1	62.9	61.9	50.5	63.8	48.1
<b>mean</b>	<b>50.4</b>	<b>46.2</b>	<b>95.5</b>	<b>97.1</b>	<b>91.9</b>	<b>92.5</b>	<b>51.3</b>	<b>46.1</b>	<b>55.7</b>	<b>47.8</b>	<b>11.8</b>	<b>9.2</b>	<b>12.1</b>	<b>8.7</b>

Figure 24 and Figure 25 shows these percentage performances in the first graph. The rest of graph shows the confusion matrix between CHART#1 and CHART#2 separately for each ECG findings.

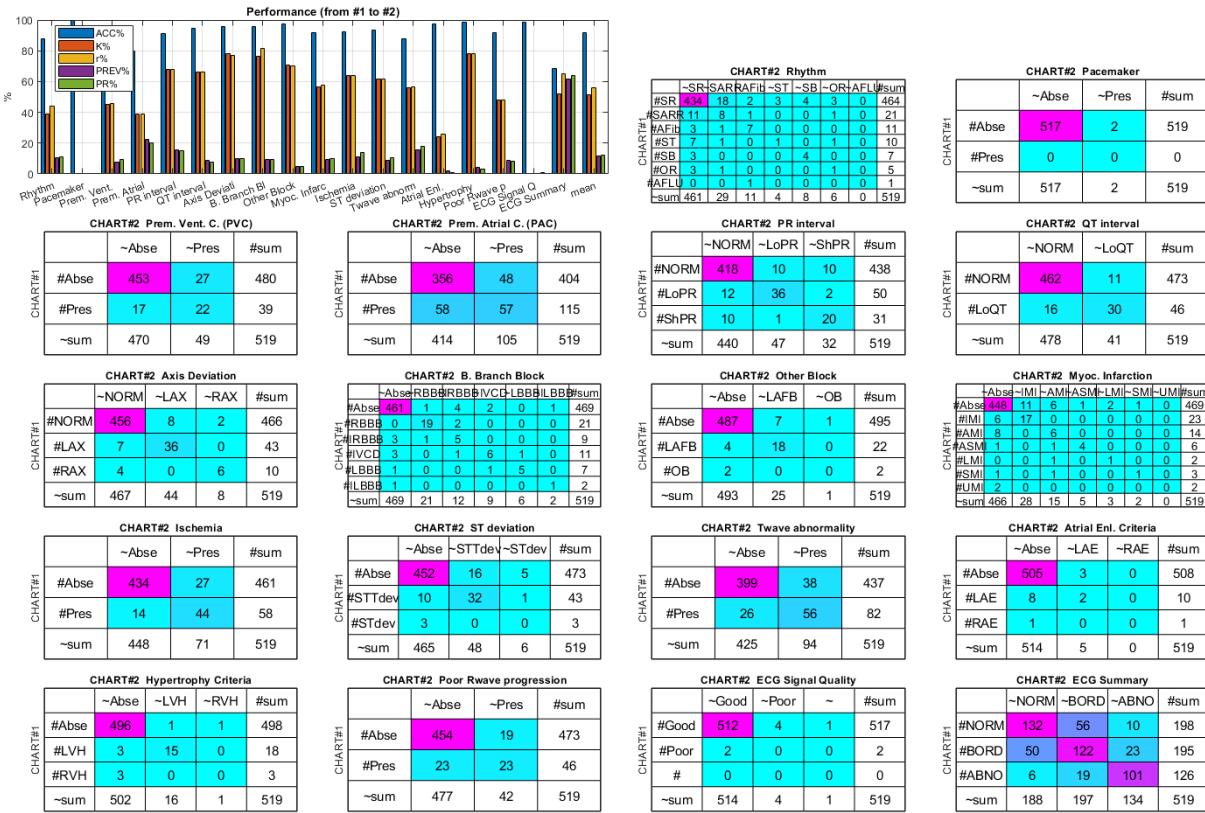


Figure 24- CPA ECG finding reproducibility analysis between CHART#1 and CHART#2

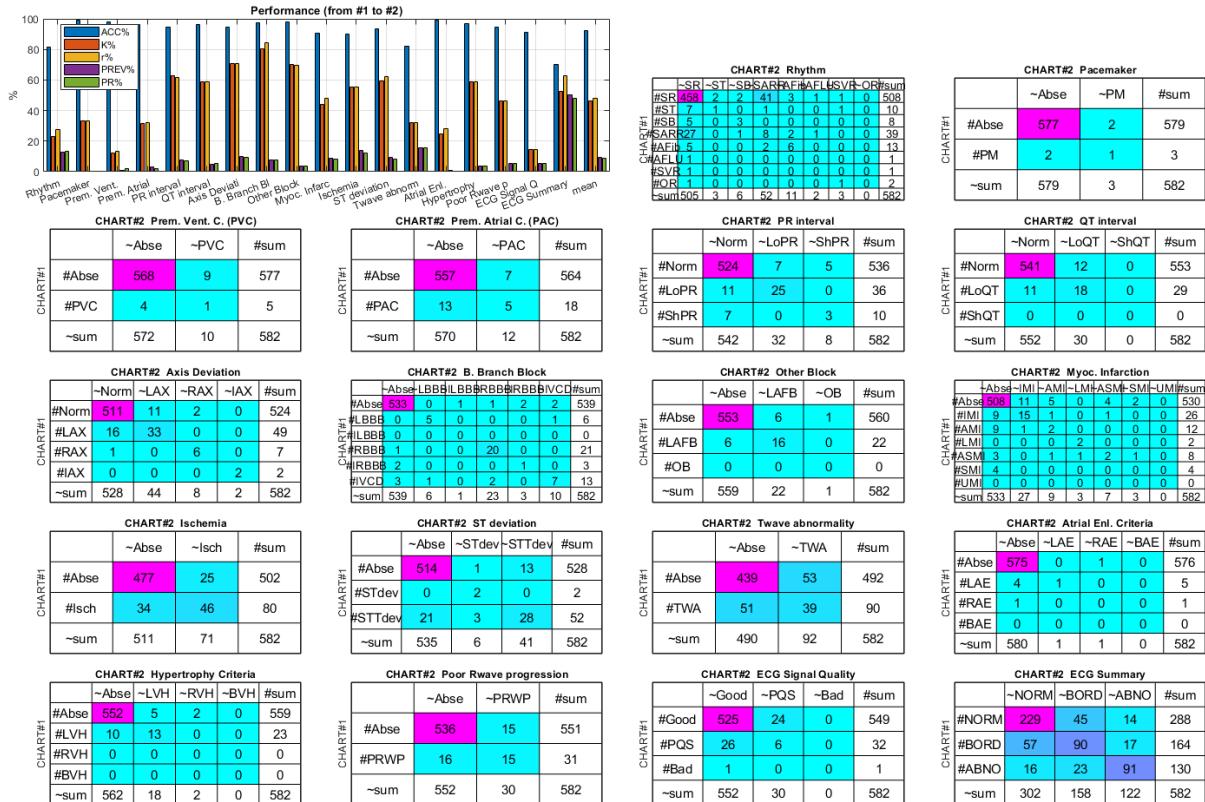


Figure 25- Predicate ECG finding reproducibility analysis between CHART#1 and CHART#2

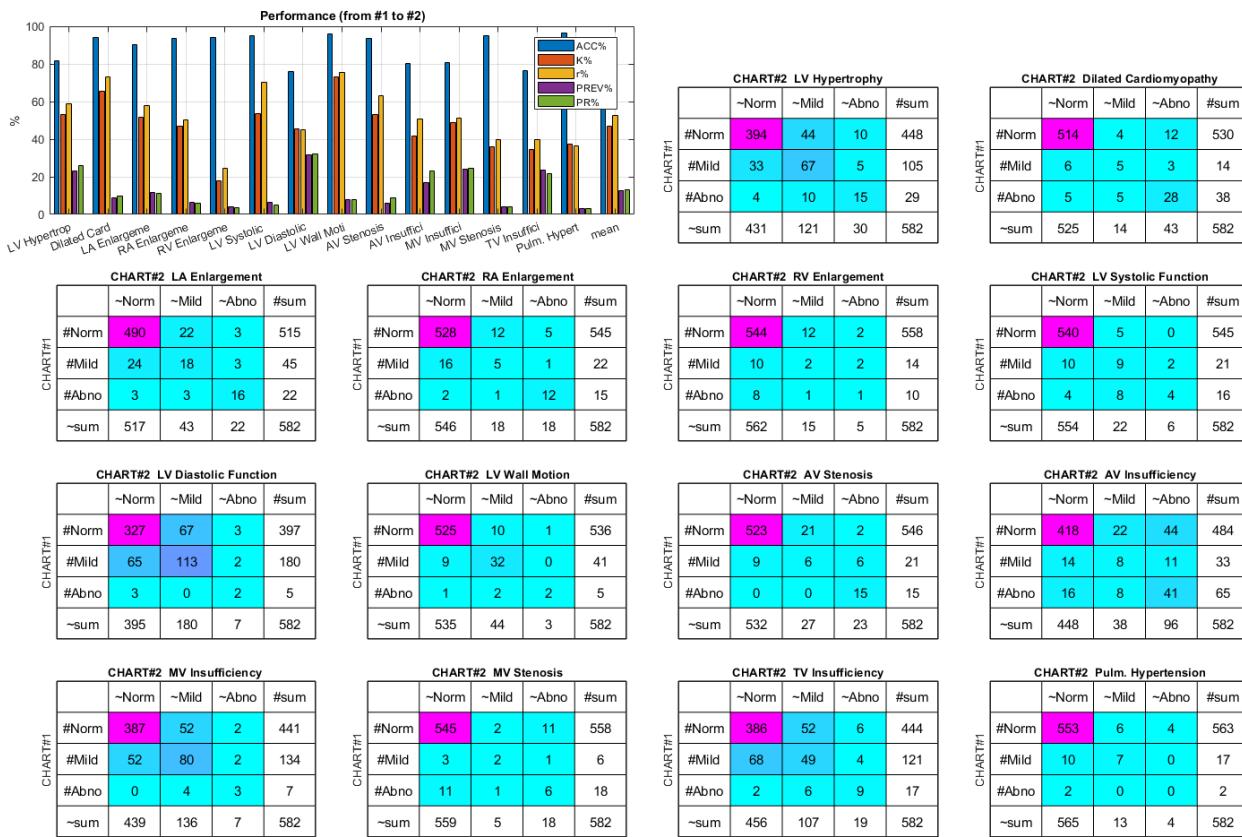


### 8.4.6.3 HART finding

This is same analysis as in ECG findings (section 8.4.5.2) applied to 14 HART findings.

**Table 19- List of metrics for CPA-HART**

HART Finding	PPA%	NPA%	ACC%	K%	r%	PR CHART#1 (PREV%)	PR CHART#2 (PR%)
<b>LV Hypertrophy</b>	62.1	92.6	81.8	53.3	58.8	23.0	26.0
<b>Dilated Cardiomyopathy</b>	76.3	97.1	94.0	65.3	73.3	8.9	9.8
<b>LA Enlargement</b>	66.2	97.0	90.0	51.8	57.8	11.5	11.2
<b>RA Enlargement</b>	65.7	97.9	93.6	46.8	50.5	6.4	6.2
<b>RV Enlargement</b>	17.5	98.4	94.0	18.1	24.7	4.1	3.4
<b>LV Systolic Function</b>	43.6	99.4	95.0	53.4	70.4	6.4	4.8
<b>LV Diastolic Function</b>	51.6	90.8	76.0	45.5	45.0	31.8	32.1
<b>LV Wall Motion</b>	59.1	98.9	96.1	73.3	75.4	7.9	8.1
<b>AV Stenosis</b>	87.5	97.2	93.5	53.3	63.3	6.2	8.6
<b>AV Insufficiency</b>	66.2	87.9	80.2	41.5	50.7	16.8	23.0
<b>MV Insufficiency</b>	53.0	93.5	80.8	48.6	51.3	24.2	24.6
<b>MV Stenosis</b>	37.5	97.8	95.0	36.2	39.9	4.1	4.0
<b>TV Insufficiency</b>	51.1	92.6	76.3	34.7	39.5	23.7	21.7
<b>Pulm. Hypertension</b>	18.4	98.8	96.2	37.2	36.6	3.3	2.9
<b>mean</b>	54.0	95.7	88.7	47.1	52.7	12.7	13.3



**Figure 26- HART finding reproducibility analysis between CHART#2 and CHART#2**



#### 8.4.6.4 Performance Comparison

The reproducibility performances are compared between the three set of findings in Figure 27.

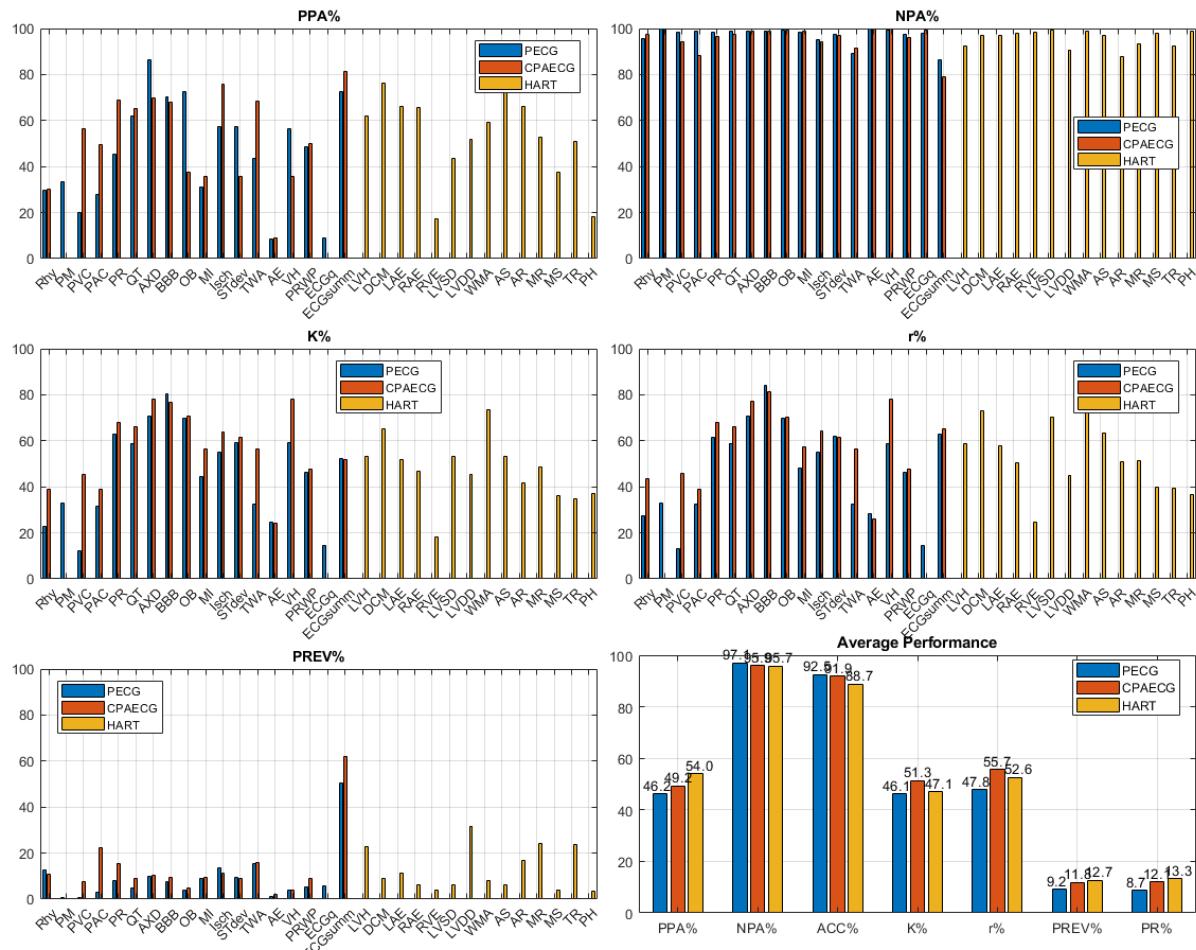


Figure 27- Comparing reproducibility performance of Glasgow ECG, CPA-ECG and HART findings

#### 8.4.6.5 Reproducibility in function of time delay

Reproducibility can be simply measured as the number of different findings between CHART#1 and CHART#2. This is expressed by change rate on the vertical axis Figure 28:

Change rate = number of different finding / all findings

The horizontal axis is the time delay between CHART#1 and CHART#2 in days. The colored cell represents the frequency.

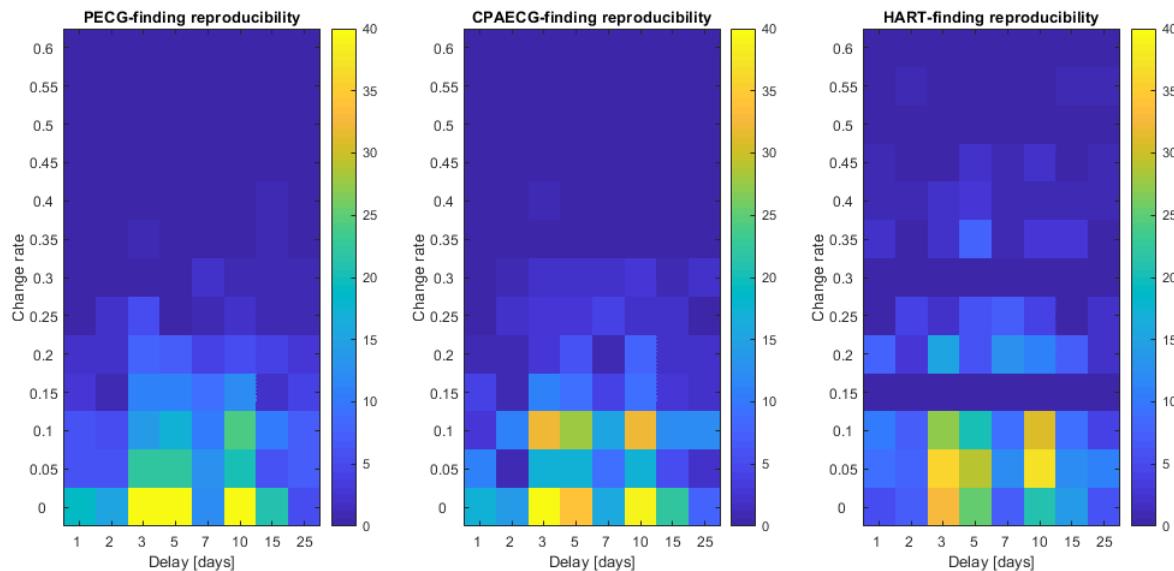


Figure 28- ECG and HART findings reproducibility in function of time delay between CHART#1 and CHART#2

#### 8.4.6.6 Conclusion

- The CPA-ECG is similar to the predicate ECG in reproducibility performances, however CPA-ECG has slightly higher performance on average. There are a few findings having significantly higher prevalence with CPA (PVC, PAC), which was expected, and the reproducibility is smaller for these findings in accordance with prevalence. The occurrence of premature beats is variable, and the reproducibility is not expected as much as other findings.
- The HART findings show slightly smaller reproducibility performances compared to ECG on average: HART has ACC=89%, K=47%, r=52%, ECG has ACC=92%, K=51%, r=56%. This 5-10% loss of reproducibility is expected for HART findings, because HART's classifiers include the use of PCG and MCG signals, where MCG is less reproducible than ECG, while PCG is less reproducible than MCG.
- The reproducibility as a function of time delay is similar between ECG and HART findings. This confirms that the variability depends on the bio-signals used for the classification of ECG and HART findings.
- There are differences between ECG findings and between HART findings in reproducibility:
  - ECG – Axis deviation and Heart blocks show the highest reproducibility.
  - ECG - Pacemaker, Quality, Atrial Enlargement and Rhythm show the lowest reproducibility. Our interpretation is: no patient having truth pacemaker, so this was a false positive; the quality is normally varied independently from heart functioning; atrial enlargement criteria should be reconsidered; the rhythms are likely to vary after a few days.
  - HART – Dilated Cardiomyopathy and LV Wall Motion shows the highest reproducibility – the classifier relies more on ECG signal for this finding
  - HART - RV Enlargement shows the lowest reproducibility – the classifier relies more on PCG signal



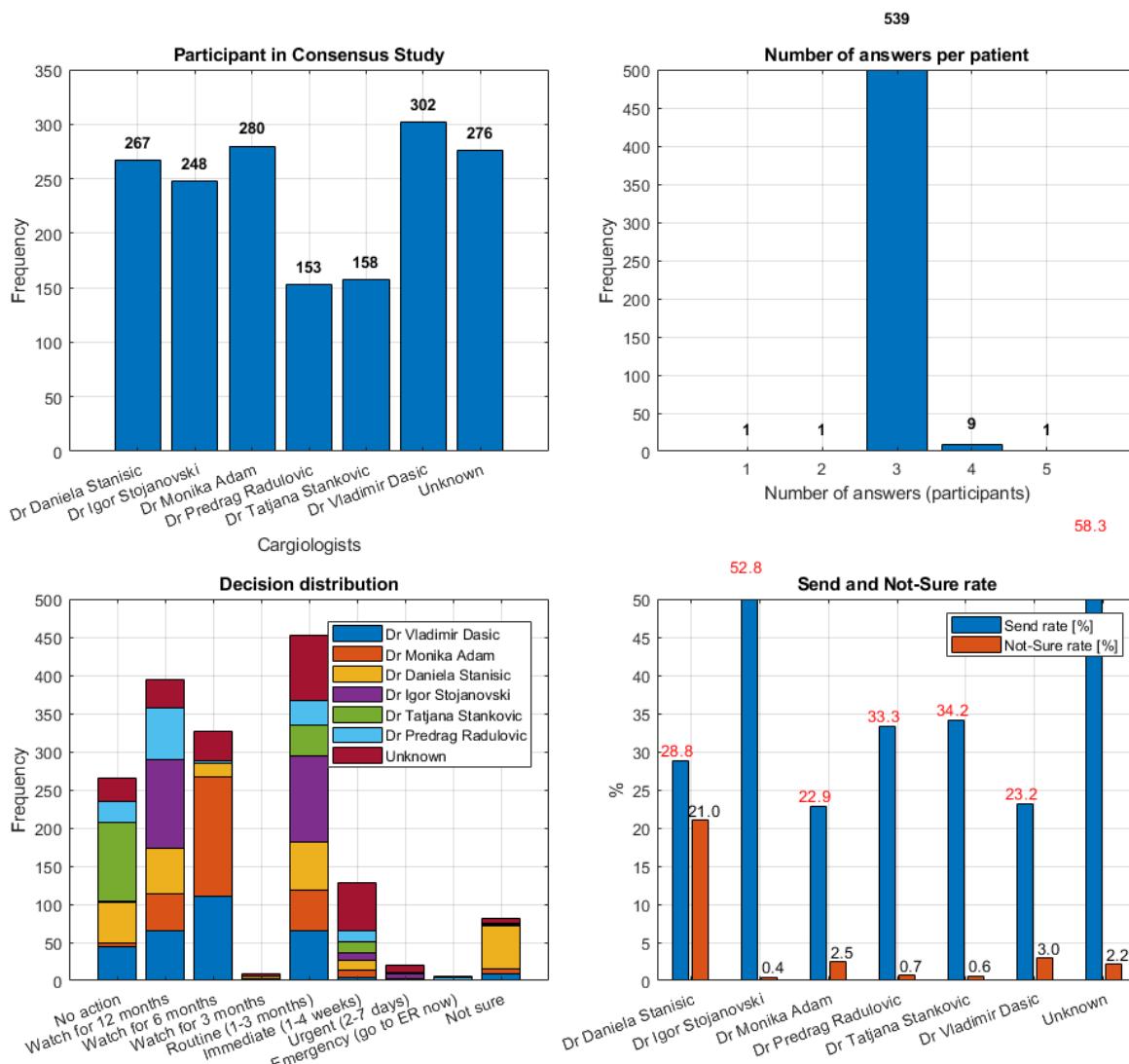
- HART – LV Diastolic Function shows average reproducibility - the classifier relies more on MCG

## 8.5 Consensus Based ground truth

### 8.5.1 Consensus Input Statistics

There are 1684 consensus records for establishing the Ground Truth (GT) of 550 patients. Most of the patients have three answers/participants, and 6 cardiologists are involved from at least two centers.

Figure 29 shows the distribution of participants, the number of answers, the distribution of referral decision answers, and the send-rate by cardiologists.



**Figure 29- General statistics of consensus (upper graphs), and referral decision statistics (lower graphs)**



## 8.5.2 Referral Decision

### 8.5.2.1 Aggregation method

The aggregation method selects the most probable decision answer based on the three cardiologist answers. The ground truth of decision should be detailed in a binary format too.

*Table 20- The consensus or agreement performance*

Performance	Description	Detailed examples	Binary examples
1/3	in case of three answers are different	Don't:Wa12, Don't:Wa3, Send:Rout	Don't, Send, Not Sure (or missed)
2/3	Two answers have agreement	Don't:Wa3, Send:Imme, Send:Imme	Don't, Send, Send
3/3	Three answers are in consensus	Don't:Wa12, Don't:Wa12, Don't:Wa12	Don't, Don't, Don't

To solve the aggregated decision a method is required, as the detailed decisions are mostly not in consensus<sup>27</sup>, only in case of binary decision is it possible to select the majority vote - see consensus performance of detailed and binary decision on Figure 30.

The aggregation method is following:

1. Replace the Not Sure to Routine – it mimics the real situation, as when GPs are not sure they are cautious and will tend to send. Also, because of uncertainty they send the patient with routine priority. (If we disable the not sure answers from the consensus instead of the proposed solution, the other two answers will decide only, would not show significant difference, the overall distribution of Send/Don't would be the same. However, the major vote definition in case of two answers is not possible – this is why minimum three answers are collected.)
2. Transfer the detailed decision to binary:
  - a. NoA, Wa12, Wa6, Wa3 -> Don't,
  - b. Rout, Imme, Urge, Emer -> Send
3. Select the major vote from three answers, which always exist. This will be the binary decision ground truth.
  - a. Don't + Don't + Don't -> Don't

---

<sup>27</sup> The weak consensus performance of detailed decision was expected, because of the 8 detailed priority resolution options. Since there are no strict medical rules which patient should wait to 6 months, which 3 months, etc. These decisions are based on individual medical experience.



- b. Don't + Don't + Send -> Don't
  - c. Don't + Send + Send -> Send
  - d. Send + Send + Send -> Send
4. Select the highest priority decision from the major group of binary decision – it assumes that a higher priority decision has more certain medical evidence compared to a lower priority decision. This will be the detailed ground truth.
- a. Don't:NoA + Don't:Wa12 + Don't:Wa3 -> max(NoA, Wa12, Wa3) = Wa3
  - b. Don't:NoA + Don't:Wa12 + Send:Imme -> max(NoA, Wa12) = Wa12
  - c. Don't:NoA + Send:Rout + Send:Urge -> max(Rout, Urge) = Urge
  - d. Send:Rout + Send:Rout + Send:Imme -> max(Rout, Rout, Imme) = Imme

### 8.5.2.2 Consensus Output Statistics

The result of aggregation can be seen in Figure 30.

The overall consensus performance is measured by the percentage of total agreements (3/3) cardiologists in the whole set of patient cases. The average consensus performance of binary referral decision is 87.5%, or cardiologists agreement in Send/Don't is 87.5% and for priority/watch is 60%.

The average consensus performance of detailed referral decision is 60.6%. (this accounts for variability is the different watch periods)

The overall send rate is 43,0%.

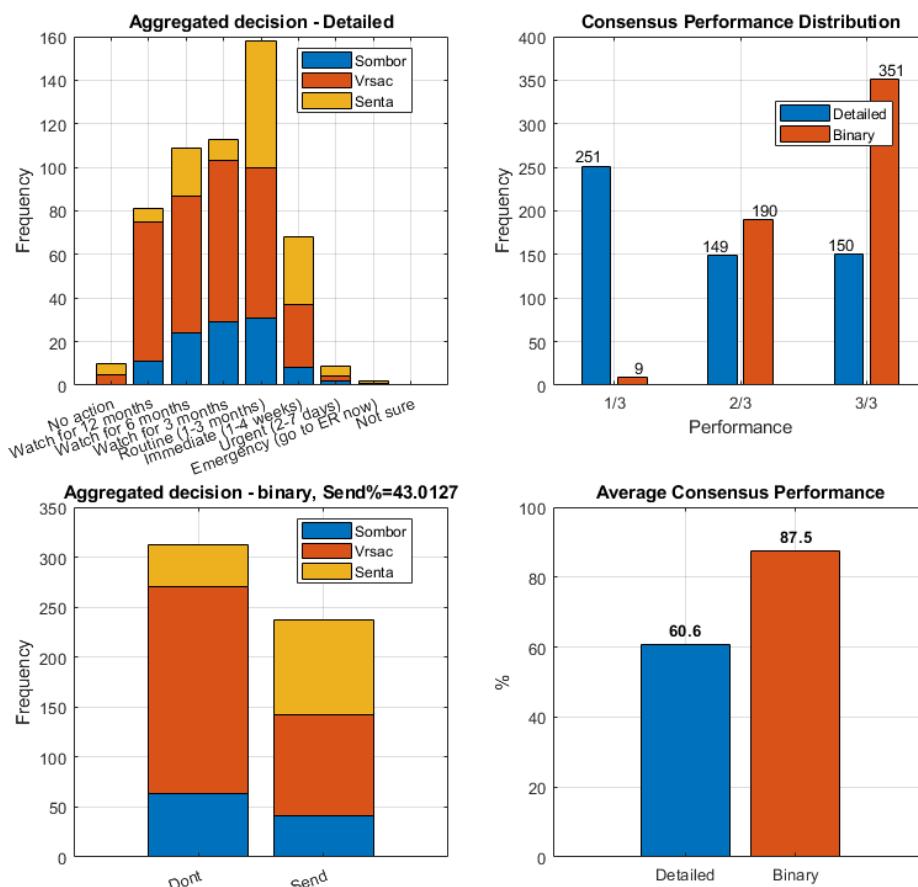


Figure 30- Referral decision aggregation and performance in consensus study



### 8.5.3 ECHO Findings

There are 20 ECHO findings included in the consensus forms that define and establish the ground truth diagnosis. See Table 21.

#### 8.5.3.1 Aggregation method

The aggregation of ECHO findings is based on averaging. The four kinds of class have each their own representative number, on which the rounded average can be calculated:

- Normal:  $x = 1.0$
- Mild:  $x = 2.0$
- Moderate:  $x = 3.0$
- Severe:  $x = 4.0$

The  $x$  denotes the input classes,  $y$  the output class, and  $g$  is the aggregation function.

$$y = g(x|x \in \mathbb{R}^3) = \left\lfloor \frac{1}{3} \sum_{i=1}^3 x_i \right\rfloor$$

The performance is calculated by followings:

$$p = \frac{1}{3} \sum_{i=1}^3 \begin{cases} 1 & x_i = y \\ 0 & x_i \neq y \end{cases}$$

Examples:

Normal + Normal + Normal	$y = (1+1+1) / 3 = 1$	Normal	p=3/3
Normal + Normal + Mild	$y = (1+1+2) / 3 = 1.33 \sim 1$	Normal	p=2/3
Normal + Normal + Moderate	$y = (1+1+3) / 3 = 1.66 \sim 2$	Moderate	p=0/3
Normal + Mild + Moderate	$y = (1+2+3) / 3 = 2$	Mild	p=1/3
Normal + Mild + Severe	$y = (1+2+4) / 3 = 2.33 \sim 2$	Mild	p=1/3
Normal + Moderate + Severe	$y = (1+3+4) / 3 = 2.66 \sim 3$	Moderate	p=1/3
Mild + Moderate + Severe	$y = (2+3+4) / 3 = 3$	Moderate	p=1/3
Moderate + Moderate + Severe	$y = (3+3+4) / 3 = 3.33 \sim 3$	Moderate	p=2/3
Moderate + Severe + Severe	$y = (3+4+4) / 3 = 3.66 \sim 4$	Severe	p=2/3



### 8.5.3.2 Consensus Output Statistic

Table 21 lists the 20 ECHO findings, its abbreviation, the number of classes for all the cardiologist input and for the aggregated output, and the average consensus performance.

**Table 21- ECHO findings Consensus Statistics**

ECHO Finding	Abbr.	All the input				Aggregated output				Consensus Perf.
		Normal	Mild	Mod.	Sev.	Normal	Mild	Mod.	Sev.	
<b>LV Systolic Function</b>	LVSD	1392	199	86	8	431	66	24	2	0.915
<b>LV Dilatation</b>	DCM	1566	65	43	11	462	31	23	7	0.943
<b>LV Hypertrophy</b>	LVH	1297	305	64	19	405	95	23	0	0.874
<b>LV Geometry</b>	LVG	1449	158	57	21	441	75	7	0	0.846
<b>LV Wall Motion Abnorm.</b>	WMA	1536	105	43	1	469	49	5	0	0.959
<b>Pericardial Effusion</b>	PEE	1669	14	2	0	518	5	0	0	0.989
<b>Left Atrial Enlargement</b>	LAE	1452	157	66	10	457	48	17	1	0.929
<b>Right Atrial Enlargement</b>	RAE	1606	57	20	2	499	18	5	1	0.971
<b>Right Ventricular Enlarg.</b>	RVE	1493	142	48	2	463	55	5	0	0.902
<b>Aortic Valve Stenosis</b>	AS	1588	62	22	13	493	20	7	3	0.962
<b>Aortic Valve Insuff.</b>	AR	1479	150	53	3	436	56	30	1	0.952
<b>Aortic Ascend. Dilation</b>	AOD	1666	13	6	0	517	6	0	0	0.985
<b>Mitral Valve Stenosis</b>	MS	1670	13	1	1	509	12	1	1	0.984
<b>Mitral Valve Insuff.</b>	MR	1174	388	109	14	363	126	30	4	0.941
<b>Mitral Valve Prolapse</b>	MVP	1683	2	0	0	523	0	0	0	0.999
<b>LV Diastolic Function</b>	LVDD	687	858	124	16	202	290	25	6	0.836
<b>Tricuspid Valve Stenosis</b>	TS	1673	11	1	0	521	2	0	0	0.992
<b>Tricuspid Valve Insuff.</b>	TR	1222	377	82	4	375	129	18	1	0.931
<b>Pulmonary Hypertension</b>	PH	1573	96	15	1	490	31	2	0	0.952
<b>Pulmonic Valve Insuff.</b>	PR	1583	102	0	0	488	35	0	0	0.989
<b>average</b>		1473	164	42.1	6.3	453	57.5	11.1	1.4	0.943

The average performance is plotted in right-top graph of Figure 31. The right bottom graph shows the histogram of all the performances. The last graph shows the three input classes by colors, where the vertical axis represents the individual patients.

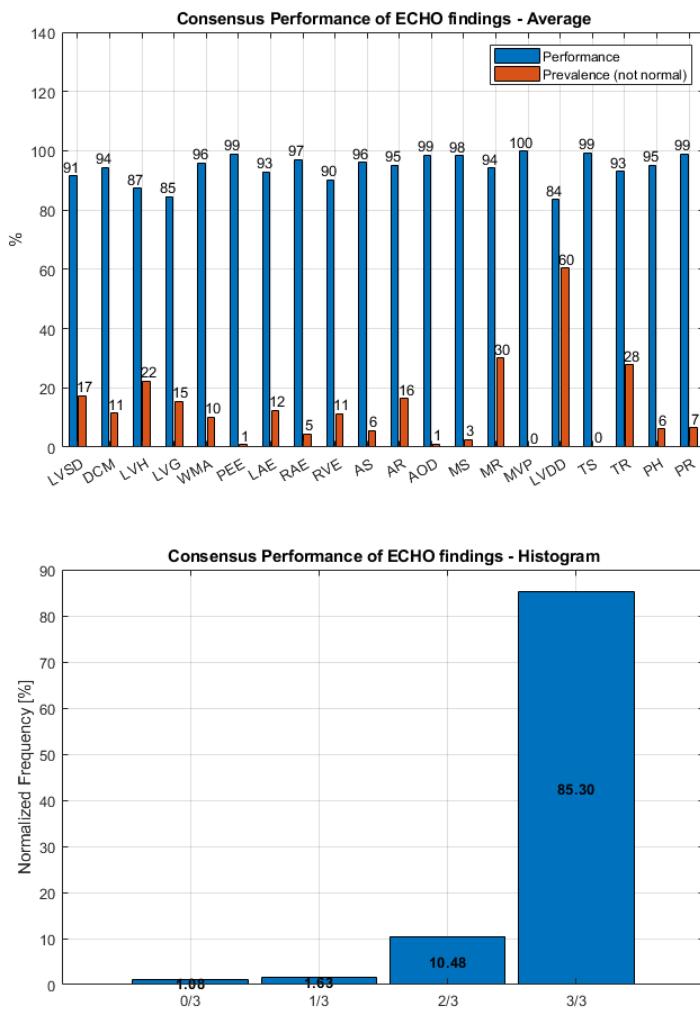


Figure 31- ECHO findings aggregation and consensus performance

#### 8.5.4 Conclusion

- Of the patients participating in the CUUS study, 43% are referred to Send to cardiologist care as the result of consensus-based ground truth. This is an expected number and consistent with the GP's, ORC' and RC's decision.
- The agreement or consensus performance of detailed decision is weak; however, it was expected, because of the 8 detailed priority resolution options. But the performance of binary (Send/Don't) decision is acceptable, 63.8% have 3/3 performance, 34.6% have 2/3 performance, and only 1.6% have 1/3 performance.
- The consensus performance of ECHO findings is acceptable, the perfect consensus is reached in 85.3% (performance = 3/3), the acceptable issues is reached 96% (performance  $\geq 2/3$ ). The performance per findings depends on the finding prevalence, but this is expected, because the biggest cardiologist uncertainty is between the severity of finding (mild or moderate or severe) and not between normal and abnormal.



- The prevalence of all the ECHO findings are expected, except the pulmonary regurgitation (PR) with the 7% prevalence. In the training and validation databases of CHART this was under 1%, and this was the reason to exclude this finding from 14 HART findings. Further investigation of this is required as it is likely that the reason for the low prevalence in the training and validation datasets was that the medical teams were not looking for it, so its GT could not be identified like in the CUUS.
- The most frequent abnormality is the “LV diastolic function = Impaired relaxation” with the 60% prevalence. Except PR, all the frequent findings (PREV>3%) are addressed by the 14 HART findings.

## 8.6 HART findings classification performance

### 8.6.1 Performance by ECHO findings

The 14 ECHO findings can be easily transformed into triple-class format making them compatible with the 14 triple HART findings, see Table 22. Thus, ECHO-findings used as reference ground truth in performance evaluation of HART findings.

*Table 22- ECHO-findings transforming to HART-findings (LVDD different, listed at the end)*

ECHO findings	1. class	2. class	3. class	4. class
LVH	Absent	Mild	Moderate	Severe
DCM	Absent	Mild	Moderate	Severe
LAE	Absent	Mild	Moderate	Severe
RAE	Absent	Mild	Moderate	Severe
RVE	Absent	Mild	Moderate	Severe
LVSD	Normal	Mild	Moderate	Severe
WMA	Absent	Mild Hypok.	Hypok./Akines.	Dyskin./Aneurism
AS	Absent	Mild	Moderate	Severe
AR	Absent	Mild	Moderate	Severe
MR	Absent	Mild	Moderate	Severe
MS	Absent	Mild	Moderate	Severe
TR	Absent	Mild	Moderate	Severe
PH	Absent	Mild	Moderate	Severe
HART findings	<b>1. Normal</b>	<b>2. Mild</b>	<b>3. Moderate/Severe</b>	
ECHO findings	1. class	2. class	3. class	4. class
LVDD	Normal	Impaired Relax.	Pseudonorm.	Restrictive Filling
HART findings	<b>1. Normal</b>	<b>2. Mild</b>		<b>3. Moderate/Sev.</b>

The performance and prevalence results listed in the following table.



Table 23- Performance of HART findings by ECHO findings

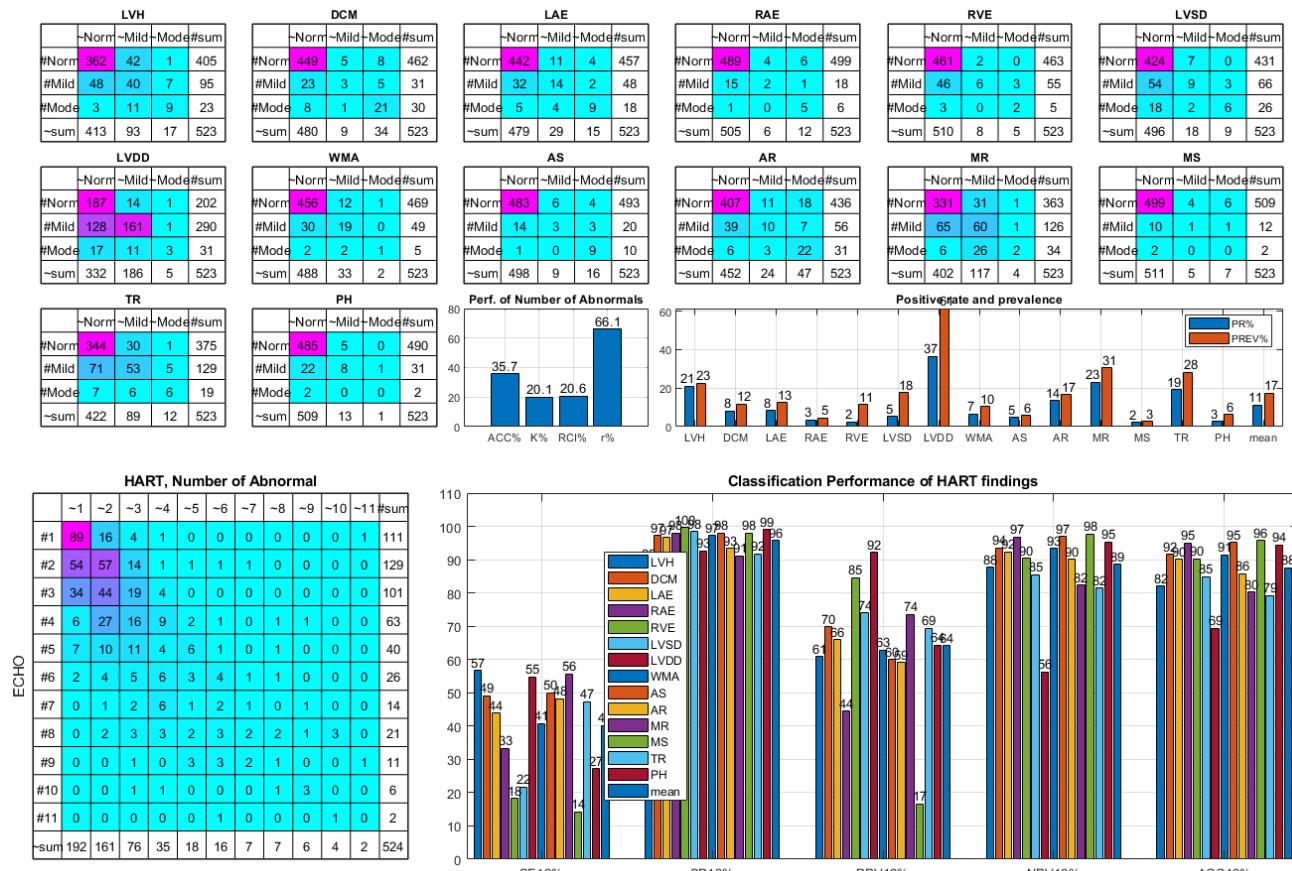
HART finding	SE12%	SP12%	PPV12%	NPV12%	ACC12%	K%	PR%	PREV%	LR+	LR-
<b>LVH</b>	56.8	89.4	60.9	87.7	82.0	39.6	21.0	22.6	5.3	0.48
<b>DCM</b>	49.2	97.2	69.8	93.5	91.6	48.2	8.2	11.7	17.5	0.52
<b>LAE</b>	43.9	96.7	65.9	92.3	90.1	42.7	8.4	12.6	13.4	0.58
<b>RAE</b>	33.3	98.0	44.4	96.8	95.0	33.9	3.4	4.6	16.6	0.68
<b>RVE</b>	18.3	99.6	84.6	90.4	90.3	23.5	2.5	11.5	42.4	0.82
<b>LVSD</b>	21.7	98.4	74.1	85.5	84.9	24.7	5.2	17.6	13.4	0.80
<b>LVDD</b>	54.8	92.6	92.2	56.3	69.4	41.0	36.5	61.4	7.4	0.49
<b>WMA</b>	40.7	97.2	62.9	93.4	91.4	42.9	6.7	10.3	14.7	0.61
<b>AS</b>	50.0	98.0	60.0	97.0	95.2	47.1	4.8	5.7	24.7	0.51
<b>AR</b>	48.3	93.4	59.2	90.0	85.9	40.4	13.6	16.6	7.3	0.55
<b>MR</b>	55.6	91.2	73.6	82.3	80.3	39.7	23.1	30.6	6.3	0.49
<b>MS</b>	14.3	98.0	16.7	97.7	95.8	9.9	2.3	2.7	7.3	0.87
<b>TR</b>	47.3	91.7	69.3	81.5	79.2	39.4	19.3	28.3	5.7	0.57
<b>PH</b>	27.3	99.0	64.3	95.3	94.5	33.8	2.7	6.3	26.7	0.73
<b>average</b>	<b>40.1</b>	<b>95.7</b>	<b>64.1</b>	<b>88.6</b>	<b>87.5</b>	<b>36.2</b>	<b>11.3</b>	<b>17.3</b>	<b>14.9</b>	<b>0.6</b>

The triple class form is simplified to binary and the performance is calculated for binary case at the performance metrics denoted with “12”: SE12%, SP12%, ACC12%, PPV12%, NPV12%. Practically the 2. Mild and 3. Abnormal are joined into a general “2. Abnormal” class, creating a simple binary classification and confusion matrix, from which the performance is then calculated.

Each individual HART finding is evaluated, except for one overall statistics, the number of abnormal findings is analyzed as well. This represents the “how many abnormalities exist within the 14 HART findings”.

Figure 32 illustrates the:

- triple class confusion matrixes where the vertical reference # denote the ECHO and horizontal ~ denotes the HART findings;
- binary prevalence;
- the performance of HART findings: binary sensitivity (SE12%), binary specificity (SP12%), binary PPV, binary NPV, binary accuracy (ACC12%), Kappa statistics (K%)
- and the confusion matrix and performance related to “Number of Abnormal” – this is important to evaluate HART findings total impact compared to ECHO-findings total impact.



**Figure 32- HART findings performance by ECHO findings:** 14 top graphs show the confusion between triple HART and triple ECHO-findings; 15<sup>th</sup> graph show the performance of “Number of Abnormals” table; middle right graph compares the prevalence of HART and ECHO findings; left bottom graph show the confusion of “Number of Abnormals” between HART and ECHO; right bottom graph shows the performance metrics for 14 HART findings

### 8.6.2 Performance by Referral decision

In this section we analyze the HART findings prediction capability as to whether the patient should be referred to the cardiologist, the Send/Don’t decision, as established by consensus, see section 8.5.2.

The 14 distribution graphs in Figure 33 related to 14 HART findings illustrate that most of mild and moderate classification of HART belong in the “Send” decision. This means the positive prediction capability of HART findings is really high.

Therefore, the positive predictive values (PPV%) of HART findings are plotted in the left bottom graph in Figure 33. The yellow color represents the predictive value of Mild or Moderate class by HART findings, white pink color represents the predictive value of Moderate class.

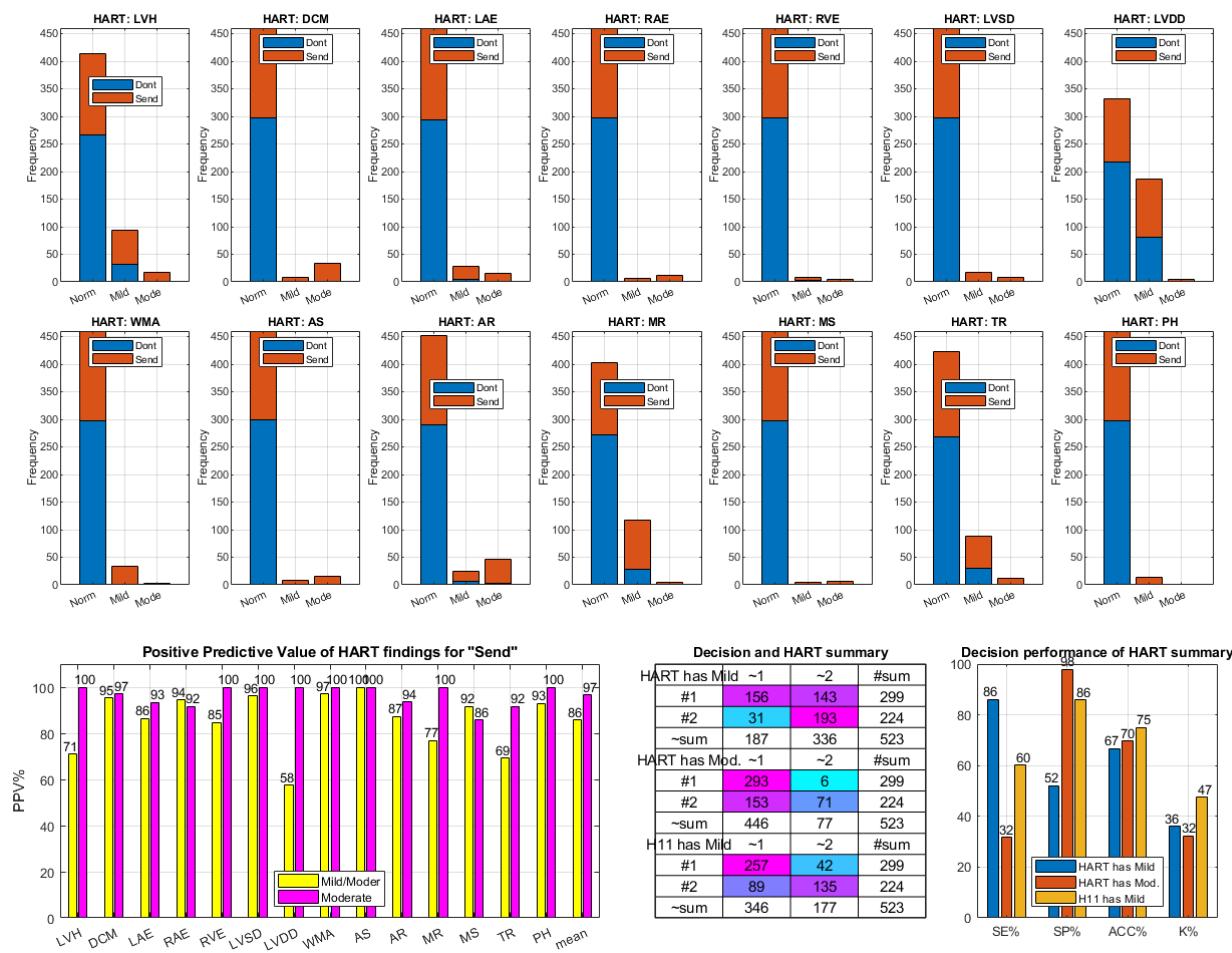


Figure 33- HART findings prediction capability to referral decision

Three kinds of HART findings aggregation interpretation are evaluated separately, and its confusion matrix and performance are plotted in the right bottom graphs in Figure 33:

- "HART has Mild" – in this interpretation or aggregation method, HART suggest "Send" if any of HART shows Mild or Moderate class. This reached high sensitivity of decision (SE=86%) but low specificity (SP=52%). Consequently, this interpretation of HART findings would be super sensitive, but creates too many false positive.
- "HART has Mod." – in this interpretation or aggregation method, HART suggest "Send" if any of HART shows Moderate independent from Mild classes. This reached low sensitivity (SE=32%) but very high specificity (SP=98%). Consequently, if HART shows a Moderate/Severe abnormality then the reader will be right 98% of the time with a Send decision.
- "H11 has Mild" – From the PPV analysis of 14 HART findings, three findings show lower values (LVH, LVDD, and TR). In this interpretation we select 11 HART findings from 14 (exclude 3 findings), and do the same aggregation under point a) – if any of 11 HART findings shows Mild or Moderate. This reach a balanced decision with sensitivity (SE=60%) and specificity (SP=86%) compared to reference decision. Complement this with the ECG, PCG and MCG findings,



sensitivity would likely be increased and creating the best interpretation of HART findings with respect to the referral decision – see more details in conclusion.

### 8.6.3 Conclusion

The overall performance of HART findings is acceptable and close to the expected values, which is derived from the external and independent validation performances from CPA-HART findings. The average accuracy ACC=87%, and average Kappa K=36%

The classification performance of three HART findings are weaker than the expected level, having similar strict threshold type weakness. As reference for comparison purpose, the average SE=40% with SP=96%.

- RVE – low sensitivity (SE12=18.3%) with strong specificity (SP12=99.6%)
- LVSD – low sensitivity (SE12=21.7%) with strong specificity (SP12=98.4%)

In case of these low performance findings the positive classification means some abnormality even if it is not the correct one. For example, when HART: MS show Moderate, but really there is no mitral stenosis, then patient is likely to have aorta stenosis, and or mitral sclerosis. This can be observed by analyzing the HART findings performance together, described in section 8.6.2. This is consistent with the nature of heart disease where there is never just one disease.

We can conclude that HART findings have a strong positive predictive value in referral decision of patients. Using the 11 most predictive HART findings<sup>28</sup> the HART findings-based decision shows ACC=75% accuracy, SE=60% sensitivity and SP=86% specificity – which is similar performance reached by GP, ORC and RC on CHART report.

## 8.7 Usability evaluation

A Human Factors Validation Study was performed during the Clinical Study, with a total of 16 participants- Medical Assistants (MA) and General Practitioners (GP) which has the contact with the Device (whole system), to demonstrate the usability of the user interface.

The study enrolled 2 user groups:

- Group 1- User divided – MA and GP
- Group 2- Users age 20-49 and 50+

Each group included participants with and without Computer skills, as well as participants who use some similar ECG instrument before.

Testing identified critical tasks as those tasks where the user does not understand working with the device properly, or limitations of the device. Each task was assessed for completion and success criteria were clearly defined. Users had the accompanying

---

<sup>28</sup> DCM, LAE, RAE, RVE, LVSD, WMA, AS, AR, MR, MS, PH



documentation available for use. Testing also collected subjective feedback in a written open response questionnaire (according to ISO 62366-1:2015) and post-test interview.

### 8.7.1 Group descriptive statistic

#### 8.7.1.1 Group 1- MA/GP

Table 24- MA/GP descriptive statistic

		Frequency	Percent
Valid	GP	7	43.8
	MA	9	56.3
	Total	16	100.0

From 16 enrolled medical professionals, 9 are working as Medical Assistant, and 7 as General Practitioners.



Graph 34- MA/GP distribution

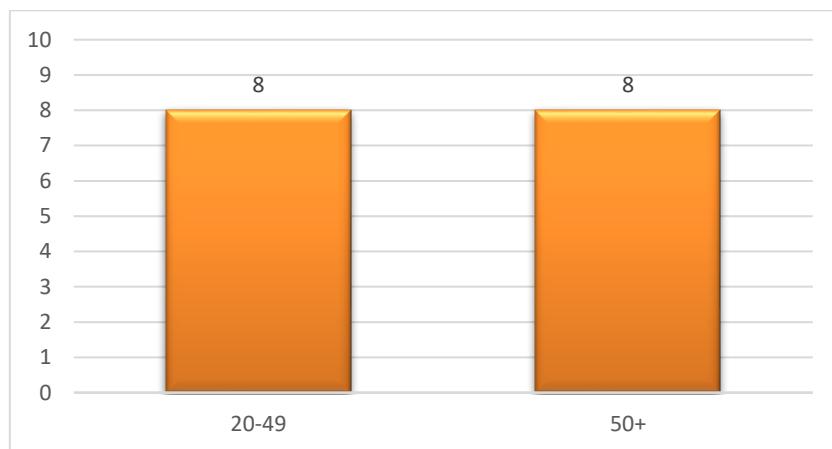


### 8.7.1.2 Group 2- Age distribution

Table 25- Age distribution descriptive statistic

		Frequency	Percent
Valid	20-49	8	50.0
	50+	8	50.0
	Total	16	100.0

There were 16 medical professionals working with the device. 8 of the participants have 20-49 years, and 8 has more than 50 years.



Graph 35- Age distribution descriptive statistic

### 8.7.1.3 Usability scale distribution

Table 26- Usability scale distribution

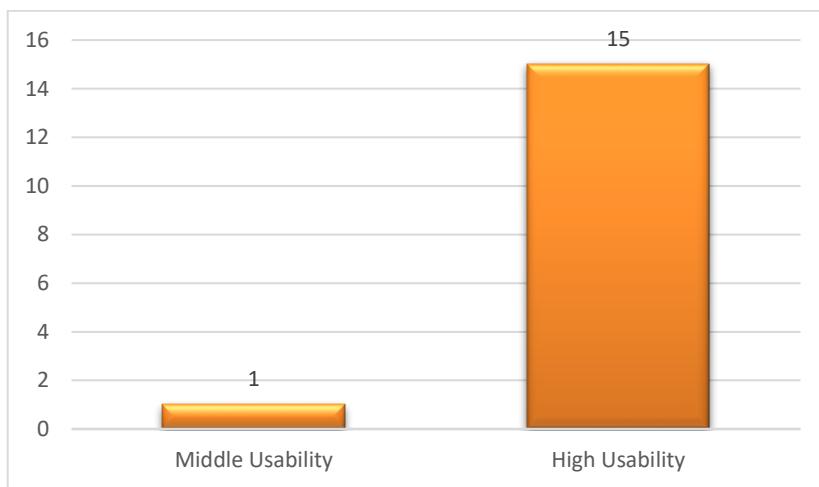
		Frequenc y	Percen t
Valid	Low Usability	0	0
	Middle Usability	1	6.3
	High Usability	15	93.8
	Total	16	100.0

In Table 26 we show the answers of the users regarding to the Usability of the device. Usability was measured with the test which contains the 17 questions related to the usefulness of the device. Some of them are:



- It is simple to use CHART system/Cardio-TriTest v6.5 medical device
- I can efficiently complete medical test using Cardio-TriTest v6.5
- I feel comfortable using Cardio-TriTest v6.5
- It was easy to learn how to use CHART system
- I believe I became more productive using CHART system
- Whenever I make a mistake using the system, the system recovers easy and quick
- The additional provided information are clear.
- The interface of the system is pleasant.
- It is easy to find the information I needed.

Result show us that non user grade the device or the system as useless (low usability). 93.8% of the users grade the device and its interface as high useful, in its intended use.



**Graph 36- Usability distribution**

#### **8.7.1.4 Usability Effectiveness- Overall**

$$\text{UE} = \frac{\text{Number of MT completed successfully}}{\text{Total number of MT undertaken}} \times 100$$

$$\text{UE} = \frac{409}{421} \times 100 = 97.23 \%^{29}$$

Looking from the significant point of view, using the given equation, we calculate that 97.23% of the measured usability was effective. This state is confirming our expectations, on 0.027 level of significance.

---

<sup>29</sup> level of significance – 0.027



### 8.7.1.5 Group 1 Calculation

**Table 27- Descriptive statistic**

			Health specialty		Total
			GP	MA	
Low usability	Count	0	1	1	
	% of Total	0.0%	6.3%	6.3%	
Middle usability	Count	3	2	5	
	% of Total	18.8%	12.5%	31.3%	
High usability	Count	4	6	10	
	% of Total	25.0%	37.5%	62.5%	
Total	Count	7	9	16	
	% of Total	43.8%	56.3%	100.0%	

Comparing medical professionals' groups with the Usability scale to see if there is a statistical significance between GP's and MA in answers on the Usability scale, we can confirm that the significant difference doesn't exist (level of significance 0.504). This result tells us that the device/system was easy to use to GP's as well as to the MA, regardless of the level of education or Computer skill level.

**Table 28- Chi-Square Tests**

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	1.371	2	.504
Likelihood Ratio	1.740	2	.419
Linear-by-Linear Association	.003	1	.960
N of Valid Cases	16		



### 8.7.1.6 Group 2 Calculation

**Table 29- Descriptive statistic**

			Usability			<b>Total</b>
			Low usability	Middle usability	High usability	
20-49	Count	0	1	7	8	
	% of Total	0.0%	6.3%	43.8%	50.0%	
50+	Count	1	4	3	8	
	% of Total	6.3%	25.0%	18.8%	50.0%	
<b>Total</b>	Count	1	5	10	16	
	% of Total	6.3%	31.3%	62.5%	100.0%	

**Table 30- Analysis of Variance - ANOVA**

	<b>Sum of Squares</b>	<b>df</b>	<b>Mean Square</b>	<b>F</b>	<b>Sig.</b>
<b>Between Groups</b>	1.563	1	1.563	5.000	.042
<b>Within Groups</b>	4.375	14	.312		
<b>Total</b>	5.937	15			

As for the result from the Table 29 and 30, where we wanted to test if there is a significant difference between Age categories and Usability scale, according to the level of significance (Sig.) of 0.042, we can conclude that the significantly differences between the age categories regardless the usability scale exists. From the Table 29 we can see how are those differences distributed.

There are some positive and negative comments we captured during the usability testing which we will gladly accept, and work on them in the future to improve our system/device and to reach much higher level of usability.

### 8.7.1.7 Positive comments/Negative comments

**Table 31- Positive comments**

<b>Positive comments</b>	<b>Frequency</b>
Easy to use, the 3-test recording itself without Risq questionnaire is short	6
Understandable	7
Working with the device is understandable and not complicated	3



Forms and reports are user friendly and very helpful	6
System gives a lot of information about the conditions of the patient health, with reference to the correspondent aspect	8
More information	8
More information and therefore the diagnosis are more accurate	4
Quick and easy availability of the information, not necessary refer every patient to cardiologist because of the better picture of the patient health status	7

**Table 32- Negative comments**

Negative comments	Frequency
If the sensors are not connected properly, the test cannot be done, if cables are disconnected during the testing, software blocks, and the test cannot be continued, software crashes	4
Lack of the explanation of the CHART report terms, it was very difficult to understand the short names of the findings	5
Internet connection problems	6
On CHART report the heart rate is missing, and that can be very helpful for the diagnosis	3
Too much time, it takes a lot of time from established routine time scheduled by SOC	4
Long Risq questionnaire, the examinations last more than expected	4
Cable connection problems	6

Based on the comments got from the users in CUUS, we would test the usability of the device and the whole system again to establish the risks, which we already consider as present and prepare the future requirements which will help us improve our products on a higher level.

## 8.8 Corrections and diagnosis and decisions

During the study there were some unclear understandings of the protocol and this resulted in a biased answer and statistics. During analysis in detail of the data, the biases were found, the reason investigated by interview with the doctor and corrections made with the approval of affected user.

### 8.8.1 Dr. Brankica referred everybody

#### 8.8.1.1 Bias observed

Dr Brankica is a GP, when working on the ECG protocol, she referred almost everybody in the first 150 patients.



Her decision answers show large bias toward “Send” compared to other GPs in the referral decision on ECG report. Related, this skewed ECG as more sensitive compared to CHART, which was not expected.

#### **8.8.1.2 Reason investigation**

After an interview with GP, the following reason was found:

She thought all the patients were to be Send to cardiology to verify diagnosis for Ground Truth, irrespective of her decision. She therefore Send most all patients with no medical justification as Normal patients for the study.

#### **8.8.1.3 Corrective action**

All the patients affected by her decision was re-verified and the “Send/Don’t Send” decisions corrected based on her actual diagnosis. She verified the results.

Patient re-evaluated as “Don’t send” having normal diagnosis and following IDC10 codes:

- I10 - Essential (primary) hypertension
- I11 - Hypertensive heart disease
- I15 - Secondary hypertension
- I95 - Hypotension
- I99 - Other disorder of circulatory system – which code was used for the normal patients

Altogether 107 patient decisions were corrected, from which 100 patients were changed from false “Send” to “Don’t” decision category, with her approval.

#### **8.8.1.4 Preventive action**

She was re-educated on the correct interpretation of the protocol, and her remaining patients were correctly processed. When she switched to the CHART protocol, her bias disappeared.

### **8.8.2 Dr. Daniela- switched from CHART to ECG protocol in the GP evaluation**

In the initial phase of the CUUS, Dr. Daniela was first put on the CHART protocol. Midway through the CUUS, she was then switched to the ECG protocol.

#### **8.8.2.1 Bias observed**

In the referral diagnosis (ICD10 code) we found CHART+ based diagnosis evaluated on ECG-only report from Dr. Daniela.

RisQ and ECG report does not provide any medical justification to diagnose these the following heart disease that CHART+ is engineered to do:



- I05.0 - Rheumatic mitral stenosis
- I05.2 - Rheumatic mitral stenosis with insufficiency
- I06.8 - Other rheumatic aortic valve diseases
- I07.0 - Rheumatic tricuspid stenosis
- I07.1 - Rheumatic tricuspid insufficiency
- I08.9 - Rheumatic multiple valve disease, unspecified
- I34.0 - Nonrheumatic mitral (valve) insufficiency
- I35.0 - Nonrheumatic aortic (valve) stenosis
- I35.1 - Nonrheumatic aortic (valve) insufficiency
- I35.2 - Nonrheumatic aortic (valve) stenosis with insufficiency
- I35.8 - Other nonrheumatic aortic valve disorders
- I36.1 - Nonrheumatic tricuspid (valve) insufficiency
- I42.0 - Dilated cardiomyopathy

#### **8.8.2.2 Reason investigation**

In the initial phase of study, the online processing system to capture clinical study notes, the online documentation system did not hide the OTHER report, if the ECG protocol the CHART report was also present in the patient file. If the CHART protocol, then the ECG report was also present in the patient file. It was expected that depending on the protocol they would open that report. However, the GP was able to open either or both reports, irrespective of the protocol. Although their instructions were to only use the report in the protocol they were streamed in, they were not prevented from access to the other report.

After an interview with GP, the following was found:

Dr. Daniela had initially been on the CHART protocol, she had become reliant on CHART. So, when switched to the ECG protocol, whenever she felt that the ECG report did not adequately explain what she suspected her patient was suffering from, she simply reached into the file and read the CHART report. Dr Daniela was first attached to the CHART stream, and had completed some 150 patients using the CHART report only. She was well aware of the additional information the CHART report provided and when the ECG report did not provide her the diagnostic assistance, she had become accustomed to, she simply reached into the patient file and opened the CHART report. She had simply become accustomed to using the CHART report and found it not only very useful and necessary to answer various diseases such as Valve disease, PCG and MCG related diagnostic questions, but it also provided her with more confidence in her overall decisions – so she used it.

#### **8.8.2.3 Corrective action**

Altogether 17 patients were found to have CHART+ based diagnosis recorded by her on ECG forms, where the CHART+ based diagnoses were deleted.



From these 17 patients, 8 decisions were corrected from “Send” to “Don’t” decision category, where there was no ECG diagnosable abnormality only CHART+.

#### **8.8.2.4 Preventive action**

The qmsWrapper software program was corrected to present both reports being available in any one stream. In this case, the CHART report was no longer available in patients file if in the ECG only diagnosis stream.

### **8.8.3 Dr. Tatjana, Cardiologist- too many “No-Action” decisions in consensus**

#### **8.8.3.1 Bias observed**

The distribution analysis of decision shows a bias of Dr. Tatjana’s decision in the consensus forms. She has significantly higher No-Action decision and lower Watch decisions.

#### **8.8.3.2 Reason investigation**

After an interview with the Cardiologist, the following reason was found:

She considered the No-action and Watch options under Don’t send decision as relating to the cardiologist No-action and Watch decision, not what a GP should have done. In her opinion No action means nothing to do at the cardiology level, because the patient should be treated at the GP level.

It was a simple misunderstanding, because in the case of Don’t Send, the No-action really means GP Send the patient home without any treatment or control. She was to determine whether the referral decision and its priority should be taken as to what a GP should do, and not what a cardiologist should do.

#### **8.8.3.3 Corrective action**

All the original decisions of “Don’t: No-Action” by Dr. Tatjana were disabled where the diagnosis had some abnormality, and those tests were re-evaluated by her. The ground truth of the decision was established after based on re-evaluated answers for these patients.

Altogether 103 patients were classified as No-Action by her.

From which 45 patients have abnormally diagnosed by ECG or ECHO, and Tatjana’s decision were re-evaluated for these patients.

#### **8.8.3.4 Preventive action**

She was counselled. She has re-evaluated the decision for the concerned patient.

### **8.8.4 Hypothesis Acceptance/Rejection**



#### 8.8.4.1 Null-Hypothesis

Since the results evaluated in previous chapters showed that the CHART system is more effective than ECG-only analysis in various points, we can reject established Null-Hypothesis.

#### 8.8.4.2 Primary Hypothesis

Hypothesis point	Result
<p><b>“CHART analysis provides better diagnostic decision support in clinical practice compared to ECG-only analysis leading to better outcomes.”</b></p> <p><b>We are going to prove this statement by</b></p> <p>1) measuring the reduction in FP and FN rates, and</p>	<p>Hypothesis is accepted!</p> <p>False negative and False Positive of GP's referral decision is significantly reduced by CHART compared to ECG, according to a 95% CI.</p> <p>See details in section 8.3.1.</p> <p>This is reinforced by the ORC and RC statements about the ECG over CHART usability. See details in section 8.4.2</p> <p><b>Error! Reference source not found.</b></p>
<p>2) reduction in “no-evidence” and “Not Sure” answers,</p> <p><b>between ECG-only and CHART analysis, which will be verified by consensus decision ground truth, calculated according to a 95% confidence interval.</b></p>	<p>Hypothesis is accepted!</p> <p>“Not sure” diagnosis is significantly reduced by CHART compared to ECG, according to a 95% CI.</p> <p>See details in 8.4.1.</p> <p>This is reinforced by the ORC and RC statements about the ECG over CHART usability. See details in 8.4.2.</p>
<p><b>“CTT bio-signals are valid and consistent in accordance with the accepted standards in clinical practice.” This will be proved by:</b></p> <p><b>1) ECG - measurements to the IEC60601-2-25 Standard for Diagnostic ECG.</b></p> <p><b>2) PCG/MCG - measurements to the Signal Standard for PCG</b></p>	<p>Hypothesis is accepted!</p> <p>The bio-signals provided by CTTv6.5 device is verified and validated by the signal performance standards.</p> <p>The repeatability results (10-50%) are generally showing smaller deviation compared to reproducibility results (20-70%). This was expected, because the source of the signal (patient heart) has</p>



<p><b>and MCG bio-signal measurement Standard.</b></p> <p><b>3) Reproducibility and Repeatability testing of ECG, PCG and MCG between:</b></p> <ul style="list-style-type: none"><li><b>a. different CHART tests between Clinical Practice and Cardiology (CHART 1 and CHART 2)</b></li><li><b>b. different CHART test between Different Centers.</b></li></ul>	<p>changed more within few days than few minutes.</p> <p>The ECG consistency trend is two type:</p> <ul style="list-style-type: none"><li>• The standard average-beat based measurements show really strong consistency values (~10% repeatability, and ~30% reproducibility)</li><li>• The heart rate related features show relatively high consistency value (~40% repeatability, and ~50% reproducibility).</li></ul> <p>This is because the rhythm is frequently changed, sometimes after few second, but the ECG morphology is more consistency, since it depends on the heart myocardial conditions, which can only change more slowly.</p> <p>The PCG and MCG signals inherit these both type of ECG trend, however the most of features are similar to the ECG heart rate related features.</p> <p>The channel-based comparison shows higher deviation for ECG, than PMCG. This is because ECG leads has rotated shape (captured from different phases), sometimes having opposite sign, while PMCG signals are more similar captured on the four cardiac auscultation points (Aortic, Pulmonic, Tricuspid and Mitral). The heart sound has smaller-bigger amplitude on these auscultation points, but generally the waves radiating with same phases.</p> <p>Understanding these conditions all the results are expected and acceptable, and provides justification about PCG and MCG bio-signals has similar consistency as ECG signals.</p>
---	--



	<p>All the three type of bio-signals show similarly high consistency based on the distribution analysis between centers.</p> <p>The p-value &gt; 0.001 shows for most of features.</p> <p>The exception is only the ECG heart rate related features. This is because patients in Vrsac has less frequently arrhythmias compared to patients in other two centers. But this bias is correctly represented by the bio-signals.</p> <p>The files sent were exactly the same files received. There were NO alterations, tampering or other changes, resulting from the interoperability of the systems at start point to endpoint.</p> <p>The CHART system is deemed interoperability safe.</p>
--	---

#### 8.8.4.3 Secondary Hypothesis

Hypothesis point	Result
<p><b>1) CHART analysis provides better prioritization and referral medical justification than ECG-only analysis.</b></p> <p><b>This will be accomplished by comparing the prioritization results and the diagnostic findings between the Clinician and the cardiologists, comparing CHART with ECG-only analysis.</b></p> <p><b>The significant increase of effectiveness measuring by answers of the referral cardiologists about prioritization and medical justification between CHART and ECG-only, with 95% confidence interval.</b></p>	<p>Hypothesis is accepted!</p> <p>The performance of GP's detailed decision or prioritization significantly increased compared to ECG, when the reference detailed decision is from ORC.</p> <p>In case of RC as reference decision the performance increase is borderline significant according to a 95% CI.</p> <p>See details in 8.3.2.</p>



<p><b>2) CHART can provide an effective starting point for Echo examinations.</b></p> <p><b>This will be accomplished by Comparing the diagnoses of CHART#2 and ECHO to show that cardiac status is correctly indicated by CHART analysis to show adequate starting point for ECHO examination. This will be confirmed by summing the cardiologists' statements about the effectiveness of CHART analysis to this end.</b></p> <p><b>This effectiveness will be measured by summing the cardiologists' statements about the effectiveness of CHART analysis to this end.</b></p>	<p>Hypothesis is accepted!</p> <p>Based on RC statements, CHART can provide an effective starting point for ECHO examination in 98.7%. This means from 97.3% to 99.5% according to a 95% CI. See details in 8.4.5.</p> <p>This statement is reinforced by the HART findings classification performance measured with the consensus-based ECHO findings as reference. See details in 8.6.</p>
<p><b>3) CHART medical tests are reproducible and repeatable.</b></p> <p><b>This will be accomplished by comparing the results between CHART#1 and CHART#2, according to 95% confidence interval. This will be confirmed by summing the cardiologists' statements related to their comparison between the two tests.</b></p>	<p>Hypothesis is accepted!</p> <p>Based on RC statements, CHART is reproducible and repeatable at 90-100% level in 80.9%. This means from 72.9% to 89.2% according to a 95% CI.</p> <p>See details in 8.4.4.</p> <p>This meet with the statistics on findings. The ECG and HART findings reproducibility performances are accepted, HART has ACC=89%, K=40%, r=44%, ECG has ACC=92%, K=50%, r=57%.</p> <p>See details in 8.4.6.</p>
<p><b>4) MA have effective clinical understanding of CHART usability</b></p>	<p>Hypothesis is accepted!</p> <p>Based on the Usability tests and interviews report results, CHART system is Easy to use and understandable to end users (MA) on a level of 97.23%-significance 0.042.</p> <p>See details in 8.7.</p>



## 8.9 Requirement on safety

This CER had determined that the benefit/risk ratio of the device is most favourable for use in Primary Care. There were no adverse events during the Clinical Study, which was consistent with expectations and with previous Clinical Studies.

The device should be allowed to be marketed.

No additional risks were identified in the Clinical Data appraisal that are not in the risk analysis. The risk analysis was accordingly not updated.

No change was expected in the risk analysis, and the clinical study confirmed these expectations. No is there a different rate of risk compared to known predicate devices.

Based on the clinical data, no additional warnings should be added to the instructions for use. The IFU correctly identifies the intended purpose and contains information to reduce the risk of user error; residual risks; and appropriate warnings, precautions and contraindications.

A clinical investigation, of which this report is the result, was conducted in Europe. No further additional clinal investigations are required prior to CE Mark certification of the device. Post-market clinical follow-up in Europe should be monitored IAW ISO 13485:2016.

The Summary of the Benefits/risk include:

1. No additional risk compared to predicate devices. Device is very safe to use.
2. Increased effectiveness for a wider range of disease conditions.
3. Easy to use, little or no additional operator training required for existing operators of predicate devices.
4. Fits into current workflows for normal Standards of Care (SoC).
5. Suitable for use in Primary Care, and in most all patient care situations where disease diagnosis, but not monitoring, is required.
6. Is an effective assistant to Primary Care physicians in better understanding their patient's cardiac status.
7. Reduction of FN by %.
8. Reduction of FP by %
9. Widespread benefit to all patients attending Primary Care, or patient care clinics.
10. Helps Cardiologists with Collaborative triage of patient appointment priorities.
11. Helps Cardiologists identify start point for Echo examination, saving time, and costs.
12. Better able to detect and confirm onset of heart disease earlier when treatment options are more effective and cost-efficient.



## 8.10 Acceptable benefit/risk

No side-effects or other risks were identified. The nature and duration of the benefit to patients is earlier detection of a wider range of heart disease onset in Primary Care, with life-long benefits.

Diagnostic uncertainty levels are comparable to predicate devices and pose no additional risks. They are compatible with the highest level of protection of health and safety and offer corresponding justification for its adoption into widespread use.

### Annex G- Benefit Risk Analysis CUUS.

The benefits can be fully assessed based on the data available from the Clinical Study. No limitations, gaps or uncertainties were observed. Nothing was assumed. The available data allows for the adequate assessment of performance, gaps and possible but unseen, uncertainties. The evidence from the clinical Study was sufficient for every intended performance.

No side effects were detected, and none are expected.

## 8.11 Supporting Documents

- **Annex A-** Technical File Cardio-TriTest v6.5;
- **Annex B-** Technical File CHART Processing Algorithm;
- **Annex C-** Intended Use and Indication for Use;
- **Annex D-** Risk Assessment for CUUS Clinical Study;
- **Annex E-** Steps of CUUS;
- **Annex F-** Statistical plan for CUUS;
- **Annex G-** Benefit Risk Analysis CUUS;
- **Annex H-** Clinical Investigation plan- CIP;
- **Annex I-** Regulatory approval- ALIMS;
- **Annex J-** PMS/PMCF;
- **Annex K-** Qualification of the responsible evaluators.



## 9 Conclusions

More than 500 patients' clinical results are evaluated to measure the diagnostic and decision support capability of CHART report compared to ECG report.

Patient population is representative of the target population according to intended use of CHART. Some 43% of patients are classified as Sent to referral cardiology based on the consensus-based ground truth.

CHART resulted in a decrease of 16% in false negative and a decrease of 5% in false positives in the patient referral decision by GP compared to ECG only based decision.

In case of CHART report, the decision consistency between GP, ORC and RC is higher compared to ECG based decision. The send/don't decision agreement between GP and ORC increased from PPA=52% to 67%, NPA=77% to 81%, and between GP and RC from PPA=56% to 68%, NPA=78% to 80% by CHART.

GP, ORC and RC were significantly more sensitive in their referral decision based on CHART report compared to ECG. This is confirmed by the positive diagnosis rate, which was increased from 8.5% to 23-27% by CHART. This means the number of recognized abnormalities increased by 250% compared to ECG. The uncertainty rate of diagnosis is significantly decreased on average from 24% to 2% in accordance with previous statement.

It has been shown that CHART really has an extended range of detectable abnormalities suggestive of heart diseases compared to ECG.

This benefit of CHART is reinforced by the questionnaire results both from GP, ORC and RC groups. In 64% of cases, CHART reduced diagnostic uncertainty and provided a better understanding of the patient's overall cardiac status. RC and ORC changed their Send/Don't referral decision in 19%-22% of the patients compared to primary ECG based decision.

The reproducibility of findings in ECG and CHART reports show 90% on average, and this is acceptable taking into account the 10-day average time delay between the two CTT recordings. The CHART ECG report has better reproducibility ( $K=50\%$ ) compared to the predicate ECG ( $K=45\%$ ), but HART findings has lower ( $K=40\%$ ) on average.

The consensus performance of Send/Don't binary decision and ECHO findings are meet with the expectation, both showing a 98% with a minimum 2/3 agreement.

### 9.1 Benefit-Risk Determination

The CHART system is intended to be used to assist General Practitioners (GP's) to better understand and determine the Cardiac State of their patients.



CHART system is intended to provide objective assessment and diagnostic interpretation support to clinicians in patient care situations to aid in the evaluation of cardiac status and prognosis.

Because the CHART system, or its Cardio-TriTest v6.5 device is very similar to ECG device, there are minimal safety concerns. The probable risks associated with using the device are nearly all related to false results in CPA detection or human use errors. False negative results may falsely reassure the GP and cause delay or inappropriate changes in medical evaluation and treatment. A false positive result can lead to additional unnecessary medical procedures. In the clinical study, CHART decreased the false positive rate from 21% to 16%, and also more significantly decreased false negatives from 50% to 34% in referral decision.

There is also a risk of misinterpretation of the output/reports by the user-GP which can be compounded by false positive/negative results. However, this risk can be mitigated through labeling to help GP's to understand how the system output should be interpreted and when to seek further care from a cardiologist.

The system provides the user a convenient and readily accessible means to record perform medical test during the any state of the person condition- during regular medical control examinations or when the symptoms are already present. The output/report can then be reviewed by a medical professional (GP) to determine if the symptoms may be related to cardiac abnormalities or establish some new risks that can lead to cardiac disease if not treated properly. This is especially valuable for users with recurrent, transient but infrequent symptoms, which can be difficult to catch with the traditional ECG, or during the regular medical examinations. The information can be helpful to make the medical evaluation more efficient and obviate some unnecessary procedures. Establishing the risk factors based on the report, early detection and prompt treatments are likely to improve clinical outcomes.

Overall, the probable benefits outweigh the probable risks given the available information concerning the benefits and risks. There is reasonable assurance of the safety and effectiveness for this system for the intended use.

The following statements are made for the device under evaluation and in view of the evaluated clinical data.

***The existing data are/are not sufficient to verify that the device is in conformity with all the Essential Requirements pertaining to clinical performance and clinical safety.***

The benefit/risk profile according to current knowledge/ the state of the art in the medical fields concerned and according to available medical alternatives is acceptable.



***The intended use and corresponding risk reduction measures are adequate and the product information is suitable for the intended users and sufficiently covers all usability aspects.***

All claims foreseen by the manufacturer are identified and any discrepancy and no gaps were identified in the clinical data.

***There is full consistency between the clinical data, the information materials supplied by the manufacturer and the risk management documentation for the device under evaluation***

Residual risks and uncertainties are sufficiently identified. The acceptability for CE-marking is sufficiently discussed and follow-up measures during PMS are addressed. (This includes uncertainties regarding medium- and long-term performance, safety under wide-spread use, residual risks such as side-effects and complications occurring at rates below detection possibilities of currently available clinical data, others).

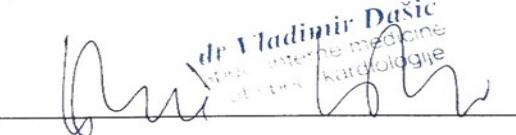
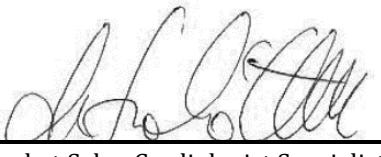
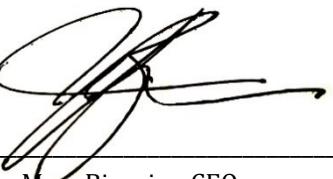
***PMS activities (including PMCF studies)***

For more details see **Annex J**- Post-Market Surveillance Plan.

## ***10 Date of the next clinical evaluation***

Date of next update	Justification of the date
As required.	No future Clinical Evaluation is planned or required at this time.

**11 Signatures of the responsible evaluators**

Signatures of the responsible evaluators		
Responsibility*	Date	Signature
<b>Coordinating clinical Investigator (Principal Investigator Vrsac)</b>	16 <sup>th</sup> May, 2019	 Dr Vladimir Dasic, Cardiologist Specialist
<b>Principal Investigator Sombor)</b>	16 <sup>th</sup> May, 2019	 Dr Tatjana Stankovic, Cardiologist Specialist
<b>Principal Investigator Senta</b>	16 <sup>th</sup> May, 2019	 Dr Aniko Berta-Szabo, Cardiologist Specialist
<b>Co-principal Investigator Senta</b>	16 <sup>th</sup> May, 2019	 Dr Erzsebet Sabo, Cardiologist Specialist
<b>Manufacturer</b>	17 <sup>th</sup> May, 2019	Cardio-Phoenix Inc. 44 Rosemead Close Markham, Ontario, Canada, L3R 3Z3 Phone: 416-595-0795  Marc Bisnaire, CEO

\*By signing this document, responsible agree with the contents of the report.



## 12 Qualification of the responsible evaluators

See Annex K- Qualification of the responsible evaluators.

## 13 References

- Abramowitz M., Stegun I.A.: Handbook of Mathematical Functions. Government Printing Office. 1964: 26.6.2
- Clinical Performance Assessment: Considerations for Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data- Premarket Approval (PMA) and Premarket Notification 510(k) Submission. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-performance-assessment-considerations-computer-assisted-detection-devices-applied-radiology>
- Dodd, Lori E., and Margaret S. Pepe. "Partial AUC estimation and regression." *Biometrics* 59.3 (2003): 614-623.
- E9(R1) Statistical Principles for Clinical Trials: Addendum: Estimands and Sensitivity Analysis in Clinical Trials, to determine the suitable statistical analysis we are going to use in our CS. Available from: <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM582738.pdf>
- GoffJr D.C., Lloyd-Jones D.M. et al. "ACC/AHA Guideline on the Assessment of Cardiovascular Risk." (2014). Available from: <https://www.ahajournals.org/doi/full/10.1161/01.cir.0000437741.48606.98>
- Guidelines on Medical Devices- Clinical Evaluation: A guide for manufacturers and notified bodies under directives 93/42/EEC and 90/385/EEC. MEDDEV 2.7/1 revision 4. Available from: <http://ec.europa.eu/DocsRoom/documents/17522/attachments/1/translations/en/renditions/native>
- HIPAA regulations, concerning Ethical considerations. Available from: <https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM582738.pdf>
- Hripcsak, George, and Adam S. Rothschild. "Agreement, the f-measure, and reliability in information retrieval." *Journal of the American Medical Informatics Association* 12.3 (2005): 296-298.



- Knowledge of Implementation of Cardiovascular Risk Clinical Practice by General Practitioners and Specialists. Available from:  
<http://www.revespcardiol.org/en/knowledge-and-implementation-of-cardiovascular/articulo/13092248/>
- Kraemer, H.C., et al. "Tutorial in Biostatistics- Kappa coefficients in medical research." Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, California, U.S.A (2002):2109-2129. Available from:  
<https://pdfs.semanticscholar.org/65c2/1cd4f2f33d7a0e4797c151cbf988b8c98d20.pdf>
- Lang, Thomas A., and Michelle Secic. "How to Report Statistics in Medicine: Annotated Guidelines for Authors." *The Nurse Practitioner* 22.5 (1997): 198.
- MatWorks, Syntax for determination of hypothesis calculation. Available from:  
<https://www.mathworks.com/help/stats/binofit.html>
- Safari, Saeed, et al. "Evidence Based Emergency Medicine; Part 5 Receiver Operating Curve and Area under the Curve." *Emergency* 4.2 (2016): 111.
- SigmaYone- "Understanding Binomial Confidence Intervals" to extract Confidence Interval. Available from:  
[http://www.sigmaxone.com/binomial\\_confidence\\_interval.htm](http://www.sigmaxone.com/binomial_confidence_interval.htm)
- Simundic, A.M. "Measures of Diagnostic Accuracy: Basic Definitions." US National Library of Medicine National Institutes of Health (2009): 203-211. Available from:  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4975285/>
- Sindhwan, V., Bhattacharjee, P., & Rakshit, S. (2001). Information theoretic feature crediting in multiclass support vector machines. In First SIAM international conference on data mining (ICDM'01). Chicago, IL, April 5-7.
- Statistical Guidance on Reporting Results from Studies Evaluation Diagnostic Tests. Available from: <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/statistical-guidance-reporting-results-studies-evaluating-diagnostic-tests-guidance-industry-and-fda>
- Statistical Principles for Clinical Trial- to establish valid Hypothesis. Available from:  
[http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Efficacy/E9/Step4/E9\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf)



- UCSF CHR criteria for qualification of the study risk. Available from:  
<https://irb.ucsf.edu/recruitment>
- Wei, Jin-Mao, et al. "A novel measure for evaluating classifiers." Expert Systems with Applications 37.5 (2010): 3799-3809.
- WMA Declaration of Helsinki-Ethical Principles for medical Research Involving Human Subjects. Available from: <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>