UNIVERSITETET
I OSLO

Project thesis

How many checklist items are required to determine if an ECG is normal or abnormal?

A retrospective cross sectional study

Åge Frivoll and Vetle Mørland

Medicine, profession
University of Oslo, Faculty of Medicine

Abstract 4
Objectives 4
Method 4
Results 4
Conclusion 4
Introduction 5
ECG Interpretation and the use of Checklists 5
Theory and Background 5
Introduction to the Cardiac Conduction System 5
The Normal Electrocardiogram (ECG) 6
Materials and Methods 8
Design 8
Checklist Material 8
Selecting Checklist Items 8
Regarding the Sources Used for the Compiled ECG Checklist 8
Compiling the Checklist 9
Included Checklist Items 9
Excluded Checklist Items 17
ECG Material 19
Database Selection 19
SCP codes 20
Database Export 20
Database Cleaning, Selection, and Randomization 20
Record Retrieval - Visualization and Export 21
ECG Interpretation using Compiled Checklist 21
Proofing Checklist and Method 21
Individual Interpretation 21
Comparison of Interpretations 22
Preparation for Analysis 22
Analysis of ECG Interpretation 23
Outcome Measures 23

Statistical Analysis 23
Prospective Verification of our Checklist 23
Results **25**
Discussion **26**
Clinical Relevance 26
Future Directions 26
Doctors or Machines? 26
Limitations 27
Conclusion **28**
Acknowledgments **28**
Bibliography **29**
Arbeidsfordeling - prosjektoppgave **31**

Abstract

Objectives

The purpose of this retrospective cross sectional study is to identify the minimal amount of checklist variables you must consider to diagnose an ECG as either normal or abnormal, and warrant further investigation. The aim is to combat the wide variability of checklists used by clinicians and health care professionals when interpreting ECG's by creating an evidence based checklist that both increase productivity and decrease the amount of error during ECG interpretation.

Method

We compiled a comprehensive set of ECG checklist items from well known sources such as UpToDate and American Heart Association that covers the most common ECG phenomenon. We then exported 308 adult ECGs with various diagnosis and normal ECG's from the PTB-XL database on physionet.org.

Further, we analyzed each of the 308 ECG's using the comprehensive checklist and categorized each variable as true or false according to diagnostic criteria, as well as each ECG as a whole with their status as normal or abnormal. From this comprehensive checklist we generated all possible sub-combinations of checklist items and selected the lists with the least amount of items necessary for diagnosing an ECG as abnormal in at least 95 % of the cases.

Following this we verified the selected checklists against the same 308 ECG's, only this time using the clinical remarks from the cardiologists annotating the database instead of our own interpretation. Using the cardiologists interpretation as gold standard, we want to find the specificity and sensitivity of the final checklists in the specific dataset, resulting in the most optimal checklist with regards to specificity and sensitivity.

Results

By evaluating every combination of checklist variables, excluding QTc, P amplitude, and P duration due to substantial variance between the interpretations, we derived six checklists, each comprised of seven variables. All lists exhibit similarity in the percentage of correctly identified true abnormalities, ranging from 95.3 % to 95.7 % of cases.

Further analyzing these lists against the clinical remarks we see that the list; «1) Rhythm, 2) frequency, 3) axis, 4) T inversion, 5) ST depression, 6) ST elevation and 7) Sokolow-Lyon's criteria for left ventricular hypertrophy» was slightly superior with a sensitivity of 89.0% and specificity of 77.8%

Conclusion

Recognizing the degree of bias in our study, and acknowledging many aspects that could be improved for future studies, we believe we can screen ECG's faster with a more condensed checklist. Thus using more time on the ECG's we mark as abnormal, of which a more thorough assessment is warranted. To finally evaluate this list, it should be tested out in a prospective clinical setting by health care profession with little ECG training, utilizing cardiologist as the study Gold Standard. Especially verifying the sensitivity and specificity of those lists, and determining an acceptable rate of false normal ECGs.

Introduction

ECG Interpretation and the use of Checklists

Electrocardiography is the most commonly performed cardiovascular diagnostic test (1), and interpretation of the electrocardiogram (ECG) is regarded as an important skill for most physicians. The ECG records the electrical activity of the heart and generates a visual representation, known as an ECG waveform or tracing, which displays the heart's electrical patterns on a millimeter grid. The ECG is an essential diagnostic tool used to detect cardiac abnormalities, such as arrhythmias, ischemia, and structural heart problems, providing valuable information for assessing cardiac function and guiding medical treatment decisions. Interpretation of an ECG is a complex task that requires knowledge of anatomy, electrophysiology, and pathophysiology, visual pattern recognition, and diagnostic reasoning.

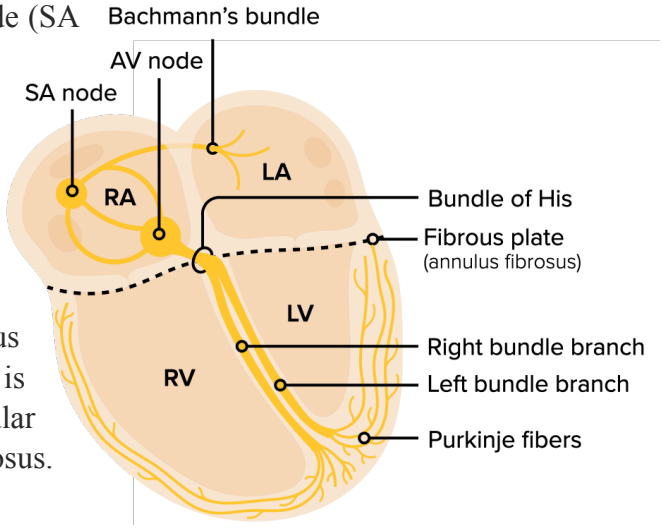
Considering this complexity, a checklist can be a valuable tool when interpreting ECG's, as it can provide a structured and systematic approach that reduces errors and enhances clinical communication. It may help to ensure that healthcare professionals does not overlook critical components of the ECG and form structure around the interpretation process. By following a checklist, it is believed that the chances of errors during interpretation will be reduced. This is particularly important during emergency situations when quick and accurate diagnoses and measurements are crucial. Checklists can also be helpful for healthcare professionals who are learning to interpret ECG's. It can provide a clear framework for learning and can be a good support while getting more comfortable and fluent in interpreting ECGs. Using a checklist can also save time as it guides one through the process and reduces the need for a second opinion. However, checklists should be used with care, maintaining a balance between checklist use and clinical judgment to ensure the most optimal patient care.

Theory and Background

Introduction to the Cardiac Conduction System

The cardiac conduction system is a large network of specialized cells and conduction pathways that coordinates the contraction and relaxation of the heart.

The electrical signal is generated in the sinoatrial node (SA node), located in the right atrium (RA). The SA node is often referred to as the heart's natural pacemaker. It generates electrical impulses at a regular rate, initiating each heart cycle. The electrical impulses spread throughout the right atrium and is conducted to the left atrium (LA) through a specialized electrical conduction pathway known as the Bachmann's bundle (7), leading to near simultaneous contraction of both the atria. The electrical signal is simultaneously blocked from entering the ventricular muscle by a thick fibrous sheath called annulus fibrosus. The only path through the fibrous sheath is through the atrioventricular node (AV node) located in the atrial septum. The AV node serves as a relay station, delaying the electrical impulses briefly



to allow the ventricles to fill with blood from the atria before contracting. Delaying the electrical conduction ensures proper coordination between atrial and ventricular contractions, and an optimal cardiac output. From the AV node the signal is relayed through a common conduction pathway, called the bundle of His, that serves as the only pathway between the atria and ventricles. The bundle of His quickly splits into the left and right bundle branches that runs in the ventricular septum, between the left and right ventricle. At the apex of the heart, the bundle branches split into multiple fine and rapidly conducting fibers called Purkinje fibers, that spread throughout the ventricle walls. During a heart contraction cycle, which includes both systole (contraction phase) and diastole (relaxation phase), it is essential that these components work in perfect synchronization (7).

The Normal Electrocardiogram (ECG)

Electrode placement and the 12-lead ECG

To reliable capture the electrical signals in the heart, electrodes are to be positioned at specific anatomical locations, registering the electrical potentials generated by the heart. Since the different electrodes record the electrical activity at specific locations, they serve as distinct "viewpoints," of the heart. The different "viewpoints" are shown in 12 different recordings, known as leads, on the echocardiogram. The standard 12-lead ECG consists of 3 limb leads (I, II, and III), 3 augmented limb leads (aVR, aVL, and aVF), and 6 precordial leads (V_1 through V_6).

Limb leads: The foundational trio of limb leads — I, II, and III — plays a vital role in monitoring the electrical activity in the frontal plane. This frontal plane perspective offers insights into the horizontal movement of electrical impulses within the heart (1).

Augmented limb leads: Complementing the limb leads are the augmented limb leads — aVR, aVL, and aVF. These are derived mathematically from the limb leads, hence giving additional perspectives and enhancing the diagnostic capabilities of the ECG (1).

Precordial leads: Precordial leads specifically capture the electrical activity in the horizontal plane, providing crucial information about the anterior, lateral, and inferior walls of the heart (1).

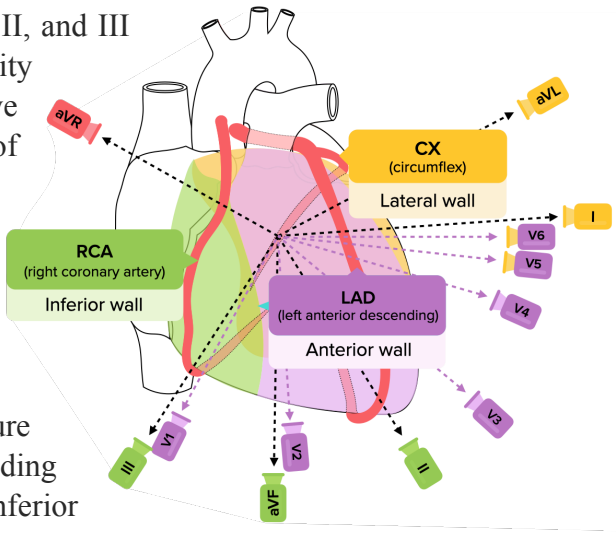


Illustration: © MedEasy Group AS

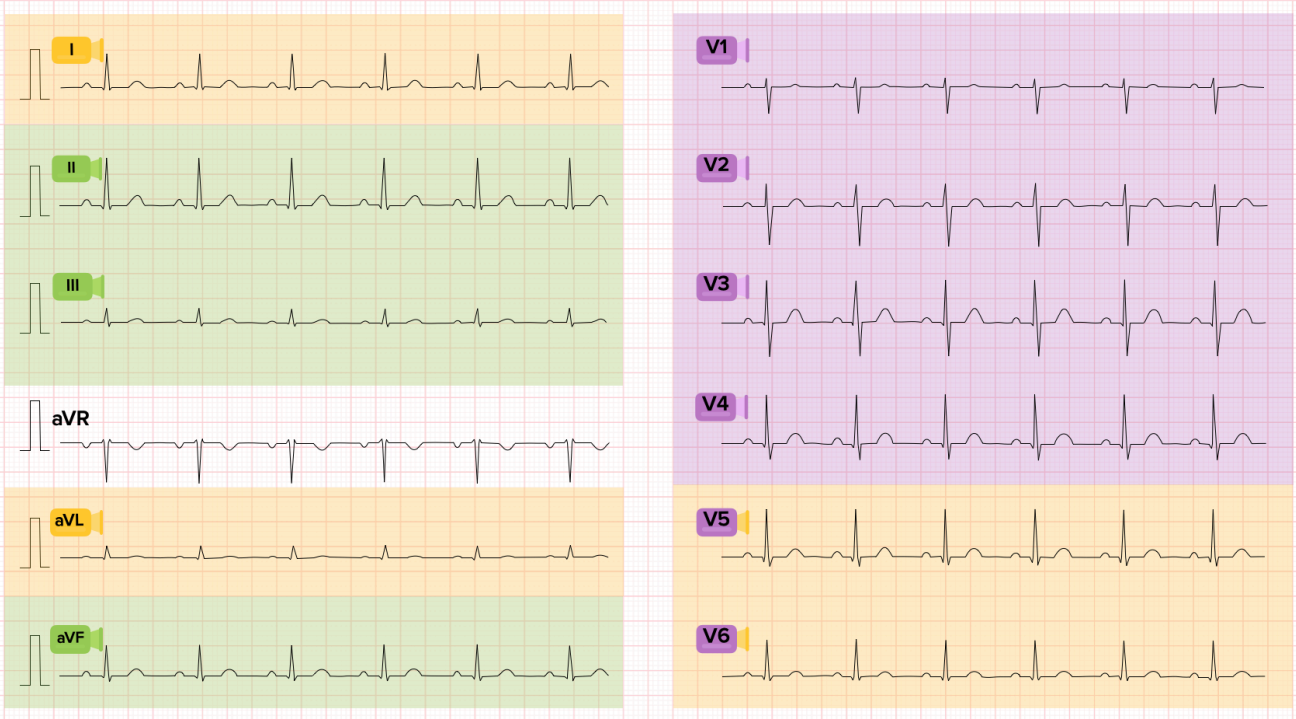
Anatomically contiguous leads

Anatomically contiguous leads are two or more leads that look at adjoining areas of tissue, which can aid in diagnostic accuracy as well as assessing the degree of the pathology.

Inferior leads: Leads II, III and aVF are leads that have their positive electrode located at the left foot. They are contiguous leads that all look at the inferior wall of the left ventricle (1).

Lateral leads: Leads I and aVL are leads that have their positive electrode located on the left arm. These leads views the superior lateral wall of the left ventricle. Leads V_5 and V_6 are situated on the left lateral aspect of the chest and view the inferior lateral wall of the left ventricle. Since Leads I, aVL, V_5 and V_6 all view the lateral wall of the left ventricle they are considered contiguous (1).

Figure 1 - Anatomically Contiguous Leads



The figure shows the different anatomically contiguous leads.
Anterior leads (purple): V1-V4. **Lateral leads (yellow):** I, aVL, V5-V6. **Inferior leads (green):** II, III and aVF
Illustration: © MedEasy Group AS (copied with permission)

Anterior leads: Leads V_1 and V_2 are positioned on each side of the sternum and can "look through" the right ventricle and see the septal wall. Leads V_3 and V_4 are on the anterior wall of the left chest which correlates with the viewing the anterior wall of the left ventricle (1).

Electrical axis of the heart

In electrocardiology, a vector shows both the size and direction of the electrical signal from a heart muscle cell. When all these vectors are added up, we get the electrical axis. Hence, we can interpret the axis for both the atrial and ventricular contraction. The left ventricle normally is normally the biggest part of the heart muscle and creates the largest vector on the ECG. Since this is the easiest to interpret and often carries the most clinical value, we usually talk about the ventricular axis when interpreting the electrical axis. All though this is most commonly used in regular medical practice, special placements of electrode to evaluate the atrial axis can be useful in some cases aswell. The cardiac axis can be appreciated by evaluating the QRS complexes, representing the ventricular activity, in specific leads. This will be discussed further within the materials section (1).

Materials and Methods

Design

We began by looking into different methods of ECG interpretation from different sources. This included a thorough dive into different systematic reviews, resulting in a compiled checklist that included many important parameters which may reveal pathology. Secondly, the focus shifted to the extraction and synthesis of the ECG material, including database selection, sub-sectioning, randomization, and visualization. Thirdly, we interpreted and analyzed the ECG's using the compiled checklist, noting any abnormalities and categorizing each ECG as normal or abnormal. In the fourth phase, we analyzed the data, bringing together the insights gained from the preceding phases, with the aim of putting together an efficient and standardized checklist. In the final step, we aimed to validate the checklist by analyzing our interpretation of the ECG with the selected checklist against cardiologist remark on each ECG provided by the database.

Checklist Material

Selecting Checklist Items

A challenge when creating a checklist that aims to confirm whether or not an ECG is normal or abnormal, is that there is varying degree of consensus as to what makes an ECG normal. Normality can rarely be based on simple parameters alone, and needs to be assessed together with factors such as age, physical health, gender and underlying conditions. Normality in ECG interpretation is therefore at a philosophical crossroads between diagnostic standardization and clinical perspective and experience.

Given the limited clinical information available, comprising of only age, sex, and the ECG itself, our approach will focus on standardizing as much as possible, in order to evaluate the ECG as normal or abnormal. When categorizing an ECG as normal, our main objective will be to have a high level of confidence in that diagnosis, aiming to avoid overlooking any potential pathology that could pose a significant risk to the patient. Accordingly, for abnormal ECGs, we recommend reinterpretation of the ECG by an expert. While an ECG may initially appear abnormal, an experienced clinician will also consider the clinical context in order to evaluate the abnormality as within the range of normality for this particular patient, or as a true pathology that needs to be addressed.

Regarding the Sources Used for the Compiled ECG Checklist

UpToDate

As a widely respected online clinical practice textbook, UpToDate provides healthcare professionals and researchers with up-to-date and evidence-based information on a wide range of medical topics. It serves as a comprehensive clinical decision support tool, offering in-depth articles, guidelines, reviews, and clinical updates on diseases, treatments, diagnostic methods, and more. It is regarded as the top level of information in clinical practice by the pyramid search at helsebiblioteket.no.

From UpToDate we have evaluated the follow articles:

- ECG tutorial: basic principles of ECG analysis (8)
- ECG tutorial: ST and T wave changes (9)
- ECG tutorial: Myocardial ischemia and infarction (10)

American Heart Association (AHA)

The American Heart Association (AHA) is a well-established nonprofit organization dedicated to cardiovascular health and reducing the impact of heart disease and stroke. It is a trusted source for both healthcare professionals and the general public seeking reliable information and guidance on matters related to heart health and stroke prevention. In the time periode of 2007-2009 they released a series of 6 documents regarding ECG standardization, interpretation and diagnosis which works as a basis for our checklist (1-6).

University of Oslo - online ECG course

The University of Oslo, representing the largest medical school in Norway, provides an online ECG course, highlighting the physiological processes which are the basis of ECG readings, and their clinical implications. The course aims to deepen the understanding of the ECG among medical students and early-career doctors. Main authors are professor Knut Gjesdal and cardiologist Mathis K. Stokke. (7)

Compiling the Checklist

After carefully examining all sources, we have established criteria for each checklist item outlined in the theory and background. This process initially led to the creation of a 22-item checklist. Due to low standardization and poor sensitivity and specificity of some checklist items, we eliminated 3 items. Our refined checklist ended up including 19 items, represented in Table 2. We will present the item criteria and our reasoning for inclusion and exclusion in the following section.

Included Checklist Items

Rhythm

The rhythm refers to the regularity or pattern of the heartbeats, which can either be a sinus rhythm or arrhythmic. A sinus rhythm in an ECG signifies a normal and regular cardiac cycle originating in the sinus node. It is characterized by upright and consistent P-waves, a normal PR interval, a narrow QRS complex and a regular and consistent rhythm.

UpToDate suggests fives steps in the rhythm analysis: 1) Locating the P waves, 2) Establishing the relationship between P waves and the QRS complex, 3) Analyzing QRS morphology, 4) Assessing the regularity of QRS complexes and 5) Interpreting the rhythm in the clinical setting present (8).
As our main objective is to capturing abnormal findings and not specific diagnoses, we have modified and merged some of the steps to create a more concise rhythm interpretation. As the QRS complex duration is included in a separated checklist item (table 1-4) and due to a lack of clinical information about the patients, we have chosen to exclude step 3 and 5 in UpToDate's approach. This leaves us with steps 1, 2 and 4 which is incorporated in the following checklist items.

One P wave precedes each QRS (steps 1 and 2)

To streamline and standardize our checklist, we have combined the first and second item in UpToDate's list into 'One P wave precedes each QRS complex throughout the lead strip'. "One P wave" to identify any arrhythmia where there is either to many P waves (eg. atrial fibrillation) or non at all (sinus arrest), and "precedes each QRS" as certain arrhythmias (eg. AVnRT and AVRT) may display a P wave after the QRS complex. Analyzing the entire lead strip is important for capturing some arrhythmias such as second degree AV block's and ectopic heart beats.

QRS regularity (step 4)
There is no universally established standard for defining 'normal QRS regularity' or 'acceptable RR interval variability' in the literature. Consequently, this checklist item introduces a degree of subjective interpretation bias. To mitigate this bias, efforts have been made to minimize it by cross-referencing results and comparing them to diagnostic coding in the ECG database.

Frequency
The frequency can be calculated by counting the number of millimeters between each QRS complex (ventricular depolarizations), also called the RR interval. A heart rate of 50-100 bpm is considered within the normal range. Tachycardia is generally defined as > 100 bpm (> 15 mm between each QRS complex). While some consider bradycardia to be a heart rate < 60 bpm (> 25 mm), Norway and most other countries consider rates of < 50 bpm (> 30 mm) to represent bradycardia (11). Hence, our checklist uses the criterion of < 50 bpm for bradycardia detection.

Table 1-1

	Checklist item	Normal	Abnormal
Rhythm and frequency			
1	One P wave precedes each QRS complex throughout the lead strip	Yes	No
2	QRS regularity	Yes	No
3	Frequency	50-100 bpm	< 50 bpm, > 100 bpm

P wave
The P wave represents atrial contraction and is typically a small, rounded wave that precedes the QRS complex. As we have already assessed the P waves quantitatively within the "rhythm" checklist item, our main focus is on abnormal P wave morphology. This can be caused by enlargement of either the right or left atria. As discussed earlier, the normal sinus P wave demonstrates depolarization originating in the sinus node. This depolarization spreads through the right atrium, creating the initial portion of the P wave. Simultaneously, the electrical signal is conducted through a specialized pathway known as Bachmann's bundle. However, it's important to note that the left atrial depolarization lags behind that of the right atrium. Consequently, the P wave may exhibit a notched appearance when the contraction delay is increased. This notched appearance may best be appreciated in lead II, hence this lead serves as the basis for our checklist criteria (5).

AHA proposes three criteria for capturing atrial enlargement and interatrial conduction disturbances. When the right atrium enlarges, the initial part of the P wave increases in amplitude. This is due to the involvement of a greater number of muscle cells contracting simultaneously. According to AHA, a P wave amplitude ≥ 0.25 mV in lead II is indicative of right atrial enlargement (5). The amplitude is measured from the isoelectric line to the highest point of the P wave (12). In the case of left atrial enlargement or interatrial conduction disturbances, the conduction rate in the Bachmann’s bundle is slowed. This results in a prolonged duration of the P wave, also amplifying the notched appearance in lead II at the same time. AHA recommends defining abnormal P wave duration as ≥ 120 ms, and the duration between the two peaks (the notched appearance to be ≥ 40 ms (5). The P wave duration is measured by tracing two vertical lines on the frontal plane leads, one marking the onset and the other the offset or the end of the P wave in any lead (12).

Table 1-2

Checklist item		Normal	Abnormal
PR interval			
4	Duration	< 120 ms	≥ 120 ms
5	Duration between peaks (lead II)	< 40 ms	≥ 40 ms
6	Amplitude (lead II)	< 0,25 mV	≥ 0,25 mV

PR interval

The PR interval represents the time it takes for the electrical impulse to travel from the sinus node, via the atria, through the AV node, and ultimately to the ventricles. It is measured from the beginning of the P wave to the beginning of the QRS complex (which may be a Q wave or R wave). PR intervals < 120 ms is defined as short and > 200 ms is defined as prolonged (8, 13). Measuring PR interval has importance in discovering AV bundle conduction disturbances (AV blocks) which generally increases the PR interval. Most of the AV blockages will give a sudden loss of QRS complex, which would be discovered by the previously mentioned rhythm criteria. However, a first degree AV block with consistent prolonged PR interval with no loss of QRS would not be discovered. This is the rationale behind the inclusion of the PR interval as an item.

Table 1-3

Checklist item		Normal	Abnormal
P wave			
7	PR interval	120-200 ms	< 120 ms, > 200 ms

QRS complex

The QRS complex visualizes the ventricular depolarization (contraction), and its width on the ECG measures the time it takes for the electrical impulse to travel through the ventricles. The QRS width is usually measured in the precordial leads where it at its widest. To emphasize the focus on bundle branch conduction and any abnormal conduction shunts, the QRS width are to be measured in a QRS following a P wave. This is to ensure the duration of conduction is not overestimated in cases of ventricular extra systoles (VES).

A normal QRS duration is typically less than 110 milliseconds, according to the American Heart Association (8). If the QRS duration extends beyond 110 milliseconds, it is considered abnormal. The QRS width of 110-119 are defined as incomplete bundle branch blockages, which is a very narrow range to interpret with certainty manually using the ECG grid. Hence, the incomplete cases are more easily assessed using automatic electronic interpreting software as they usually calculate and compare all leads simultaneously. For out part, when clinically diagnosing a complete bundle branch block, the threshold for abnormality is usually set at 120 milliseconds (3, 27). UpToDate also follows this 120-millisecond definition for an abnormal QRS length and results in our checklist criteria (8).

Table 1-4

	Checklist item	Normal	Abnormal
QRS complex			
8	Duration	≤ 120 ms	> 120 ms

ST segment

The ST segment is normally a flat, isoelectric segment on the ECG following the QRS complex and preceding the T wave due to very low electrical activity at that time point. The amplitude of the ST segment is measured by the difference in amplitude between the isoelectric line and the J-point (7). The isoelectric line can best be estimated in the segment after the T-wave and before the P wave. The J point is the rapid change of slope, or junction, between the end of the QRS and the beginning of the ST segment. ECG checklists are typically aimed to reveal elevation or depression of the ST segment, as these can both be caused by a number of different conditions, but most noteworthy ischemic heart diseases.

ST segment elevation refers to an upward shift of the J-point from the isoelectric line and functions as a crucial marker in diagnosing specific cardiac conditions that might need urgent treatment. The threshold values for ST segment elevation vary depending on factors such as gender, age, and the specific ECG lead being examined (6). Our checklist items for assessing abnormal ST elevation consider these variations as seen in Table 1-5. To diagnose ST elevation myocardial infarction (STEMI), it is required that the ST elevation changes are present in two or more anatomically contiguous leads as discussed earlier (6). However, changes in only one lead are also considered abnormal, as they may indicate nonspecific ischemia and non-ST elevation myocardial infarction (NSTEMI). Therefore any ST segment change is categorized as abnormal as we simply aim to identify abnormal changes.

The morphology or shape of the ST segment can provide valuable insights for differentiating between various cardiac diagnoses. Upsloping ST elevations are for example considered benign compared to flat or down sloping segments (6). However there are no standardized and specific criteria for assessing the morphology. Therefore we have not included this in our checklist.

Table 1-5

	Checklist item	Normal	Abnormal
ST segment			
9	ST elevation		> 0,1 mV in all leads, except for V_2/V_3 where it is: > 0,15 mV in females, > 0,2 mV in men older than 40 years and > 0,25 mV in men less than 40 years
10	ST depression		< - 0,1 mV in all leads, except for V_2/V_3 where it is < - 0,05 mV

T wave

The T wave follows the ST segment on the ECG and represents the repolarization of the ventricles. The normal T wave points the same way as the major deflection of the QRS (either R- or S-wave) and has an amplitude within the normal range. As we've established the depolarization of the ventricle (QRS complex) begins at the endocardial surface and spreads to the epicardium. The repolarization wave, on the other hand, begins at the epicardial surface and spreads to the

endocardium. Hence, the direction of ventricular depolarization is opposite to that of ventricular repolarization. The fact that opposite vector directions gives the same deflection on the ECG might seem counterintuitive at first. However it can be explained by the fact that the repolarization wave is negatively charged, thereby opposite to the positive charge in the depolarization wave. With the charge being negative, the lead deflection will be opposite to the direction of the repolarization. Thus, the T wave vector on the ECG normally is in the same direction as the major deflection of the QRS. Another way of saying this is that the QRS and T wave axes are concordant.

When the QRS and T wave point opposite to each other, they are disconcordant or inverted, and this is an abnormal feature. However, T wave inversion is considered normal in some leads, but the specific leads in question vary depending on the source. AHA uses a quantitative approach and defines T wave inversion as significant when the T-wave amplitude is ≥ - 0.1 mV in leads I, II, aVL, and V_2 to V_6. Inversions in leads III, aVF, aVR, and V_1 are therefore considered normal by AHA criteria (4). AHA's definition does not account for the polarity of the QRS complex. If the QRS complex was negatively deflected, a negative T wave would be concordant to the QRS, but defined as disconcordant by the AHA's criteria. UpToDate are pragmatically oriented and defines T wave inversion simply as abnormal when the T wave is opposite to the QRS deflection in all leads except in leads V_1-V_3 (9). While UpToDate does not provide specific amplitude values, it emphasizes the discordant relationship of the T wave with the QRS complex.

Interpreting isolated T-wave abnormalities presents a considerable challenge. The interpretation of such abnormalities in isolation can result in ambiguity and inaccurate diagnoses, especially concerning myocardial ischemia and infarction. Some references suggest a connection between cardiac pathology and the presence of two or more T-wave inversions in adjacent leads (7). Consequently, our checklist is designed to identify T-wave discordance in two or more anatomically contiguous leads to reduce false abnormal ECG's.

Symmetrically peaked T waves with increased amplitude are also considered abnormal and may indicate severe hyperkalemia. Though specific criteria may vary, AHA suggests that T waves in V_2 are abnormal if they exceed 1.4 mV in males and 1.0 mV in females, with variations for different age groups (1.6 mV in males aged 18-29) (4).

Biphasic T-waves are T waves characterized by a two-phase pattern. It is generally associated with ischemia, and particularly associated with proximal stenosis in the left anterior descending artery (LAD) when seen in V_2 and V_3 (14).

Table 1-6

	Checklist item	Normal	Abnormal
T wave			
11	Inversion	No	≥ 2 disconcordant T waves to preceding QRS in leads I, II, III and V_3-V_6.
12	Amplitude		> 1.4 mV in V_2 (1.6 mV in 18-29 years), females: > 1.0 mV in V_2
13	Biphasic	No	Yes

U wave

The U wave is a small, often subtle wave that may be visible after the T wave in the ECG. Even though the origin and clinical significance of U waves is not entirely understood, multiple hypotheses have been proposed. Late depolarization, delayed repolarization from M-cells and stretch induced depolarization during the filling phase are some of the well established attempts to understand the U wave (15). This reflects the uncertainty with regards to the mechanism of the U-waves. The normal U wave are usually best seen in leads V_2 and V_3, and are apparent with a slower heart rate and in individuals with chronic hypertension. This is why U waves are observed in over 90% of patients with a heart rate below 65 bpm. (4,15).

Abnormal U-waves can be observed in a broad spectrum of heart diseases such as coronary artery disease with ongoing myocardial ischemia or infarction, ventricular hypertrophy, congenital heart disease, primary cardiomyopathy and valvular defects. U waves are considered abnormal if they are inverted relative to the QRS complex or if the amplitude is the same as, or higher than the preceding T wave (16). The American Heart Association (AHA) recommends specific criteria for identifying abnormal U waves. These include inversion relative to the QRS complex or having an amplitude equal to or higher than the preceding T wave.

Table 1-7

	Checklist item	Normal	Abnormal
U wave			
14	Inversion		Disconcordant to preceeding QRS
15	Amplitude		≥ amplitude of preceding T wave

QTc interval

The QT interval is the segment from the initiation of the QRS complex to the end of the T wave and includes the entire ventricular depolarization and repolarization process. As the QT interval varies in relation to heart rate, we utilize QTc which is the QT segment time corrected for heart rate. A prolonged QTc interval is associated with an increased risk of ventricular arrhythmias, and therefore of clinical relevance.

Accurately measuring the QT interval can be challenging due to variations in ECG placements on the skin. The onset of the QRS complex and the end of the T wave may be challenging to pinpoint exactly, and the duration may differentiate across different leads, and also within the same lead. In practice, variances up to 50 ms in QT interval duration are often considered normal (4). Since the QT interval in leads V_2 and V_3 tends to be of the longest duration (4), and we are assessing the intervall in a single lead, we measure the QT interval in one of these leads.

Together with the lead to lead variability when assessing the QT interval, the normal QT interval is influenced by factors such as gender, heart rate and QRS duration (4). The normal QT interval ranges differ between men and women to account for gender-related differences. To correct for heart rate as mentioned earlier, different formulas are used to best calculate the QTc interval. Bazett's formula is considered the standard for heart rates between 50 and 100 bpm, while Fridericia's formula is preferred for other heart rate ranges (16). As the QT interval is measured from the onset of the QRS complex to the end of the T wave, QRS duration can affect the

estimation of the QT interval. Therefore, in cases of prolonged QRS duration we use the Bogossian's formula for a more accurate correction (17), followed by a correction of heart rate using the previously mentioned Bazetts formula.

Table 1-8

Checklist item		Normal	Abnormal
QTc interval*			
16	Duration	Males: 390-450 ms Females: 390-460 ms	Short: < 390 ms Prolonged: males: > 450 ms, females > 460 ms

*Formulas for correcting QT interval:
Bazetts formula: when heart rate is 50-100 bpm: $QT_{time}/RR^{1/3}$.
Friderichias formula: when heart rate is <50 and > 100: $QT_{time}/\sqrt{RR_{interval}}$.
Bogossians formula: when QRS > 0.12 ms: $QT_{time} - (0.485 * QRS_{ms})$, followed by QTc correction using Bazetts formula exclusively

Axis

The cardiac axis represents the overall direction of electrical conduction in the heart. As previously mentioned, we mainly focus on the ventricular axis which is given by the QRS complexes. It is usually given in degrees and indicates whether the heart's electrical activity is primarily oriented towards the right, left, or within the normal range. Deviations from the normal axis can suggest heart conditions, but is highly dependent on the clinical setting and presenting complaint as it can be a normal physiological sign.

The normal electrical axis of the ventricles typically falls within the range of +90 to -30 degrees. Deviations from this range can signal underlying cardiac conditions and guide clinical diagnosis. An axis exceeding +90 degrees is categorized as a right axis deviation. Conversely, an axis falling below -30 degrees is termed a left axis deviation. In cases where the axis spans the region from +180 to -90 degrees, is categorized as an extreme deviation, commonly referred to as a northwest axis.

An efficient method for approximating the electrical axis involves an evaluation of leads I and aVF in the quadrant test (18). By assessing the predominant deflections (positive or negative) in these leads, an approximate axis can be determined. For precise axis determination in degrees, a more comprehensive analysis, incorporating additional limb leads, may be required. However, for the fundamental purpose of diagnosing normal and abnormal ECG's, this technique provides an efficient approach.

Table 1-9

Checklist item		Normal	Abnormal
Axis			
17	Ventricular axis	Positive lead I, positive lead aVF	**Left:** Positive lead I, negative lead aVF **Right:** Negative lead I, positive lead aVF **Extreme:** Negative lead I, positive lead aVF

Left ventricular hypertrophy

Left ventricular hypertrophy (LVH) is a condition characterized by the thickening of the muscular wall of the left ventricle in the heart. Evaluation of QRS amplitude in ECG's primarily reflects the combined electrical activity of the left ventricle. An increased amplitude of the QRS in specific leads can be indicative of ventricular hypertrophy. The reasoning why the amplitude reflects primarily the left wall, is that physiologically the left ventricle is larger and thicker then the right, resulting in a much bigger electrical output on the ECG then the right ventricle. While left ventricle hypertrophy usually is physiological in younger age groups, this cardiac adaptation can occur in response to various factors and diseases, including chronic high blood pressure, valvular heart disease, or genetic predisposition. Determining LVH based on ECG is difficult, as there are multiple confounding factors as age, obesity, gender etc. This affects the ability to accurately indicate LVH based on ECG alone. As such, the ECG must be seen as a screen tooling, that have to take clinical information into account to warrant further investigation (19).

Several methods of assessing the QRS amplitude have been proposed in order to most accurately predict LVH. The sensitivity of these methods, which represents the ability to correctly identify LVH when it is present, is generally low and often falling short of 50 % for many criteria (5). This means that a negative finding for LVH, cannot rule out hypertrophy in itself. On the other hand, the specificity, which measures the ability to correctly identify patients without LVH, is typically high, ranging from 85 % to 90 % (5). However, it's important to note that the sensitivity and specificity of each criterion can differ significantly, affecting the overall diagnostic accuracy (5). This variation in sensitivity and specificity leads to patients meeting one set of LVH criteria while not meeting others.

The two most used criteria for assessing LVH today is the Sokolow-Lyon and Cornell criteria, which assesses QRS amplitude in aVL and specific precordial leads. Because of the above-mentioned poor sensitivity of the individual criteria, we have chosen to include both of these criteria in our study. The Sokolow-Lyon criteria is assessed either by adding together S wave amplitude of V_1 with the R wave amplitude of the tallest of V_5 or V_6, or solely by the R wave amplitude in aVL (5,8,20). The Cornell criteria is the sum of the R wave in aVL and the S wave of precordial lead V_3. It is important to note that the S wave of the precordial leads is measured from the isoelectric baseline, not from the peak of the R wave.

The normal ranges varies depending on factors such as age, gender and ethnicity. Typically, the established QRS voltage criteria are designed for individuals aged 35 years and older. However, standards for those between the ages of 16 and 35 remain less firmly established, and relying solely on voltage-based criteria for diagnosing (5). LVH in this age group yields low accuracy, especially in trained athletes. The Cornell criteria differentiate between female and male normal values, while the Sokolow-Lyon criteria uses identical threshold for both genders.

Table 1-10

	Checklist item	Normal	Abnormal
Left ventricular hypertrophy			
18	Sokolow-Lyon's criteria		S wave in V_1 + R wave in V5/V6 ≥ 3.5 mV R wave in aVL ≥ 1.1 mV
19	Cornell's criteria		S wave in V_3 + R wave in aVL: Men: > 2.8 mV, women: > 2.0 mV

Excluded Checklist Items

R progression

R progression refers to the evolving pattern of the R waves in the precordial (chest) leads of the ECG. Normally, R waves become progressively larger as you move from lead V_1 to V_6. This is a highly subjective interpretation and difficult to quantitatively measure, therefore excluded from the final checklist. Pre-analytic bias can also make it difficult to assess the R progression as wrong precordial electrode placement can affect the QRS amplitude in the different leads.

Q wave

Q waves are small, initial downward deflections of the QRS complex. Pathological Q waves are typically wider and deeper and can indicate previous or acute onsetting myocardial infarction as well as specific genetic hear abnormalities. However there is a lack of standardization as to what is considered normal or pathological Q-waves.

Right ventricular hypertrophy

Excluding right ventricular hypertrophy (RVH) as a checklist item is justified due to the limited sensitivity of ECG in detecting RVH. Echocardiograms face challenges in measuring the complex 3-dimensional shape of the right ventricle and the thickness of the free right ventricular wall. While RVH can alter the QRS vector and cause delays in right precordial leads, the dominance of left ventricular activation in a normal hear, and especially in the case of left ventricular hypertrophy, makes it difficult to identify RVH using ECG alone (5).

Table 2 - Compiled Checklist

	Checklist item	Normal	Abnormal
Rhythm and frequency			
1	One P wave precedes each QRS complex throughout the lead strip	Yes	No
2	QRS regularity	Yes	No
3	Frequency	50-100 bpm	< 50 bpm, > 100 bpm
P wave			
4	Duration (lead II)	< 120 ms	≥ 120 ms
5	Duration between peaks (lead II)	< 40 ms	≥ 40 ms
6	Amplitude (lead II)	< 0,25 mV	≥ 0,25 mV
PR interval			
7	Duration	120-200 ms	< 120 ms, > 200 ms
QRS complex			
8	Duration	≤ 120 ms	> 120 ms
ST segment			
9	ST elevation		> 0,1 mV in all leads, except for V_2/V_3 where it is: > 0,15 mV in females, > 0,2 mV in men older than 40 years, > 0,25 mV in men less than 40 years
10	ST depression		< - 0,1 mV in all leads, except for V_2/V_3 where it is < - 0,05 mV
T wave			
11	Inversion	No	≥ 2 disconcordant T waves to preceding QRS in leads I, II, III and V_3-V_6.
12	Amplitude		> 1.4 mV in V_2 (1.6 mV in 18-29 years), females: > 1.0 mV in V_2
13	Biphasic	No	Yes
U wave			
14	Inversion		Disconcordant to preceeding QRS
15	Amplitude		≥ amplitude of preceding T wave
QTc interval*			
16	Duration	Males: 390-450 ms Females: 390-460 ms	Short: < 390 ms Prolonged: males: > 450 ms, females > 460 ms
Axis			
17	Ventricular axis	Positive I, postive aVF	Left: Positive I, negative aVF Right: Negative I, positive aVF
Left ventricular hypertrophy			
18	Sokolow-Lyon's criteria		S wave in V_1 + R wave in V5/V6 ≥ 3.5 mV R wave in aVL ≥ 1.1 mV
19	Cornell's criteria		S wave in V_3 + R wave in aVL: Men: > 2.8 mV, women: > 2.0 mV

*Formulas for correcting QT interval: Bazetts formula: when heart rate is 50-100 bpm: $QT_{time}/RR^{1/3}$. Friderichias formula: when heart rate is <50 and > 100: $QT_{time}/\sqrt{RR_{interval}}$. Bogossians formula: when QRS > 0.12 ms: $QT_{time} - (0.485 * QRS_{ms})$, subsequently correcting the adjusted QT time with Bazetts formula exclusively independent on heart rate.

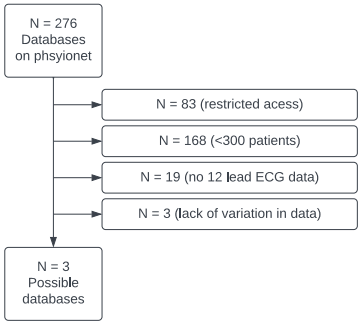
ECG Material

Database Selection

Databases on physiological data are popularly accessed through PhysioNet due to their open access, easy accessibility and programming libraries that support analysis, export and visualization.

Their overview lists 276 different databases including various physiological parameters, diseases, recording situations, etc (20). When selecting a database we wanted it to be free of charge, ensure adequate variability in the data set and contain relevant data. Variability in age, gender and disease are important as we want to test our checklist in regards to most scenarios that could occur in a normal clinical setting for the general population. Hence, we consecutively excluded databases with the following criteria:

1. Restricted access (n = 83)
2. < 300 patients (n = 168)
3. No 12 lead ECG data (n = 19)
4. Lack of disease variation (n = 3)



From this we further investigated the following three databases:

MIMIC-IV-ECG: Diagnostic Electrocardiogram Matched Subset

The database contains approximately 800 000 diagnostic ECG's from around 160 000 unique patients. The database provides ECG information, but to obtain further clinical information such as demographics, diagnosis, medication, a premium access to the MMIC-IV clinical database was required. The database was therefore regarded as restricted in regards to essential information (sex and age) we needed to interpret the ECG's and was consequently excluded.

A large scale 12-lead electrocardiogram database for arrhythmia study

The 12 lead ECG database for arrhythmia study is a database originally consisting of 10 646 patients and is based in the .csv format shared at FigShare. With the conversion to PhysioNet the database has been expanded to 45 152 ECGs and converted to the WFDB format. This makes data handling and export easier, as the python WFDB library supports handling of many parameters including ECG's. Additionally, each record has a header file providing important information such as sex, age, and SNOMED-CT codes (diagnostic codes annotated by cardiologists). More information on the database can be obtained in the reference (21).

PTB-XL, a large publicly available electrocardiography dataset

The PTB-XL database is a wide database of 21 799 12 lead ECGs from 18 869 patients annotated by two cardiologists with SCP-ECG diagnostic standard codes (mentioned later). With extensive annotation to each ECG, it is designed to train machine learning algorithms. More information on the database can be obtained in the reference (22).

As the MIMI-IV-ECG database was ruled out, the remaining two databases was evaluated. With both having the combination of annotation by cardiologists and a vast dataset with numerous ECG phenomena, they see equal suitability for our study. Nevertheless, the metadata and data structure of the PTB-XL database was deemed easier to comprehend and work with when planning to export the ECG's. Additionally, in the proof-of-concept phase before initiating the project, a preliminary script to export the PTB-XL database, Lobachevsky database, and MIMIC-Demo database was already

created. Therefore we choose the PTB-XL database as it provided an efficient way of going forward with the study at no obvious disadvantage.

SCP codes

SCP-ECG (standard communications protocol for ECG) is a way of conveying the necessary information for ECG analysis and interpretation in a standardized way. SCP codes are assigned ECG recordings to highlight specific findings, diagnosis's and abnormalities. This provides a standardized way to describe various cardiac abnormalities, arrhythmias, conduction disturbances etc. that can easily be accessed and used by programs in the analysis of the ECG. However, the specific codes and their meanings may vary depending on the version of the SCP standard or the particular ECG interpretation system or software being used (23).

Database Export

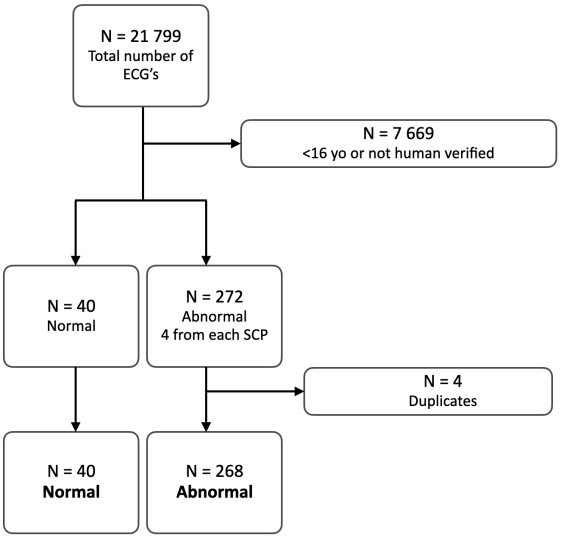
The overall database data structure was downloaded directly from PhysioNet as a .csv file named on the 3rd of October 2023. The file have the MD5 encryption code of: 62f764c8be6aa5df7202230fd51 23064 for later reference and replication.

Having obtained this file to a local system, we used Python (v. 3.9.7) in the PyCharm editor (v. 2021.3.1) to extract the data of all 21 799 ECG records. An important part of this data are the SCP-codes, which will provide us with the ability of ensuring diversity in our dataset.

Database Cleaning, Selection, and Randomization

Once imported, preliminary data cleansing was carried out to ensure data variability and relevance. To avoid potential biases arising from repeated patient records, we deleted any recording other than the first instance from the same patient. We also excluded all patients under the age of 16 as ECG interpretation in pediatric patients follow different criteria as to what is normal in adults.

Furthermore, we only included ECG's that had been verified by a human as this might give us valuable post-analysis insight.



Out of the 21 799 ECG records, 14 130 met our inclusion and exclusion criteria above. Out of these 14 130 we wanted to choose approximately 300 ECGs as an arbitrary number that stretches our capacity besides studies to a maximum. To ensure variability when severely limiting the database we did a randomization and selection based on the SCP codes.

The script was configured to randomly select 40 normal ECG's and 4 ECG's for each SCP code. This would ensure a diverse and representative sample for further analysis, with all possible diagnosis included. For each SCP code we made a list of all ECG's that are assigned with that specific code. This list was then randomly shuffled, using Python's random library, and extracting the top 4 ECG's. From this the script provided us with 272 abnormal ECGs. Following this we checked for duplicates due to the fact that some ECG where

tagged with multiple SCP codes, where a total of 4 duplicate ECG's were deleted. Lastly all the normal and pathological ECG's were collected in one list and shuffled to avoid grouping of similar diagnoses.

The meta-data now included the ECG ID, gender and sex, and was exported as a CSV together with the original ECG identifiers in order to reverse the dataset shuffle post-analytically.

Record Retrieval - Visualization and Export

For each selected ECG, the script retrieve the desired record from the main database online and visualize the ECG using the graphical WFDB library. After visualization, each ECG was saved as a PDF using the MatPlot library. The files were stored in a shared directory in the cloud service iCloud with both authors, with each filename bearing the correct age and sex as they are important variables for the interpretation.

This process was verified before the final export by manually comparing a test export of 10 ECG's with the publicly available waveform visualization at the PhysioNet webpage (24). The complete script, covering the process from importing the database to exporting the ECGs, can be obtained from the bibliography (25).

ECG Interpretation using Compiled Checklist

Proofing Checklist and Method

To ensure uniform understanding of each checklist item and minimize inter-individual variation, we co-interpreted 15 randomly selected ECGs from the database after individual interpretations were made. This comparison and co-interpretation resulted in necessary minor adjustments adjusting for criteria misinterpretation, variations in parameter measurement practices (e.g. QT interval), and adjusting to minor differences in the ECG export format compared to our normal clinical setting.

Individual Interpretation

A total of 308 ECG's were independently analyzed by each interpreter to prevent influence from each other. The numerical and boolean values for each checklist item was noted in identical and

Table 3 - Values for Simplifying the Checklist Values

Checklist item	Changed to 1	Changed to 0	Checklist item	Changed to 1	Changed to 0
P before each QRS	y	n	Axis	n	l, r, ex
QRS regularity	r	u	Sokolow	no	yes
Frequency	normal	takycardia, bradycardia	Cornell	no	yes
QRS complex duration	normal	prolonged, short	P peak interval	normal, N/A	abnormal
PR interval	normal	short, abnormal	T amplitude	no	yes
ST elevation	no	yes	T biphasic	no	yes
ST depression	no	yes	U inversion	no	yes
T inversion	no	yes	U amplitude	no	yes

separate excel sheets. The excel sheets was programmed to convert the numerical values to dichotomous values based on the specific checklist criteria as highlighted in Table 3.
The randomized association between the ECG's and their original counterparts was securely stored locally on one of the interpreter's computer and made inaccessible until the ECG interpretations were finalized.

Comparison of Interpretations

Following this we each looked at the ECGs individually and carefully compared our interpretations. In cases where we were unclear if an ECG was normal/abnormal or there was an big differences in values, we reviewed the ECG together to figure out the most accurate interpretation. If we still were in disagreement after comparing results, and looking at the ECG in question, we concluded on the average of our two values.

Preparation for Analysis

While comparing interpretations and working towards a consensus on our individual checklists, we noted significant differences in our assessments of certain checkpoint items, some exhibiting more variability than others. Due to the observed variability in our interpretations, we examined the variability of the checklist items more closely. Our primary variables of concern were 'Cornell,' 'P amplitude,' 'P duration,' and 'QTc interval' (Table 4). These discrepancies frequently influenced whether a variable was categorized as normal or abnormal.

Table 4 - Interpretation Variance

	Checklist item	Variance		Checklist item	Variance
1	U amplitude	0 %	11	ST elevation	11 %
2	Frequency	4 %	12	ST depression	11 %
3	Sokolow	4 %	13	P peak interval	12 %
4	QRS regularity	5 %	14	P before each QRS	15 %
5	Cornell	6 %	15	T inversion	19 %
6	T amplitude	6 %	16	PR interval	23 %
7	U inversion	6 %	17	P amplitude	27 %
8	QRS time	7 %	18	QTc interval	33 %
9	T biphasic	8 %	19	P duration	42 %
10	Axis	9 %			

- *Cornell:* One interpreter inconsistently interpreted the S wave in V_3, opting for the larger of the S or R wave instead. In order to mitigate this issue, all ECG's with significant Cornell was re-checked by the interpreter in V_3 before consensus was discussed.
- *QTc interval:* As mentioned earlier, QTc is calculated using various algorithms adjusted for heart rate. Small variances in QT time can affect the significance level of QTc. Additionally, QTc is typically machine-calculated, averaging multiple leads, which is impractical in a regular ECG interpretation scenario. We also observed that the slightest variance in QTc in one lead could result in an difference in interpretation of normal, short, or long when calculating QTc, even within the same lead.
- *P amplitude and P duration:* here there were also major difference between the interpreters, mostly due to subjective ability to interpret p-waves where they are not clearly present.

Due to the QTc interval, P duration, and P amplitude showing a variability exceeding 25% (Table 4), we opted to exclude these variables. Including them would introduce significant noise into the

dataset and would probably lead to the result that many ECG's would be labelled abnormal, without showing true abnormality (false positives).

Analysis of ECG Interpretation

Outcome Measures

Our primary outcome is to create a checklist which detects 95 % of abnormal ECG's. Hence, our objective is to to find the shortest combination of all possible checklist items combinations that ensures this level of accuracy.

Statistical Analysis

For every possible combination of variables, we assess all ECG's to determine whether the combined checklist items in that specific combination categorize the ECG as normal or abnormal. Subsequently, we calculate the percentage of abnormal ECG's identified by that particular combination compared to including all checklist items. Based on these results, we choose the shortest lists for further investigation.

Prospective Verification of our Checklist

In the concluding phase, we aim to verify the top checklist combinations generated in the previous phase by comparing it with the SCP codes and clinical remarks annotated in the database. A script will run through all thee ECG's and automatically assigned the ECG as normal or abnormal based on SCP codes where 'NORM':100 was interpreted a completely normal ECG, while SCP codes in Table 5 where regarded as abnormal.

From this overview of normal and abnormal categorization based on the SCP codes, we compared the normal/abnormal status from our specific checklists against the normal/abnormal status from the SCP's interpretation, and calculated the sensitivity and specificity with the following formulas:

- Sensitivity = true positive / (true positive + false negative)
- Specificity = true negative / (true negative + false positive)

The Python script as a whole, along with adjacent files underpinning the analysis of the ECG's, can be obtained from the bibliography (26).

Table 5 - SCP Diagnostic Statements regarded as Abnormal

SCP	Meaning	SCP	Meaning	SCP	Meaning
1AVB	first degree AV block	**INVT**	inverted T-waves	**LVH**	left ventricular hypertrophy
2AVB	second degree AV block	**IPLMI**	inferoposterolateral myocardial infarction	**PMI**	posterior myocardial infarction
3AVB	third degree AV block	**IPMI**	inferoposterior myocardial infarction	**PSVT**	paroxysmal supraventricular tachycardia
AFIB	atrial fibrillation	**ISCAL**	ischemic in anterolateral leads	**RVH**	right ventricular hypertrophy
AFLT	atrial flutter	**ISCAN**	ischemic in anterior leads	**SARRH**	sinus arrhythmia
ALMI	anterolateral myocardial infarction	**ISCAS**	ischemic in anteroseptal leads	**SBRAD**	sinus bradycardia
AMI	anterior myocardial infarction	**ISCIL**	ischemic in inferolateral leads	**STACH**	sinus tachycardia
ASMI	anteroseptal myocardial infarction	**ISCIN**	ischemic in inferior leads	**STD_**	non-specific ST depression
BIGU	bigeminal pattern (unknown origin, SV or Ventricular)	**ISCLA**	ischemic in lateral leads	**STE_**	non-specific ST elevation
CLBBB	complete left bundle branch block	**LAFB**	left anterior fascicular block	**SVARR**	supraventricular arrhythmia
CRBBB	complete right bundle branch block	**LMI**	lateral myocardial infarction	**SVTAC**	supraventricular tachycardia
ILMI	inferolateral myocardial infarction	**LNGQT**	long QT-interval	**TRIGU**	trigeminal pattern (unknown origin, SV or Ventricular)
IMI	inferior myocardial infarction	**LPFB**	left posterior fascicular block	**VCLVH**	voltage criteria (QRS) for left ventricular hypertrophy
INJAL	subendocardial injury in anterolateral leads	**LPR**	prolonged PR interval	**WPW**	Wolf-Parkinson-White syndrome

SCP codes obtained from the PTB-XL database metadata (22).

Results

The database export and our interpretation of the ECG's can be found at the thesis Github page (26). From this, the calculation excluded all possible combinations of <7 variables from the results as they didn't fill the requirements of catching > 95% of the abnormal ECG's. Out of the checklists containing 7 items, only 6 combinations made it through the requirement, shown in Table 6.

Table 6 - Checklists Capturing 95 % of Abnormal ECG's

Checklist	1st item	2nd item	3rd item	4th item	5th item	6th item	7th item	% abn. captured
1	Rhythm	Frequency	Axis	T inversion	ST depression	Sokolow	Cornell	95.7
2	Rhythm	Frequency	Axis	T inversion	ST depression	Sokolow	ST elevation	95.3
3	Rhythm	Frequency	Axis	T inversion	ST depression	Sokolow	PR interval	95.3
4	Rhythm	Frequency	Axis	T inversion	ST depression	Cornell	ST elevation	95.3
5	Rhythm	Frequency	Axis	T inversion	ST depression	Cornell	PR interval	95.3
6	Rhythm	Frequency	Axis	T inversion	ST depression	Cornell	P peak interval	95.3

Note: Rhythm is short for the "one P wave precedes each QRS complex throughout the lead strip" checklist item (Table 2).

The script then denoted each ECG with an abnormal or normal tag in respect to each of the checklists above, as well as the cardiologists evaluation (SCP and remarks). When automatically assigning the cardiologists evaluation based on SCP codes as previously mentioned, 21 ECG's were not marked due to missing normal/abnormal SCP's. These were manually marked by us based on the clinical note attached in the GitHub repository (26) and looking at the ECG.

From this point we evaluated the six possible checklists above against the gold standard cardiologist evaluation. A 'true normal' denotes a consensus between the specified checklist and the cardiologist, indicating that the ECG is normal. Conversely, a 'false normal' signifies a mismatch where the specified checklist failed to identify an abnormality flagged by the cardiologist. For detailed results, please refer to Table 7.

Table 7 - Checklist Performance against SCP Codes

Checklist	True normal	False normal	False abnormal	True abnormal	Sensitivity	Specificity
2	35	29	10	234	0.890	0.778
1	33	30	12	233	0.886	0.733
3	34	30	11	233	0.886	0.756
4	33	31	12	232	0.882	0.733
5	32	32	13	232	0.878	0.711
6	32	32	13	231	0.878	0.711

Note: Checklist 1 ranks second in terms of sensitivity and specificity when compared to Checklist 2. Otherwise the results align with expectations for the other checklists.

Discussion

From analyzing all possible combinations of the checklist variables after removing QTc interval, P amplitude and P duration from the dataset due to high inter-interpreter variance and evaluating them against SCP codes highlights, the list «1) Rhythm, 2) frequency, 3) axis, 4) T inversion, 5) ST depression, 6) ST elevation and 7) Sokolow-Lyon's criteria for left ventricular hypertrophy» was slightly superior with a sensitivity of 89.0% and specificity of 77.8%. This checklist had the greatest performance with regards to the objective of our study.

Clinical Relevance

Our checklist, tailored to a minimal set of items, holds potential clinical relevance. Particularly, it offers a swift screening tool for healthcare professionals such as nurses and doctors in high-pressure hospital settings. Additionally, the standardized set of items may enhance inter-interpretion skills, illustrating the advantages of a streamlined approach in ECG analysis. The study's findings suggest practical implications for efficient and accurate ECG assessment under time constraints, emphasizing the benefits of a concise, standardized checklist in a clinical context. An appreciation should also be made that key ischemic markers as ST depression, ST elevation and T inversion are included which are appealing to us in a clinical setting, intuitively making the list more trustworthy.

Future Directions

While our study has shown promising outcomes in developing a streamlined checklist for ECG interpretation, it is crucial to emphasize the necessity of future testing on more vast and different datasets to avoid introduction of systematic or accidental bias.

Patient populations may exhibit distinct characteristics or prevalence of specific cardiac conditions, calling for further testing on different patient population to assess the external validity in addressing the local burden of disease. Conducting tests in different settings, such as primary care clinics or specialized cardiac units on certain populations, will allow also allow evaluation of the negative and positive predictive value in different populations.

Moreover, the method could be seen as a framework for further improvement where tests in different populations may lead to tailoring the checklist to better suit the specific needs and challenges of different healthcare settings. This, though creates more variability which in turn makes it difficult to switch from different settings and workplaces.

Doctors or Machines?

There are philosophical considerations when it comes to ECG interpretation that need to be discussed as well. Adopting a highly categorical, or even diatomic, interpretation of ECGs may lead to over-diagnosing patients, particularly when interpreting without the context of clinical information and the patient's medical history. ECG findings should be viewed in light of the patient's overall health status. For example, when observing a left axis deviation in a routine clinical setting, we may acknowledge it to demonstrate attentiveness, but also emphasize that the finding is unlikely to have clinical significance in a young athlete. This distinction does not come to show when interpreting only using a very numeric and categorical checklist such as the ones we use.

It's crucial to recognize that labeling an ECG as "abnormal" can provoke fear in patients and create uncertainty for other physicians when evaluating ones interpretations. Therefore, expanding the definition of a "normal ECG" is often wisely to encompass the span of normality in a population. ECG interpreters should not merely function as computers but also as medical professionals, considering the patient's well-being and maintaining a holistic approach to the ECG. The goal of any clinician should be to provide a comprehensive interpretation that goes beyond a rigid diagnostic framework, taking into account the broader clinical context and the potential impact on patients and healthcare providers.

Limitations

As we relied heavily on the AHA studies (1-6), we should point out that it was intended for for automatic machine interpretation of ECGs. Nevertheless, the criteria and physiology are the same for manual interpretation.

During the data export phase, we significantly narrowed down the chosen ECGs to ensure a manageable workload for the entire project as the analysis is very time consuming. This involved selecting only four ECGs from each SCP, with a few exclusions due to duplications. Some diagnosis are recognized by multiple SCP, and will thus be well represented in the dataset. One such example is myocardial ischemia, which have SCP on each anatomical position (eg. lateral, septal, inferior ischemia) and severity (unspecific, ST elevation, ST depression). On the contrary, other phenomenon are only represented by one SCP (eg. trigeminal pattern). This might skew our results, and will not represent the natural occurrence of each diagnosis in the population.

Importantly, we should note that we only included data from one database. Consequently, there could also be inherent biases in the SCP or clinical remarks that we do not have the possibility to evaluate.

We also acknowledge the potential for confirmation bias in our approach, as our knowledge of absolute diagnostic values may influence our interpretations to lean towards one value as we know that will be significant for a pathological finding. Furthermore, the variability between ECG interpreters that was significant in several variables that were excluded is interesting for further study, especially in newly trained personell and their ability to get consistent results.

Utilizing scripts to explore all potential combinations and calculate essential values is crucial in this context with huge datasets and many variables. However, it's important to acknowledge that these scripts can introduce significant biases and flaws to the results. Errors in algorithms or logical operations may impact the data in ways that are not immediately apparent, particularly considering the vast amount of data involved. To address this challenge, we developed multiple test files to assess the integrity of different parts of the script and the script as a whole. The script follows the same logical approach as the test script but automatically accommodate to the number of variables and number of ECGs based on the file inputted. Additionally, we conducted manual calculations on several random samples from the output to validate whether a specific checklist point should yield the indicated result or not.

To ensure transparency we have attached the GitHub repository for all python script, as well as the checklists and exported ECG’s for those who want to scrutinize the logic. Unfortunately, we have not adjusted the script to proper object-oriented programming standard as this would have taken more time then we had on our hands.

Conclusion

The study successfully developed a checklist for interpretation, demonstrating a 95.7% accuracy in identifying abnormal electrocardiograms. Additionally, the checklist exhibited an sensitivity of 89.0% when compared to SCP diagnostic coding and clinical remarks. The checklist of rhythm (one p before each QRS), frequency, axis, T inversion, and ST depression are significantly shorter than other lists commonly used in clinical settings, which might reduce screening time and time needed to learn to interpret ECG's.

As mentioned, there are several potential pitfalls in our study. This includes potential biases from ECG theory sources, ECG selection and export, and interpretation variances. Emphasizing the importance of a holistic approach to ECG interpretation, acknowledging that a numerical and categorical checklist cannot fully replace professional judgment and clinical context, but could be useful for in screening for further evaluation.

In summary, the study highlights the effectiveness of the checklist in detecting abnormal ECG's while stressing the importance of contextual medical assessment. It demonstrates that ECG interpretation involves not just pattern recognition but also an understanding of their clinical significance, affirming the need for a comprehensive, context-aware approach in medical practice.

Further, prospective studies using the checklist in the correct clinical context should be done in order to verify it's sensitivity and specificity, and evaluate positive- and negative predictive values in that population.

Acknowledgments

We wish to extend acknowledgment to Christian Tronsdad and Nils Kristian Skjærvold with their essential input to cracking our main obstacle of exporting the ECG databases to graphical representation using the WDFB python package.

Deep diving into the challenges of ECG theory Knut Gjesdal have been an incredible resourceful guide navigating the jungle of diagnostic criteria and the balancing act between diatomic criteria and clinical context. His online course for the university of Oslo have provided us with deep and thorough understanding, unparalleled to any other courses we have taken in ECG interpretation. Hence, it could be highly recommended for anyone interested in ECG interpretation as an open source knowledge.

Lastly, a huge appreciation and honorary mention must be given to our study supervisor Jonny Hisdal. He has supported our project and been invaluable in the formulation and fine tuning of the clinical question. Further, he has provided us with the expertise we needed when we got stuck on problems underway. Lastly, we wish to recognize and value the flexibility Jonny has given us to work on the project as it has suited us next to our studies and projects, being available as we have needed.

Bibliography

1.
Kligfield P, Gettes LS, Bailey JJ, Childers R, Deal BJ, Hancock EW, et al. HA/ACCF/HRS recommendations for the Standardization and Interpretation of the Electrocardiogram part 1| Circulation. Circulation. 2007 Mar 13;115(10):1306–24.

2.
Mason JW, Hancock EW, Gettes LS. HA/ACCF/HRS recommendations for the Standardization and Interpretation of the Electrocardiogram part 2| Circulation. Circulation. 2007 Mar 13;115(10):1325–32.

3.
Surawicz B, Childers R, Deal BJ, Gettes LS. HA/ACCF/HRS recommendations for the Standardization and Interpretation of the Electrocardiogram part 3| Circulation. Circulation [Internet]. 2009 Mar 17 [cited 2024 Jan 18];119(10). Available from: https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.108.191095

4.
Rautaharju PM, Surawicz B, Gettes LS. HA/ACCF/HRS recommendations for the Standardization and Interpretation of the Electrocardiogram part 4| Circulation. Circulation [Internet]. 2009 Mar 17 [cited 2024 Jan 18];119(10). Available from: https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.108.191096

5.
Hancock EW, Deal BJ, Mirvis DM, Okin P, Kligfield P, Gettes LS. HA/ACCF/HRS recommendations for the Standardization and Interpretation of the Electrocardiogram part 5| Circulation. Circulation. 2009 Mar 17;119(10):e251–61.

6.
Wagner GS, Macfarlane P, Wellens H, Josephson M, Gorgels A, Mirvis DM, et al. HA/ACCF/HRS recommendations for the Standardization and Interpretation of the Electrocardiogram part 6 | Circulation. Circulation [Internet]. 2009 Mar 17 [cited 2024 Jan 18];119(10). Available from: https://www.ahajournals.org/doi/10.1161/CIRCULATIONAHA.108.191098

7.
Gjesdal K, Stokke MK, Hurdal MB. Elæring fra med.fak. UiO [Internet]. [cited 2024 Jan 18]. Available from: https://studmed.uio.no/elaring/fag/hjertesykdommer/index.shtml

8.
ECG tutorial: Basic principles of ECG analysis | UpToDate [Internet]. [cited 2023 Oct 11]. Available from: https://www.uptodate.com/contents/ecg-tutorial-basic-principles-of-ecg-analysis?search=ECG%20tutorial:%20Basic%20principles%20of%20ECG%20analysis&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1

9.
ECG tutorial: ST and T wave changes | UpToDate [Internet]. [cited 2024 Jan 17]. Available from: https://www.uptodate.com/contents/ecg-tutorial-st-and-t-wave-changes?search=ECG%20tutorial:%20Basic%20principles%20of%20ECG%20analysis&topicRef=2115&source=see_link

10.
ECG tutorial: Myocardial ischemia and infarction - UpToDate [Internet]. [cited 2024 Jan 18]. Available from: https://www.uptodate.com/contents/ecg-tutorial-myocardial-ischemia-and-infarction

11.
Spodick DH. Normal sinus heart rate: sinus tachycardia and sinus bradycardia redefined. Am Heart J. 1992 Oct;124(4):1119–21.

12.
Chen LY, Ribeiro ALP, Platonov PG, Cygankiewicz I, Soliman EZ, Gorenek B, et al. P Wave Parameters and Indices: A Critical Appraisal of Clinical Utility, Challenges, and Future Research—A Consensus Document Endorsed by the International Society of Electrocardiology and the International Society for Holter and Noninvasive Electrocardiology. Circulation: Arrhythmia and Electrophysiology. 2022 Apr;15(4):e010435.

13.
Magnani JW, Wang N, Nelson KP, Connelly S, Deo R, Rodondi N, et al. Electrocardiographic PR interval and adverse outcomes in older adults: the Health, Aging, and Body Composition study. Circ Arrhythm Electrophysiol. 2013 Feb;6(1):84–90.

14.
Forselv GC, Vik-Mo H. EKG-forandringer ved akutt koronarsyndrom uten ST-elevasjon. Tidsskrift for Den norske legeforening [Internet]. 2007 Sep 6 [cited 2024 Jan 17]; Available from: https://tidsskriftet.no/2007/09/originalartikkel/ekg-forandringer-ved-akutt-koronarsyndrom-uten-st-elevasjon

15.
Kihlgren M, Almqvist C, Amankhani F, Jonasson L, Norman C, Perez M, et al. The U-wave: A remaining enigma of the electrocardiogram. Journal of Electrocardiology. 2023 Jul 1;79:13–20.

16.
Vandenberk B, Vandael E, Robyns T, Vandenberghe J, Garweg C, Foulon V, et al. Which QT Correction Formulae to Use for QT Monitoring?

17.
Bogossian H, Frommeyer G, Ninios I, Pechlivanidou E, Hasan F, Nguyen QS, et al. A new experimentally validated formula to calculate the QT interval in the presence of left bundle branch block holds true in the clinical setting. Ann Noninvasive Electrocardiol. 2016 Aug 26;22(2):e12393.

18.
ECG Axis Interpretation • LITFL • ECG Library Basics [Internet]. [cited 2024 Jan 18]. Available from: https://litfl.com/ecg-axis-interpretation/

19.
Left ventricular hypertrophy: Clinical findings and ECG diagnosis - UpToDate [Internet]. [cited 2024 Jan 18]. Available from: https://www.uptodate.com/contents/left-ventricular-hypertrophy-clinical-findings-and-ecg-diagnosis?search=leftventricularhypertrophy&source=search_result&selectedTitle=1~150&usage_type=default&display_rank=1#H12

20.
PhysioNet Databases [Internet]. [cited 2024 Jan 18]. Available from: https://www.physionet.org/about/database/

21.
Zheng J, Guo H, Chu H. A large scale 12-lead electrocardiogram database for arrhythmia study [Internet]. PhysioNet; [cited 2024 Jan 18]. Available from: https://physionet.org/content/ecg-arrhythmia/1.0.0/

22.
Wagner P, Strodthoff N, Bousseljot RD, Samek W, Schaeffter T. PTB-XL, a large publicly available electrocardiography dataset [Internet]. PhysioNet; [cited 2024 Jan 18]. Available from: https://physionet.org/content/ptb-xl/1.0.3/

23.
Rubel P, Fayn J, Macfarlane PW, Pani D, Schlögl A, Värri A. The History and Challenges of SCP-ECG: The Standard Communication Protocol for Computer-Assisted Electrocardiography. Hearts. 2021 Sep;2(3):384–409.

24.
LightWAVE. ECG graphical represnation software on physionet [Internet]. [cited 2024 Jan 18]. Available from: https://www.physionet.org/lightwave/?db=ptb-xl/1.0.3

25.
TagFri. GitHub: TagFri/wfdb-ecg-export [Internet]. 2023 [cited 2024 Jan 18]. Available from: https://github.com/TagFri/wfdb-ecg-export

26.
TagFri. GitHub: TagFri/ecgChecklistAnalysis [Internet]. 2024 [cited 2024 Jan 18]. Available from: https://github.com/TagFri/ecgChecklistAnalysis

27.
Fredholm K, Højlund M, Bjerregaard MR, Nielsen JC, Fink-Jensen A, Correll CU. How to Estimate QT Interval in Patients With Left or Right Bundle Branch Block: A Systematic Review. J Clin Psychopharmacol. 2021 Jun 1;41(3):323–6.

Arbeidsfordeling - prosjektoppgave

Vetle Mørland og Åge Frivoll har samarbeidet om en prosjektoppgave om EKG tolkning. Prosjektoppgaven innebærer et omfattende litteraturgjennomgang for å utarbeide en sjekkliste til bruk i EKG tolkning, Python-drevet eksport av EKG'er, tidkrevende manuell tolkning av hvert enkelt EKG og Python-drevet dataanalyse mm. Vi hadde ideen til prosjektoppgaven selv og har i stor grad utarbeidet og fasilitet studien på egenhånd.

I begynnelsen brukte vi en del tid på å planlegge metoden og kartlegge studiedesignet sammen for å unngå bias og gi den mest pragmatiske tilnærmingen til oppgaven. Begge var sterkt involvert i arbeidet, og diskuterte prosessen gjennom flere utkast før selve prosjektoppgavearbeidet kunne starte.

I litteraturgjennomgangen brukte vi begge mye ressurser og tid på å gjennomgå ulike kilders kriterier for å tolke abnormale funn. Vi diskuterte ofte utfordrende kriterier der forskningskonsensus var mangelfullt. Der vi ikke ble enige kontaktet vi eksterne fagressurser. Arbeidet resulterte i en gjennomgang og vurdering av 22 sjekklistepunkter vurdert av mangfoldige kilder. Vi skrev også introduksjonsdelen parallelt med denne fasen. Vetle fikk i oppgave å skrive mesteparten av introduksjonsdelen og innledende metodedel om sjekklistepunktenes teoretiske grunnlag og inklusjonsarguementer for studiens formål når vi var enige. Dette var nødvendig siden Åge samtidig jobbet med script grunnet hans ferdigheter innen python programmering.

For å kunne tolke EKG'er så måtte vi finne en digital database som muliggjorde eksport. Åge identifiserte databasene som var tilgjengelig ut ifra eksport mulighetene og diskuterte med Vetle hvilke som er mest aktuelle for studien. Åge gikk så videre med å scripte et eksportering og seleksjonsprogram i Python som utnyttet WFDB biblioteket og lagret utvalgte EKG'er som PDF-er med riktig format. Det var betydelig sparring frem og tilbake mellom Vetle og Åge underveis for å finne den optimale måten å selektere EKG'er, samt mest optimale fremstillingen for tolkning.

Deretter fulgte den mest tidkrevende fasen der vi skulle tolke alle de 308 EKG-ene individuelt. Tolkningsprosessen bestod av å manuelt måle og registrere 19 datapunkter (sjekklistepunkter) på hvert EKG og deretter gjøre en vurdering av om EKGet er normal eller unormalt. All data ble registrert i hvert sitt excel-ark. Tidsbruk for hvert EKG er anslått å være omtrent 10 min, som tilsvarer 51 timer hver. Deretter samtolket vi resultatene for å forhindre tolkningsbias og luke ut eventuelle kunnskapshull. Denne fasen tok ytterligere 20 timer. Dataene fra denne prosessen ble viktige for å finne variansen i tolkningene.

Når vi hadde tolket EKG'ene hver for oss og sammenstilt våre resultater, lagde Åge et analysescript i Python for å kunne utnytte datakraften vi trengte for kombinatorikk da det er utrolig mange kombinasjoner av de 19 sjekklistepunktene som er mulig. Her ble det brukt mye tid på scriptet og resultatene ble diskutert og kontrollert sammen for å sikre at vi ikke har introdusert bias i beregningene som ikke ble fanget opp innledningsvis.

Resultatene ble deretter gjennomgått sammen og nådde en konklusjon. Vi brukte god tid til å diskutere mulige feilkilder og bias og har prøvd å reflektere dette i diskusjonen for å være mest mulig transparente.

Totalt har prosjektet vært et svært spennende, tidkrevende og et lærerikt samarbeid der vi har begge vært svært delaktige i alle prosjektets faser og lært mye av hverandres ferdighetsområder.