

# Study of connected health applications and implementation of a solution for personalized and precision medicine

Ci TANG<sup>1,a</sup>

Supervised by Mr Massinissa HAMIDI<sup>2,b</sup>

<sup>1</sup>student M1CNS, Informatique et Science Numérique Graduate School, University Paris-Saclay

<sup>2</sup>lecturer, IBISC Laboratory, University Paris-Saclay

<sup>a</sup>ci.tang@etud.univ-evry.fr

<sup>b</sup>massinissa.hamidi@univ-evry.fr

May 15, 2024

## Abstract

This project aims to propose a solution based on connected health applications to address the challenges faced by personalized and precision medicine in implementation, particularly concerning the issues of delayed diagnosis and misdiagnosis, especially for delayed diagnosis. The resolution of this problem holds significant potential implications for public health and healthcare system efficiency, as misdiagnosis is a important contributor to iatrogenic harm, while delayed diagnosis poses a high risk of amplifying the health consequences of diseases substantially. The proposed solution has the potential to mitigate unnecessary adverse medical outcomes and protect lives by facilitating timely medical intervention for patients with significant diseases and reducing the occurrence of misdiagnoses. Moreover, the enhancement of accuracy and timeliness in medical diagnostics will significantly contribute to the more efficient operation of the healthcare system, thereby reducing the burden on healthcare professionals. In this project, the design of the solution involves continuous health monitoring in daily life using wearable devices and rapid responses to potential health issues.

Index Terms: Health monitoring, wearable devices, personalized and precision medicine, large language model

## Contents

<b>I. Introduction</b>	<b>1</b>
1. The goal of the TER . . . . .	2
<b>II. State-of-the-art</b>	<b>3</b>

1. Commercial products in continuous monitoring of health in daily life . . . . .	4
a. Implantable devices . . . . .	4
b. Wearable devices . . . . .	4
c. Non-contact smart solutions . . . . .	5
2. Studies in continuous monitoring of health in daily life . . . . .	7

<b>III. Proposed solution</b>	<b>7</b>
-------------------------------	----------

<b>IV. Evolution</b>	<b>10</b>
1. Simulated diagnosis example . . . . .	10
2. LLM suitability assessment . . . . .	11

<b>V. conclusion</b>	<b>12</b>
----------------------	-----------

<b>VI. Future perspectives</b>	<b>12</b>
--------------------------------	-----------

## I. Introduction

Today's physicians make efforts for precision medicine to reduce iatrogenic harm and implicit health risks among patients, with the misdiagnosis and delayed diagnosis being one of the primary challenges. A priori, misdiagnosis may entail severe medical consequences, however, historically, according to a report from the National Academy of Medicine of USA in 2015[9], few attempts have been made to quantify deaths and permanent disabilities resulting from misdiagnosis. Due to the lack of statistical analysis in this regard, the full extent of the consequences of misdiagnosis in healthcare systems worldwide remains unknown, and the medical community has underestimated the overall impact of misdiagnosis. Yet, thanks to a series of related studies published by Dr. Newman-Toker's team from Johns Hopkins

University in 2019[6], 2021[8] and 2023[7] respectively, we now partially understand the deleterious medical effects of misdiagnosis in the healthcare system of the USA.

According to his research, it is estimated that in the USA, at least 371,000[7] people die annually due to misdiagnosis, 424,000[7] people suffer permanent disabilities for the same reason, with a total of 795,000 people experiencing serious harm. Determining a specific number of all misdiagnosis incidents is challenging, as many cases of misdiagnosis go undetected, a estimated number is 31 million[8] per year, while actual numbers may be much higher—about 100-200 million[8], for example—or possibly lower. Though not all misdiagnoses necessarily result in serious consequences, misdiagnosis remains one of the most common cause of death or disability resulting from medical accidents. Among the misdiagnoses causing the most harm, 74.1%[6] of errors are attributed solely to three categories—the “big three”, as termed in his paper: cancer (37.8%[6]), vascular events (22.8%[6]), and infections (13.5%[6]). Within the “big three”, the top five diseases in each category (a total of fifteen) account for 47.1%[6] of all high-risk misdiagnoses and 63.5%[6] of all high-risk misdiagnoses within the “big three”. Specifically, strokes are the most misdiagnosed vascular disease, sepsis is the most misdiagnosed infection, and lung cancer is the most misdiagnosed cancer.

The diseases causing the most harm in misdiagnosis also present another major challenge of precision medicine—the importance of avoiding delayed diagnosis. Delayed diagnosis is one of the consequences of misdiagnosis but may not occur sequentially. Delayed diagnosis significantly exacerbates the consequences of major diseases; even if not fatal, delayed diagnosis can lead to permanent and serious consequences. For example, delayed diagnosis of breast cancer results in more patients needing mastectomies[14]. For acute conditions that are not easily detected, delayed diagnosis is highly correlated with patient survival rates. For instance, in blunt hollow viscus injury occurring after blunt trauma, a five-hour delay in diagnosis reduces the chance of survival to a very low level[5].

It is worth noting that delayed diagnosis differs from misdiagnosis; the latter is solely related to physicians, while the former is also related to patients. The delay caused by the patient’s own actions depends on when they realize their health problems and how much they are aware of the seriousness of their health problems—which affects whether and when they seek medical attention.

Traditionally, the detection of diseases highly

depends on patients’ awareness and concern for their own health, especially for latent diseases, which rely on regular or irregular physical examinations. However, not everyone undergoes regular examinations, especially in underdeveloped regions and low-income groups[16, 2], leading to significant delays in diagnosis before seeking medical treatment. Moreover, the lack of past health data for patients poses greater challenges to avoiding misdiagnosis. Continuous and long-term monitoring of patients’ health conditions (both in daily life and in intensive care units) through a solution based on connected applications/objects could provide beneficial support for precision medicine. Immediate detection and attention to suspicious health risks in patients and the collection of long-term health data can facilitate early intervention in disease occurrence and risk prediction and post-diagnosis.

Currently, continuous monitoring of patients in intensive care units has made significant progress, but continuous monitoring of patients in daily life are still in the early stages of development. Although many devices and schemes have been proposed, they have not been widely adopted among the general population. **This work will focus on discussing the following challenges in the regard of helping continuous monitoring of patients in daily life and analyzing health risks through computer science methods to provide personalized and precision healthcare:**

- How to achieve system scalability to make it easier to incorporate new patterns and functionalities?
- How to analyze data using basic machine learning models to extract signs of health conditions and predict potential risks?
- How to pay special attention to specific aspects of diseases or pathologies such as Myocardial Infarction (MI)?

A state-of-the-art review of existing research and commercial tools will be detailed in Section 1. Section 2 will be dedicated to presenting my design and describing its implementation. Section 3 covers my conclusion of this work. Finally, in Section 4 I offer insight into potential future research endeavors.

## 1. The goal of the TER

The goal of this Travail d’étude et de recherche (TER), is to propose a Internet of thing (IoT) solution using wearable devices with help of Large

Language Model (LLM)s to get medical advice quickly and in a timely manner.

In this work, particularly, I focus on find out a way to realize rapide health data analyze using LLMs, and to describe that solution’s conception base on such way.

## II. State-of-the-art

This section provides an overview of all research efforts and commercial products in the field of continuous monitoring of patient health in daily life using connected objects. Various solutions are based on the utilization of different measures.

The first category of measures involves the use of implantable devices, which has a long-standing history in human medical, yet, implantable devices for health monitoring are still in their infancy stage of development. Both their shortcomings and advantages are evident. Once installed, it works continuously for long periods of time without disturbing the user at all, it doesn’t cause any degree of friction for users, doesn’t need to be remembered to be worn every day, doesn’t need to be removed when swimming or showering, doesn’t require self-charging, and doesn’t require consideration of whether they match one’s attire, in one sentence, it eliminates the need to invest more effort or attention into it in daily life. More importantly, implantable sensors within the body hold the potential to ensure continuous monitoring of health status with high precision, in tracking certain health parameters, such as cardiac and circulatory functions, implantable sensors is more accurate compared to those deployed externally on the skin. Moreover, implantable devices offer the capability to monitor health factors that are difficult to observe with devices placed on the body surface, for instance, for continuous long-term monitoring of certain features like gastrointestinal status, the use of implantable devices might be the sole option available. However, the disadvantages of such measures include the toxicity, potential radiation, bodily rejection, and possible inflammation associated with human body invasiveness. Additionally, implantable devices usually entail high costs and require specialized surgery for implantation or removal.

The second category of measures involves the use of wearable devices, such as smartwatches, fitness bands(the main difference being that fitness bands generally have fewer features and cannot install applications), and smart patches. Wearable devices completely overcome the drawbacks of implantable devices as they are non-invasive and generally do not cause adverse reactions. They can be

conveniently started or stopped. Although they may sometimes have inferior accuracy compared to that of implantable devices, typically require users to charge the devices by users themselves, usually need to be removed when users take hot showers to prevent steam from entering the device, and may not be suitable for monitoring certain specific characteristics (such as gastrointestinal status mentioned earlier). Yet, the current development level of modern wearable devices allows for continuous long-term monitoring of some basic health features, such as heart rate, blood oxygen, and sleep. Some smartwatches can collect electrocardiograms, while some advanced smart patches can monitor complex diseases such as Parkinson’s disease, epilepsy, and heart failure, as I will detail it later.

The final category of measures involves the use of non-contact smart solutions. Traditionally, non-contact health detection mainly utilizes electromagnetic radiation or mechanical waves. Typical examples include technologies that use X-rays to penetrate the human body and image different body tissues based on the penetration results of X-rays; technologies that utilize B-mode ultrasound echo imaging; and more complex technologies that use externally applied gradient magnetic fields to detect electromagnetic waves for magnetic resonance imaging. These three technologies are commonly used for anatomical diagnosis of complex mechanical injuries, cardiovascular and cerebrovascular diseases, and cancer. They are complex to operate, have large equipment volumes, high costs, require specialized examination sites, and cannot achieve continuous long-term monitoring.

However, in recent years, a series of non-contact monitoring methods have emerged using the inherent hardware features of mobile phones, such as using the microphone to monitor breathing during sleep to estimate sleep quality, and using the phone camera to scan then analyse skin diseases with AI model. These solutions are implemented with minimal resource usage, regularly or continuously invoking mobile hardware at low power using mobile applications, and are connected to machine learning models deployed in the cloud. The fact that mobile phones are universally carried by users for long periods in modern life ensures that such solutions are not much inferior to the first two categories in terms of continuous monitoring time, but they might still not always be possible to avoid significant disadvantages in terms of accuracy, especially in the regard of requiring automatic monitoring such as sleep.

# 1. Commercial products in continuous monitoring of health in daily life

## a. Implantable devices

As of present, there are few options available of implantable device that focus on continuous health monitoring purpose in the commercial market, albeit they do exist. In general, there are three main implantation methods utilized:

- Subcutaneous implantation
- Intravascular implantation
- Intracavitary implantation (including ingestion-based implantation)

This kind of device is generally developed by professional medical device companies and are marketed under the regulatory oversight of medical device categories such as presented in table 1.

Product	Manufacturer	Description
Cordella	Endotronix	A pulmonary artery pressure sensor implanted within the artery to detect severe heart failure and advance its treatment.
BeUno	DSruptive	A subcutaneously implanted temperature detection device administered via injection.
Proteus Discover	Proteus Digital Health	An indigestible device, inserted through the esophagus into the stomach, used for treating and monitoring cardiovascular diseases, infectious diseases, and other health issues. Unfortunately, its manufacturer filed for bankruptcy in 2020.
LUX-Dx ICM	Boston Scientific	A Implantable Cardiac Monitor (ICM) implanted subcutaneously in the left chest area for monitoring heart rate, particularly for detecting arrhythmias.
BIOMONITOR IV ICM	BIOTRONIK	A novel ICM with artificial intelligence capabilities designed to enhance the efficiency of arrhythmia detection.
Reveal LINQ ICM	Medtronic	A miniature ICM with self-learning capability focused on monitoring Atrial Fibrillation (AF).

Table 1: Solutions available in the form of smart patch for daily health monitoring.

It is worth mentioning that implantable devices for continuous health monitoring in daily life, particularly focusing on cardiopulmonary function, reflect the application direction of implantable devices in this regard. This involves monitoring important health characteristics that require continuous and highly accurate attention in daily life. Devices implanted within the body ensure continuous monitoring without interruption in daily activities. Furthermore, sensors implanted within the body can effectively capture electrocardiographic signals for Electrocardiogram (ECG) mapping, while sensors implanted within blood vessels

can more accurately monitor Oxygen Saturation (SATS) levels of blood.

## b. Wearable devices

Numerous products are marketed for continuous monitoring of health in daily life, mainly fall into the following three categories:

- Smartwatches/fitness bands
- Smart patches
- Others

The most common are wearable devices in the form of smartwatches or fitness bands introduced by IT industry companies. These devices are often not designed solely for health monitoring purposes and typically do not fall under the regulatory oversight of local public authorities as medical devices. However, monitoring health is one of the main functions and sources of attractiveness of such products. There is a vast array of such products, a single manufacturer may have released hundreds of different models, but I will focus here only on the newer products from leading manufacturers as which listed in Tables 2.

Product	Manufacturer	Function
Apple Watch Series 9	Apple	ECG,SATS,HR,SG,RES,STR
Apple Watch Series 8	Apple	SATS,HR,SG,RES,STR
Apple Watch ultra 2	Apple	ECG,SATS,HR,SG,RES,STR
Apple Watch ultra	Apple	SATS,HR,SG,RES,STR
Apple Watch SE	Apple	SATS,HR,SG,RES,STR
HUAWEI Band D	HUAWEI	BPECG,SATS,HR,SG,RES,STR
HUAWEI WATCH ULTIMATE DESIGN	HUAWEI	ECG,SATS,HR,SG,RES,STR
HUAWEI WATCH GT 4	HUAWEI	SATS,HR,SG,RES,STR
HUAWEI WATCH GT 3	HUAWEI	SATS,HR,SG,RES,STR
HUAWEI WATCH GT 3 Pro	HUAWEI	ECG,SATS,HR,SG,RES,STR
HUAWEI WATCH 4	HUAWEI	ECG,SATS,HR,SG,RES,STR
HUAWEI WATCH 3 Pro	HUAWEI	ECG,SATS,HR,SG,RES,STR
HUAWEI WATCH 3	HUAWEI	SATS,HR,SG,RES,STR
HUAWEI Band 8	HUAWEI	SATS,HR,SG,RES,STR
Redmi watch 4	XIAOMI	SATS,HR,SG,RES,STR
Watch Color 2	XIAOMI	SATS,HR,SG,RES,STR
XIAOMI Watch 2	XIAOMI	SATS,HR,SG,RES,STR
XIAOMI Watch S3	XIAOMI	SATS,HR,SG,RES,STR
XIAOMI band 8	XIAOMI	SATS,HR,SG,RES,STR
Sense 2	Fitbit	ECG,SATS,HR,SG,RES,STR
Google Pixel Watch 2	Fitbit	ECG,SATS,HR,SG,RES,STR
Google Pixel Watch	Fitbit	SATS,HR,SG,RES,STR
Versa 4 (watch)	Fitbit	SATS,HR,SG,RES,STR
Versa 3 (watch)	Fitbit	SATS,HR,SG,RES,STR
Galaxy Watch Series 6	Samsung	ECG,SATS,HR,SG,RES,STR
Galaxy Watch Series 5	Samsung	ECG,SATS,HR,SG,RES,STR
Galaxy Watch Series 5 pro	Samsung	ECG,SATS,HR,SG,RES,STR
Galaxy Active 2 (watch)	Samsung	ECG,SATS,HR,SG,RES,STR
Venu 3 (watch)	Garmin	SATS,HR,SG,RES,STR
Forerunner 265 (watch)	Garmin	SATS,HR,SG,RES,STR
Forerunner 165 (watch)	Garmin	SATS,HR,SG,RES,STR
Instinct Crossover (watch)	Garmin	SATS,HR,SG,RES,STR
Instinct 2 (watch)	Garmin	SATS,HR,SG,RES,STR
epix (watch)	Garmin	SATS,HR,SG,RES,STR
fenix 7 (watch)	Garmin	SATS,HR,SG,RES,STR
vivoactive 5 (watch)	Garmin	SATS,HR,SG,RES,STR

Table 2: Marketed wearable devices with health monitoring capabilities in daily life in form of smartwatches/fitness.

These products commonly feature sensors based on Photoplethysmography (PPG) technology[4], which shines light into the skin and measures the amount of light reflected by blood flow changes to detect Heart Rate (HR) and SATS by collecting the amount of blood flow through the dermis and subcutaneous tissue. PPG sensor is the core of the health monitoring function of these products, real-time monitoring of HR and SATS levels can provide crucial foundational health data, continuous monitoring of heart rate yields data on Heart rate variability (HRV), a paramount parameter reflecting overall health status, which has been also demonstrated to be significantly associated with mortality rates related to certain cardiovascular diseases such as sudden cardiac death[11], hypertension[10]. In parallel, based on HR and SATS, algorithms can infer various health data of users, such as vascular health, Somnography (SG), Respiration (RES), and Stressor (SRT) conditions. For example, Huawei’s Trusleep sleep monitoring solution is based on a Cardiopulmonary Coupling(CPC) technique proposed by Harvard Medical School in 2005[13]. Some high-end models also have more sophisticated ECG analysis capabilities. For example, the HUAWEI WATCH ULTIMATE DESIGN collects electrocardiographic signals through electrodes on the side of the watch, and then analyzes the electrocardiogram situation and draws the electrocardiogram through the built-in ECG analysis software. Due to the lack of publicly available information, I cannot ascertain the technical details of how Huawei implements this function but only temporarily refer to it as the Huawei Electrocardiogram Acquisition Solution.

Another notable product is the HUAWEI WATCH D, which was released in December 2021. It integrates professional Blood Pressure (BP) monitoring components, including miniature pumps and inflatable cushions. The wristband of the watch itself serves as the cuff for the blood pressure monitor. It also features PPG sensors, Huawei electrocardiogram acquisition solution, and various other software-level health analysis solutions. It stands as a cutting-edge product among commercially available wearable health monitoring devices.

The combination of PPG technology and medical algorithms enables effective continuous monitoring of basic health data in daily life and can provide targeted preliminary analysis of overall health status, indicating specific disease risks. This will help both doctors and patients to obtain long-term and timely understanding of the latter’s health status, thus providing precise, personalized medi-

cal care, and aiding in reducing delayed diagnosis caused by patients themselves. In addition, these products can generally connect to smartphones then transmit data, it means that they can not only update their software firmware through the latter, but also conveniently and quickly display the monitoring status and the results obtained from software analysis on smart terminals, especially on patients’ smartphones.

Here, I will also focus on an emerging wearable device, the Smart patch. These products represent an intelligent development of traditional transdermal patches used in the pharmaceutical field. They fix a small embedded machine integrated with multiple sensors onto the skin using adhesive patches. Compared to smartwatches or fitness bands, they offer advantages such as smaller size, placement on any part of the body, and closer sensor-body integration. Smart patches can be used with or without medication. Additionally, smart patches are medical devices developed specifically for medical purposes. Unlike smartwatches or fitness bands, they do not need to balance other aspects of life, such as phone calls. Therefore, despite their smaller size, they do not have a significant disadvantage in terms of intelligence. Moreover, they often provide sensors that are more specialized for particular domains and offer more precise health monitoring solutions tailored to specific needs of patients. In fact, smart patches can be regarded as non-invasive optimization solutions for certain areas where implantable devices are utilized. Similar to the latter, smart patches that designed to accomplish specific functions constitute a part of one kind of instances of solutions sought in this project’s topic. While many solutions are still in the experimental and paper stage, some manufacturers have already been providing solutions available for daily health monitoring, some of which are quite inspirational, as illustrated in table 3.

Besides, there are also wearable devices available in other forms for health monitoring, each with its own characteristics and unique features, some manufacturers offer these products, as presented in Table 4.

### c. Non-contact smart solutions

The implementation of non-contact schemes relies on non-contact sensors. Under daily conditions, the selection of such sensors is limited but still available, the primary components are the microphone and camera of a smartphone. They may not inherently detect specific health data directly, but they can be utilized to collect data useful for analyzing certain health aspects. Sub-

Product	Manufacturer	Description
BioButton	biointellisense	Providing healthcare professionals with comprehensive and professional monitoring of patients, coupled with notifications based on data analysis algorithms, to assist in clinical decision-making.
Zio	iRhythm	An easy-to-use professional cardiac monitoring device that is to be worn by the patient for two weeks before being returned to healthcare professionals.
MCOT	Biotelemetry	An easy-to-use networked professional cardiac monitoring device capable of real-time data transmission to healthcare professionals, with each patch lasting up to three days.
Vivalink ECG	Vivalink	An easy-to-use networked professional cardiac monitoring device capable of real-time data transmission to the cloud, reusable, and rechargeable.
GUTTRACKER	G-TECH	Measuring gastrointestinal peristalsis, capable of transmitting data to the cloud and quantifying digestive organ motility using analysis algorithms.
Diabetes patch	Gentag	A Blood Glucose (BG) measurement platform connected to a mobile phone.
Fever patch	Gentag	A low-cost disposable temperature monitor that continuously transmits Body Temperature (BT) data of patients without disturbing them.

Table 3: Solutions available in the form of smart patch for daily health monitoring.

Product	Description
Intelligent Insoles	Product of Xsensor, gait and kinematic analysis uses.
Smart gloves	Studies exist for patients with rheumatoid arthritis remotely monitoring use[12].
Smart hat/helmet	Head-mounted health monitors, originally designed for pilots, LifeBEAM provides such products.
Smart armband	A version of the Smart Patch that uses an elastic armband instead of a form-fitting patch, ivalnk offers SATS monitors in this form.

Table 4: Other marketed health monitoring purpose devices.

sequently, health data can be indirectly obtained through analysis by trained baseline models. This approach has gained popularity in recent years, exemplified by software and algorithms that utilize smartphones without the need for any wearable or implantable devices to record and analyze sleep patterns. Table 5 presents practical examples of such solutions.

Generally, sleep monitoring solutions in the form of mobile apps often offer a free version with basic features and a paid subscription ver-

Solution	Description
HUAWEI TruSleep	Huawei’s sleep monitoring solution, typically used in conjunction with Huawei’s wearable devices (bands and watches), but it can also work solely through the smartphone placed near the user without any wearable devices by using the microphone. In this scenario, it can collect data on sleep cycles, continuous heart rate during sleep, sleep respiration rate, snoring, and sleep talking.
OPPO Sense Sleep	The OPPO version of HUAWEI TruSleep, while functionally similar, lacks clarity regarding whether the algorithms are the same.
Aysa	A mobile app that deploys an AI model capable of recognizing signs of skin diseases, allowing users to scan skin features using the phone camera and return dermatological recommendations.
Do I Snore or Grind	A mobile app that monitors whether users snore or grind their teeth during sleep using the microphone.
PrimeNap	A mobile app using the microphone that features detailed charts, sleep analysis graphs, and additional functions such as dream diaries. Yet, the underlying principles of its operation are not clearly elucidated.
SnoreLab	A mobile app using the microphone, it focuses solely on recording and analyzing snoring sounds, providing comprehensive and detailed information on snoring through its algorithm.

Table 5: Non-contact daily health monitoring solutions.

sion with all features included. The solutions from Huawei and OPPO are integrated into a native health-fitness tracking apps included in their respective brand’s smartphones. Solutions like Aysa are a form of AI smart lens, similar to Google AI Lens, that utilize machine vision and underlying models to recognize input images from the smartphone camera (or storage space). Modern smartphones are already capable of providing high-resolution clear images, making such solutions promising for applications in personalized and precision medicine.

## 2. Studies in continuous monitoring of health in daily life

In recent years, a lot of research concerning health monitoring based on connected applications/objects has been published, encompassing two major categories: implantable devices and wearable devices. Utilizing the Mendeley database and employing keywords such as "Health Monitoring" along with "implantable" or "wearable", the literature retrieved from 2015 to 2023 reflects the degree of attention garnered by both implantable devices and wearable devices in recent years. From the quantity of literature, an obvious fact emerges: wearable devices have been studied far more extensively than implantable devices, as illustrated in Table 6.

Year	implantable devices	wearable devices
2023	117	1604
2022	155	2099
2021	143	1784
2020	124	1474
2019	130	1212
2018	116	1060
2017	97	790
2016	80	658
2015	113	530
total	1075	11211

Table 6: The number of studies concerning implantable devices and wearable devices.

Regarding wearable devices, utilizing the Mendeley database and employing keywords such as "Health Monitoring" along with "watch band bracelet" or "smart patch," another obvious fact emerges: in the time of the Covid-19 pandemic 2019-2021, although smartwatches (including bands or bracelets) are widely distributed as consumer-grade personal electronic products, contrary to their market prevalence, smart patches have received greater attention in recent research endeavors, as depicted in Table 7. This may indicate the substantial potential of smart patches as flexible wearable devices in facilitating precise and personalized healthcare diagnostics, especially in response to diseases that focus on organ system and large-scale infectious diseases like Covid-19.

## III. Proposed solution

I propose a solution for personalized precision medicine based on connected devices: utilizing smart wearable devices to collect users' health data and transmitting the data to a LLM deployed

Year	smartwatch	smart patch
2023	52	23
2022	63	46
2021	61	2224
2020	54	39
2019	39	953
2018	37	33
2017	35	22
2016	27	19
2015	23	16
total	391	3375

Table 7: The number of studies concerning smartwatch and smart patch.

in the cloud, which then evaluates the patients' health data and outputs structured diagnostic results along with health and medical advice. The entry point for this functionality is integrated into the software of smartwatches, running passively in users' daily lives and also capable of being activated with a single click. The proposed solution consists of three main components as illustrated in Figure 1:

- Wearable device ensemble with health monitoring capabilities: These include a general-purpose smart platforms and some specific-purpose smart patch equipped with sensors such as ECG detection sensors, PPG sensors, and running programs for these sensors.
- Solution software: This software can be installed on general-purpose smart devices and communicates with the cloud via interfaces to transmit locally collected sensor data to the cloud and display diagnostic results and recommendations returned from the cloud. It should also include algorithms for automatically initiating health data collection and cloud communication to enable passive monitoring.
- Cloud-based services: These services provide communication interfaces, pre-trained LLM for diagnosis, databases storing vector representations of relevant professional domain knowledge, and specific pre-trained model for health data's feature engineering if needed, among others.

It was inspired by the research of Peikun Guo and Yu Han[17], who proposed a new Zero-Shot Retrieval-Augmented Generation (RAG) method, enabling large language models to diagnose diseases such as arrhythmia and sleep apnea by understanding and generating electrocardiogram

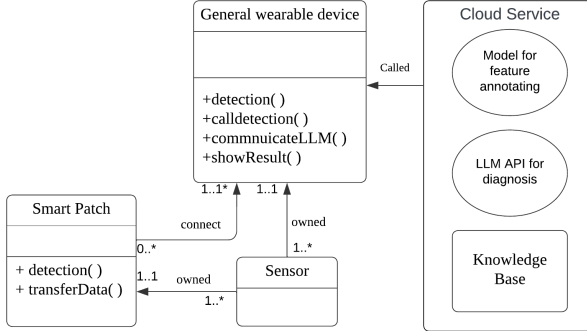


Figure 1: Solution structure

data without the need for any training data. My approach differs a bit from the one proposed by Guo and Han, which I will discuss later, my main approaches is illustrated as:

a) Constructing a professional knowledge base related to electrocardiogram diagnosis, as shown in Figure 2: Extracting professional knowledge from relevant books and papers and encoding it into vector form for storage. Thanks to Chroma technology, these vector representations can be used to capture semantics, enabling similarity-based text searches.

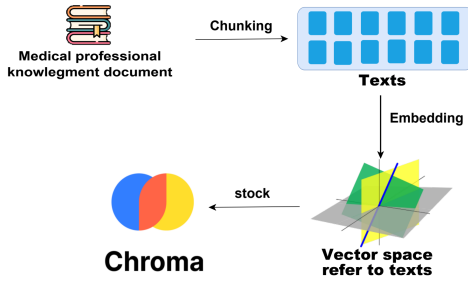


Figure 2: Step 1: Construct Database of Domain Knowledge

b) Feature extraction, as shown in Figure 3: Extracting features from raw detected health data such as ECG image, selecting key features associated with different symptoms based on diagnostic guidelines from the knowledge base, and encoding these features into text form as prompts for language models. Prompts are crucial for guiding LLMs to generate relevant responses, for example, to transform ECG into effective prompts, features need to be extracted from ECG so then they can be engineered into prompts, without considering manually extracting features, there are two approaches. The first is to utilize deep learning-based models to annotate comprehensive features of ECG, including detailed benchmark information. The second involves using LLMs with image understanding capabilities to provide coarse

descriptions of electrocardiogram features. Due to significant differences in specialization, the former is more preferable; however, regardless of the model used, both approaches represent beneficial utilization of machine learning techniques and their outcomes. On the other hand, the data collected by PPG sensors can typically be processed by algorithms to calculate a series of feature parameters, which can then be directly used as prompts. The data collected by ECG and PPG sensors already encompass many health factors.

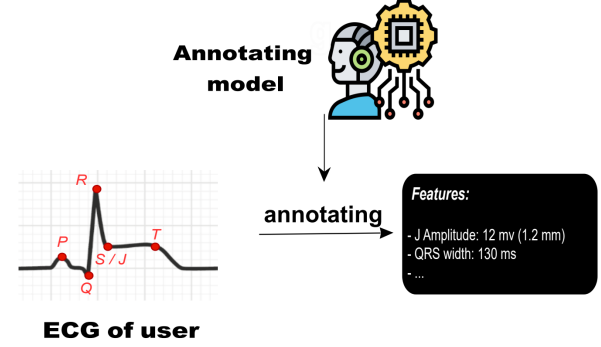


Figure 3: Step 2: Feature Extraction

c) Prompts completion and Zero-shot Diagnosis generation, as shown in Figure 4: Using the features prompts from former step, retrieving supplementary information related to observed abnormal features from the professional knowledge base, concatenating them with the feature prompts as final input prompts. Then directly using a large language model without fine-tuning to understand and generate responses, outputting structured diagnostic results. In this step, the study by Han and Guo proposes that the overall prompts should consist of five parts:

- Starter prompt: This informs the LLM that the conversation outcome involves analyzing health data for diagnosis and instructs it to comply with and refer to subsequent content.
- Diagnostic guidance: Explanatory text guiding the interpretation of diagnoses. This part can be prepared in advance by professional human doctors.
- Feature Prompts: Characteristics of the patient's health data.
- Augmenting Information: Supplementary explanations of certain details in the patient's health data features, such as explanations for features of no concern or noise.
- Format Prompts: Text explaining the rules of output format to the LLM.



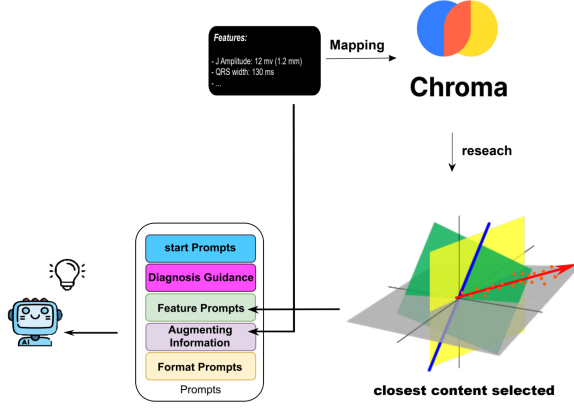


Figure 4: Step 3: Prompts completion and Zero-shot Diagnosis

I propose a way to amend “Augmenting Information” section, which involves the differences between my approach and the original approach proposed by Guo and Han: in their approach, knowledge from the knowledge base is used as reference in question-and-answer dialogues between the LLM and another entity (human or programme) regarding the domain in question, with the LLM ultimately incorporating its own responses referencing the knowledge base into the prompts as Augmenting Information. In contrast, my approach proposes directly searching the knowledge base for content similar to the features to be diagnosed, adding the returned values to into the prompts as Augmenting Information. This not only provides greater convenience but also significantly accelerates the response time for actual zero-shot diagnosis, particularly by eliminating the process of obtaining Augmenting Information through question-and-answer sessions.

In this way, LLMs can understand and diagnose complex electrocardiogram data without using any labeled data for training or fine-tuning. Guo and Han’s original method outperforms previous few-shot fine-tuning methods in tasks such as arrhythmia and sleep apnea diagnosis and even rivals supervised learning methods. Their research demonstrates the exceptional flexibility of LLMs in the field of medical diagnosis. For most diseases, preliminary diagnostic suggestions based on health data collected by wearable devices can be obtained from a knowledge base stored in vector form, using untuned language models. This will significantly reduce delays in diagnosis and help reduce misdiagnosis for patients. My solution proposed that is based on this approach and leverages

its flexibility.

I now present a specific example application scenario, as illustrated in Figure 5: Tom wears a smartwatch equipped with ECG detection sensors and ECG drawing algorithms. One day, while waiting for a bus, Tom experiences brief episodes of irregular heartbeat, chest tightness, and fatigue. He activates the software on his smartwatch, which collects 30 seconds of ECG data and sends the plotted ECG data to a large language model deployed in the cloud. The language model returns diagnostic results and recommendations to the software, which displays them on the smartwatch. Upon reading the suggestion “Atrial fibrillation, seek medical attention immediately”, Tom promptly visits the nearest hospital for medical care.

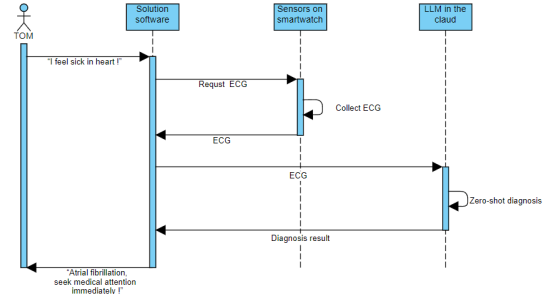


Figure 5: Scenario 1

Here’s another specific example application scenario, as illustrated in Figure 6: Bella wears a smartwatch equipped with ECG detection and PPG sensors, as well as blood oxygen measurement algorithms. The PPG sensor is programmed to periodically measure Bella’s blood oxygen levels. While she is asleep, the software determines Bella has entered the sleep state based on her blood oxygen levels. At this time, the ECG detection sensor activates at predetermined intervals and records ECG data. After Bella wakes up, the software detects her awakening via the PPG sensor and sends the ECG data to the cloud while requesting health diagnostic results. The large language model in the cloud processes the new ECG data with timestamps along with previous ECG data and returns the diagnostic results to the software, which displays them on the smartwatch. Upon reading the suggestion “positive for sleep apnea, poor sleep quality”, Bella seeks help from a medical institution nearby.

The proposed solution will help patients better understand their diseases, reducing delays in diagnosis caused by insufficient patient understanding. For patients requiring continuous monitoring, it

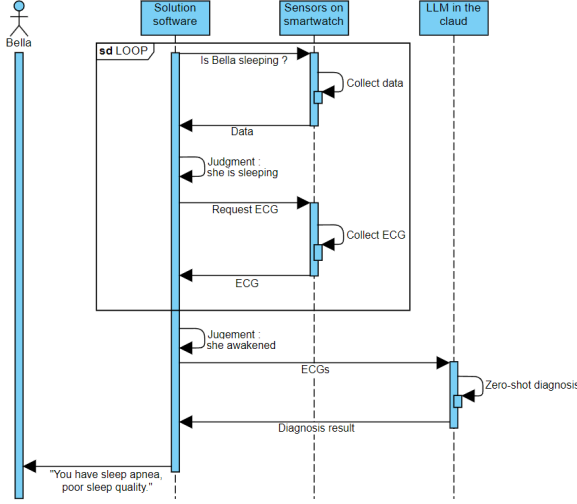


Figure 6: Senario 2

can provide communication interfaces for medical staff, allowing for the transmission of health data. In concluding, it will effectively address three out of the four challenges mentioned at the beginning:

Firstly, whether utilizing deep learning-based ECG analysis models to annotate electrocardiogram features as prompts or using specific parameters collected by PPG sensors or other sensors as prompts, the adoption of Han and Guo’s zero-shot approach for diagnosis represents a systematic and effective implementation of utilizing machine learning models for data analysis in health risk assessment.

Secondly, the solution based on wearable devices and the zero-shot approach demonstrates good scalability for adding and implementing new functionalities. For newly health risks requiring concern, when generic smart platforms (such as smartwatches) cannot collect the newly required data, data collection for the newly requirements can be achieved by adding wearable smart patches equipped with diverse sensors and connecting them to the generic smart platforms (and software on it). Furthermore, since no fine-tuning of LLM is required, enabling diagnosis of newly health risks requiring concern can be easily achieved by utilizing Chroma technology to create vector spaces storing newly additional professional knowledge; at the software level, only interface reservations are needed to accommodate the data processing requirements from different domains and the invocation of different professional knowledge vector spaces.

Similarly, the solution based on wearable devices and the zero-shot approach can easily achieve specific focus on particular aspects of specific diseases or pathologies. For any particular disease,

it only requires adding smart patches equipped with specific sensors capable of collecting specific data required for diagnosing the disease and utilizing LLMs with vector spaces storing professional knowledge in the specific field.

Additionally, in this conception of solution, it worth to note a potential challenge of securely storing user health data to respect user privacy. In this regard, commercial companies have provided valuable experiences, as their security and health protocols have been carefully considered and iterated multiple times, offering high reference value. These experiences offer the following suggestions: It is necessary to establish a central repository for health data from various devices. All interactions between this repository and health devices and software should require user permission. The data stored in this repository should be encrypted using special data protection methods, with the keys protected by deriving them from the user’s password and device UID. Access to the data should be revoked after a short period of device lockout until the next time the user enters the password or unlocks the device using facial or touch ID. If temporary log files are used to store specific health data, these temporary data should also be encrypted and deleted after being imported into the database. End-to-end encryption is preferable for data transmission. As for access permissions, a noteworthy suggestion is to utilize PPG sensors to develop more complex biometric authentication methods to verify user access to devices and data, thereby enhancing authenticity and confidentiality.

## IV. Evolution

Building upon existing approaches, I will describe simulated diagnosis examples conducted in this study, then further explore methods for evaluating the statistical accuracy of different LLMs in diagnostic tasks using the zero-shot method.

### 1. Simulated diagnosis example

In this study, I focus on the diagnosis of arrhythmia as sujet to make a simulated diagnosis example. I have selected 2 of the current most advanced LLMs as listed below:

- claude-3-opus
- gpt-4-turbo

The reasons for selecting these models include not only that they are large models with high

development investment provided by leading entity in this domain, but also an important factor, namely, due to the knowledge from the knowledge base being transmitted in the form of prompts, which will add a large number of tokens to the input, they should also have a sufficiently high token acceptance limit and processing speed. When Chroma returns 20 blocs closely related to the features from the vector space, it often results in inputs exceeding 30,000 tokens. Currently, apart from the 2 models mentioned above (and their approximate versions), other models either do not inherently support context windows of this token magnitude, or their Application Programming Interface (API) have strict token flow limitations per-minute.

I selected two professional books[1, 3] on ECG medical diagnosis and merged them into a single PDF document as a knowledge base. For knowledge base processing, I first utilized pdfminer to extract text from the knowledge base. pdfminer is an open-source PDF document processing tool written in Python, which offers a powerful interface for reading PDFs efficiently, particularly those containing large volumes of text and fields. Subsequently, I employed the chunker to segment the extracted text into semantically meaningful blocks, involving the use of a tokenizer. Specifically, the tokenizer used was cl100k\_base, provided by OpenAI, which is compatible with OpenAI’s third-generation embedding model, text-embedding-3-small. Following this, I utilized the text-embedding-3-small model as an embedding model to establish the chroma knowledge database through interfaces provided by Chroma, saving the database in its specific format locally. Although text-embedding-3-small is developed by OpenAI, the developer of the GPT series LLMs, it is sufficiently versatile for general-purpose use.

As previously described, the zero-shot diagnosis method I constructed involves searching the knowledge base for content similar to the annotated features extracted from the ECG as augmenting information. An insufficient number of returned blocks may not provide enough meaningful reference for the LLMs, while an excessive number, as mentioned earlier, will result in a heavy token load. Here, the number of returned text blocks is set to 18.

In the end, I made a simple bot using Python. It takes a textual description annotated from the original ECG as input and constructs complete prompts by combining this input as feature prompts with other parts of the prompts, including start prompts, diagnosis guidance, and format prompts. These three components directly borrow

from the examples provided in [17]. The augmenting information, as mentioned earlier, is obtained by searching for feature prompts in the knowledge base. The results of the simulated diagnosis can be found in Appendix 1 and 2. The inputted in Appendix 1 is generated by ChatGPT, in Appendix 2 is derived from real data provided in [15], specifically, it is the 501st record in the dataset provided, accompanied by the preliminary report from human physician.

The simulated diagnosis demonstrates such an example: an LLM without fine-tuning, enhanced with domain-specific expertise in ECG swiftly provides suggestive diagnoses of patients’ ECG. The diagnostic results will help patients understand their health status and prompt them to seek medical intervention promptly even after receiving concerning diagnoses, thus reducing delays in diagnosis from patients. Simultaneously, they can present the diagnostic results to healthcare professionals, aiding in drawing medical attention to potential key health issues and providing reference, thereby reducing misdiagnosis and delays in diagnosis from healthcare professionals.

## 2. LLM suitability assessment

I now propose the approach for assessing the suitability of LLMs in relevant diagnosis, it involves providing the same sets of data and knowledge bases to each LLM, conducting multiple diagnoses, and statistically analyzing the probability of their diagnoses matching those of human doctors. This process will be repeated with different knowledge bases. Subsequently, four criteria are proposed:

- **Knowledge Benchmark Accuracy (KBA):** Given a specific data sample and a set of knowledge bases, diagnoses will be conducted multiple times under different knowledge bases to determine the probability of obtaining diagnoses matching those of human doctors.
- **Simple Benchmark Accuracy (SBA):** Given a specific knowledge base, diagnoses will be conducted multiple times on a set of data samples to determine the probability of obtaining diagnoses matching those of human doctors.
- **Overall Accuracy (OA):** Given a set of knowledge bases and a set of data samples, diagnoses will be conducted multiple times on each data sample under different knowledge bases to determine the probability of obtaining diagnoses matching those of human doctors.

- Repetitive Accuracy (RA): Given a specific knowledge base and a specific data sample, diagnoses will be conducted multiple times to determine the probability of obtaining diagnoses matching those of human doctors.

It is worth noting that the accuracy of ECG annotation results directly affects the accuracy of diagnostic outcomes. In study [17], diagnostic tests were conducted using precise data provided by [15], in which, comprehensive features including the detailed fiducial information, are annotated by the original authors using both commercial and open-source algorithms. However, due to the lack of funding for the API calls required to conduct hundreds times of zero-shot diagnosis, no actual score statistics for the two mentioned LLMs were performed in this work. I hope that it could be done in the further work.

## V. conclusion

This work conducts an analysis on the continuous monitoring of health in daily life using connected health applications to aid in providing personalized and precision medicine. It particularly explores the existing commercial products on the market and reviews existing research. Present circumstances indicate that the use of wearable devices and mobile health data collection methods is gaining popularity and significant attention due to their advantages, such as minimal disruption to patients' normal lives, ease of scalability and customization, deployment and removal convenience, and affordability.

Simultaneously, this work proposes a detailed design of a scheme that integrates wearable devices-including smartwatches serving as a general platform and smart patches as expandable components, and cloud-based LLM services, to provide immediate personalized medical diagnostic advice. The proposed solution aims to reduce diagnostic delay and misdiagnoses to facilitate the implementation of precision medicine, its advantage lies in the ability for users to very rapidly obtain medical advice without the need to make special efforts to visit hospitals or schedule appointments with doctors; and its contribution lies in raising a potential possibility to addressing the issue that impeding precision medicine especially delayed diagnosis caused by the dependence of patients on their own level of awareness and concern regarding their disease and health status on the commencement of traditional medical procedures.

For the proposed solution, this work focuses on exploring methods for utilizing LLM to perform

suggestive medical diagnoses, particularly built an approach inspired by [17] with the modification on purpose of meeting the requirements for rapid diagnosis, which ultimately presents as using an LLM without fine-tuning, enhanced with knowledge base in domain-specific expertise, swiftly provides suggestive diagnoses of patients' health data detected. The advantage of this approach lies in the absence of any fine-tuning of the LLM and its ability to address issues in different domains by simply changing the knowledge base, making its application highly flexible. Additionally, this method ensures a rapid response time, as no manual enhancement prompts are required for the LLM, and diagnostic results are generated in a single session with minimal waiting. However, the disadvantage of this method lies in the heavier token load of prompts, which places higher demands on the supported context window length and language processing speed of the LLM.

## VI. Future perspectives

As mentioned earlier, the implementation of feature extraction engineering, without considering manual intervention, for raw health data, especially abstract image data such as ECG, involves either annotating and engineering prompts using specialized models based on deep learning or using LLMs with image understanding capabilities for annotation. The advantage of the former approach lies in its ability to thoroughly annotate comprehensive features, including detailed baseline information. However, its drawback is the need to call an additional model outside the LLM, and the response speed of these models may not be sufficiently fast. This could increase the time and space costs of server response to each diagnostic request, thereby prolonging the diagnostic response time. On the other hand, the advantage of the latter approach is its simplicity and speed. LLMs autonomously understand, annotate, and engineer image data into features, and then perform diagnosis. This would only marginally increase the time cost in estimation. However, the feature extraction engineering autonomously performed by LLMs may struggle to produce sufficiently accurate and detailed results in zero-shot approach, that is its drawback.

Due to my lack of medical knowledge, I did not engage in the annotation engineering of extracting medical features from the raw ECG data in this work. However, it is evident that the proposed solution requires an automated method or model for extracting medical features from the raw ECG data. The subsequent development of this work

can focus on addressing this issue by finding an approach that can automatically complete feature extraction engineering accurately and in detail, without significantly increasing the time and space costs of response. This approach should strive to retain the flexibility of zero-shot approach, meaning it should avoid fine-tuning the LLM used for zero-shot diagnosis as much as possible.

The emphasis of the proposed outcome solution lies in the rapid acquisition of health data using wearable devices and obtaining medical advice quickly through diagnosis by LLMs. Therefore, in addition to deploying a cloud-based knowledge base and LLMs, suitable IoT hardware is also required, including smartwatches and smart patches with sensors, as well as supporting software. Still, practical deployment of these parts remains to be achieved in future work.

Finally, as for the regard of the potential challenges related to securely storing user data and protecting user privacy in deploying this solution, as mentioned earlier, warrant further discussion on the practical implementation of secure methods based on existing expertise.

## Acronyms

<b>AF</b>	Atrial Fibrillation. 4
<b>API</b>	Application Programming Interface. 11, 12
<b>BG</b>	Blood Glucose. 6
<b>BP</b>	Blood Pressure. 5
<b>BT</b>	Body Temperature. 6
<b>ECG</b>	Electrocardiogram. 4, 5, 7–12
<b>HR</b>	Heart Rate. 5
<b>HRV</b>	Heart rate variability. 5
<b>ICM</b>	Implantable Cardiac Monitor. 4
<b>IoT</b>	Internet of thing. 2, 13
<b>KBA</b>	Knowledge Benchmark Accuracy. 11
<b>LLM</b>	Large Language Model. 2, 3, 7–13, 19
<b>MI</b>	Myocardial Infarction. 2
<b>OA</b>	Overall Accuracy. 11
<b>PPG</b>	Photoplethysmography. 5, 7–10

<b>RA</b>	Repetitive Accuracy. 12
<b>RAG</b>	Retrieval-Augmented Generation. 7
<b>RES</b>	Respiration. 5
<b>SATS</b>	Oxygen Saturation. 4–6
<b>SBA</b>	Simple Benchmark Accuracy. 11
<b>SG</b>	Somnography. 5
<b>SRT</b>	Stressor. 5
<b>TER</b>	Travail d’étude et de recherche. 2

## Acknowledgment

I would like to thank Mr Hamidi Massinissa, my R&D project supervisor. His consistent guidance, invaluable advice, and personal time investment have been crucial throughout the project. I deeply appreciate his unwavering support and mentorship.

## References

- [1] Tomas B Garcia. *12-lead ECG*.
- [2] Biao GENG, Wei YUE, Mengyuan LI, Yuanyuan CHEN, Tao GENG, and Zhichao LIU. Analysis of factors influencing the frequency of health check-up and health check-up behavior in healthy population. *Chinese Journal of Behavioral Medicine and Brain Science*, 32(05):455–460, 2023.
- [3] Jane Huff. *ECG Workout Exercises in Arrhythmia Interpretation*.
- [4] AAR Kamal, JB Harness, G Irving, and AJ Mearns. Skin photoplethysmography—a review. *Computer methods and programs in biomedicine*, 28(4):257–269, 1989.
- [5] Darren J Malinoski, Madhukar S Patel, Didem Oncel Yakar, Donald Green, Faisal Qureshi, Kenji Inaba, Carlos VR Brown, and Ali Salim. A diagnostic delay of 5 hours increases the risk of death after blunt hollow viscus injury. *Journal of Trauma and Acute Care Surgery*, 69(1):84–87, 2010.
- [6] David E. Newman-Toker, Adam C. Schaffer, C. Winnie Yu-Moe, Najilla Nassery, Ali S. Saber Tehrani, Gwendolyn D. Clemens, Zheyu Wang, Yuxin Zhu, Mehdi Fanai, and

- Dana Siegal. Serious misdiagnosis-related harms in malpractice claims: The “big three” – vascular events, infections, and cancers. *Diagnosis*, 6(3):227–240, 2019.
- [7] David E. Newman-Toker, Najlla Nassery, Adam C. Schaffer, Chihwen Winnie Yu-Moe, Gwendolyn D. Clemens, Zheyu Wang, Yuxin Zhu, Ali S. Saber Tehrani, Mehdi Fanai, Ahmed Hassoon, and Dana Siegal. Burden of serious harms from diagnostic error in the usa. *BMJ Quality Safety*, 33(2):109–120, 2024.
- [8] David E. Newman-Toker, Zheyu Wang, Yuxin Zhu, Najlla Nassery, Ali S. Saber Tehrani, Adam C. Schaffer, Chihwen Winnie Yu-Moe, Gwendolyn D. Clemens, Mehdi Fanai, and Dana Siegal. Rate of diagnostic errors and serious misdiagnosis-related harms for major vascular events, infections, and cancers: toward a national incidence estimate using the “big three”. *Diagnosis*, 8(1):67–84, 2021.
- [9] National Academy of Medicine of USA. Diagnostic error in health care. 2015.
- [10] Paolo Palatini and Stevo Julius. The role of cardiac autonomic function in hypertension and cardiovascular disease. *Current hypertension reports*, 11(3):199–205, 2009.
- [11] Luisa Politano, A Palladino, Gerardo Nigro, M Scutifero, and V Cozza. Usefulness of heart rate variability as a predictor of sudden cardiac death in muscular dystrophies. *Acta Myol*, 27(3):114–122, 2008.
- [12] Muhammad Wasim Raad, Mohamed A Deriche, A Bin Hafeedh, H Almasawa, K Bin Jofan, H Alsakkaf, A Bahumran, and M Salem. An iot based wearable smart glove for remote monitoring of rheumatoid arthritis patients. *Biosignals*, 2019:224–228, 2019.
- [13] Robert Joseph Thomas, Joseph E. Mietus, Chung-Kang Peng, and Ary L. Goldberger. An Electrocardiogram-Based Technique to Assess Cardiopulmonary Coupling During Sleep. *Sleep*, 28(9):1151–1161, 09 2005.
- [14] Eline L van der Veer, Joost Lameijer, Angela MP Coolen, Adriana MJ Bluekens, Joost Nederend, Maaïke Gielens, Adri Voogd, and Lucien Duijm. Causes and consequences of delayed diagnosis in breast cancer screening with a focus on mammographic features and tumour characteristics. *European Journal of Radiology*, 167:111048, 2023.
- [15] P Wagner, N Strodthoff, R Bousseljot, W Samek, and T Schaeffter. Ptb-xl, a large publicly available electrocardiography dataset (version 1.0. 1), physionet (2020).
- [16] Min YAO, Hui ZHAO, Zhuang TIAN, and Zhanfu ZHENG. Research on residents’ grass-roots medical seeking behavior and influencing factors under the background of hierarchical diagnosis and treatment system. *Chinese Journal of Public Health Management*, 35(5):585–588, 2019.
- [17] Han Yu, Peikun Guo, and Akane Sano. Zero-shot ecg diagnosis with large language models and retrieval-augmented generation. In *Machine Learning for Health (ML4H)*, pages 650–663. PMLR, 2023.

## Appendix 1. Example 1 of Prompts and Outputs

### Start Prompts:<sup>1</sup>

Identify the types of arrhythmia in the ECG signal with diagnostic guidance and the extracted features. The diagnostic guidance you should follow is detailed below. Additionally, consider the supplemental information from textbooks regarding the detected features. Please be careful about the features in the different leads.

### Diagnosis Guidance:<sup>2</sup>

When diagnosing a Myocardial Infarction (MI), various ECG changes must be considered. The ST segment elevation is a critical indicator that signals myocardial injury. For instance, if the ST elevation is observed in leads II, III, and aVF, an inferior MI is suggested, whereas ST elevation in leads V2 to V4 points to an anterior MI. Additionally, ST depressions opposite the infarct area, known as reciprocal changes, are also significant. T wave abnormalities are another aspect, where inverted or sharply peaked "tombstone" T waves can be seen in the affected leads. Lastly, the presence of Q waves, which are pathological, indicates a transmural MI and will appear in the corresponding leads of the infarct area. Conduction disturbances in the heart manifest through various changes in the ECG. A QRS complex that is wider than 0.12 seconds is indicative of a disturbance. Specifically, an RSR' or rSR' pattern in lead V1 suggests a right bundle branch block (RBBB), whereas a wide S wave or notched R wave in lead V6 indicates a left bundle branch block (LBBB). In lead III, multiple peaked QRS complexes may show localized intraventricular conduction delays. Additionally, a QRS complex wider than 0.12 seconds without the specific characteristics of LBBB or RBBB points to a generalized intraventricular conduction delay (IVCD). Hypertrophy within the heart can be detected by assessing certain ECG features. Left Ventricular Hypertrophy (LVH) is characterized by tall R waves in leads I and V5-V6, coupled with deep S waves in V1-V2. A sum greater than 35 mm of the S wave depth in V2 and the R wave height in V5 is indicative of LVH. Right Ventricular Hypertrophy (RVH) is suggested by increased R wave amplitude in V1 and a deep S wave in V6, with an R:S ratio greater than 1 in V1. Atrial enlargement is also identifiable; left atrial enlargement shows as a broad and notched P wave in lead II (P-mitrale), while right

atrial enlargement presents as a tall and peaked P wave in the same lead (P-pulmonale). (Only when there is no other diagnosable arrhythmia.) STTC is only significant when there is no other arrhythmia types, otherwise, abnormalities are diagnostic rather than STTC. Analyzing ST/T changes on an ECG requires a lookout for any obvious abnormalities. For example, wide and bizarre complexes interspersed with normal cycles may hint at ventricular tachycardia. ST depression and T waves that deviate from the normal direction are indicative of ischemia. ST elevation, which may or may not be accompanied by changes in T waves, denotes injury. The presence of pathological Q waves, especially in leads other than V1, is significant. Moreover, symmetrical or asymmetrical changes in the ST segment and T waves provide insights into the underlying pathology. Clear signs of ischemia or infarction are regional T wave abnormalities, ST segment changes, and the appearance of Q waves.

### Feature Prompts:<sup>3</sup>

I: Not Applicable  
II: Small P Wave  
III: Notched R Wave (R wave with a notch of 0.5 mm)  
aVF: Not Applicable  
aVR: Wide QRS Complex (QRS duration of 120 ms)  
aVL: Not Applicable  
V1: Not Applicable  
V2: Not Applicable  
V3: Small Q Wave (Q wave depth of 1.0 mm)  
V4: Absent T Wave  
V5: Not Applicable  
V6: Prominent T Wave (T wave amplitude of 2.0 mm)

### Augmenting Information:

Please take this data on electrocardiography (ECG) professional knowledge and refer to it: {knowledge\_base\_return}<sup>4</sup>.

### Format Prompts:<sup>5</sup>

Based on the above guidance and considering the supplemental information, please categorize the ECG into the following arrhythmia categories: - Normal ECG (NORM) ST/T Change (STTC) - Myocardial Infarction (MI) - Conduction Disturbance (CD) Hypertrophy (HYP). Please output in a Json format of NORM:, MI:, STTC:, CD:, HYP: with the explanation of each type. For

<sup>1</sup>Quote from [17].

<sup>2</sup>Quote from [17].

<sup>3</sup>Generated by ChatGPT with model gpt-3.5.

<sup>4</sup>Those blocs returned by chroma knowledge base.

<sup>5</sup>Quote from [17].

example: NORM: Result: True, Explanation: ...  
 Note: please directly provide the results (be sure with the the explanation of every each category) in a JSON format with no other explanation out of JSON format.

#### Result of GPT 4:

```
{
  "NORM": {
    "Result": false,
    "Explanation": "Presence of
      notable abnormalities such
      as notched R waves, wide
      QRS complex, and the
      presence of pathological Q
      waves disqualifies this ECG
      from being classified as
      normal."
  },
  "MI": {
    "Result": true,
    "Explanation": "The small Q
      wave in lead V3 indicates a
      localized myocardial
      infarction. Absence of T
      wave in V4 could suggest
      myocardial injury,
      potentially correlating
      with the infarcted area."
  },
  "STTC": {
    "Result": true,
    "Explanation": "The prominent
      T wave in lead V6 indicates
      ST/T change which may
      reflect localized
      repolarization abnormality
      potentially due to ischemia
      ."
  },
  "CD": {
    "Result": true,
    "Explanation": "A wide QRS
      complex as seen in lead aVR
      suggests a conduction
      disturbance, indicating
      delayed ventricular
      depolarization which could
      be due to a bundle branch
      block or other
      intraventricular conduction
      delays."
  },
  "HYP": {
    "Result": false,
```

```
"Explanation": "Although the
  notched R wave in lead III
  could imply ventricular
  hypertrophy, the absence of
  clear evidence such as
  tall R waves in V5-V6 or
  deep S waves in V1-V2 for
  LVH, and the lack of
  criteria met for RVH, leads
  to a conclusion of no
  hypertrophy present based
  on provided data."
}
```

#### Result of Claude 3:

```
{
  "NORM": {
    "Result": false,
    "Explanation": "The ECG
      shows several
      abnormalities including
      a notched R wave in
      lead III, a wide QRS
      complex in aVR, a small
      Q wave in V3, an
      absent T wave in V4,
      and a prominent T wave
      in V6. These findings
      are not consistent with
      a normal ECG."
  },
  "MI": {
    "Result": false,
    "Explanation": "While
      there is a small Q wave
      in V3 which can
      sometimes indicate a
      prior myocardial
      infarction, there are
      no ST elevations, ST
      depressions, T wave
      inversions or other
      changes suggestive of
      an acute or evolving MI
      ."
  },
  "STTC": {
    "Result": true,
    "Explanation": "The absent
      T wave in V4 and
      prominent T wave in V6
      represent non-specific
      ST-T wave changes. As
      there are no clear
      signs of ischemia,
      injury or infarction,
```



```

        and no other overriding
        arrhythmia, these
        changes can be
        classified as STTC."
    },
    "CD": {
        "Result": true,
        "Explanation": "The wide
        QRS complex of 120 ms
        in lead aVR indicates a
        conduction disturbance
        , likely an
        intraventricular
        conduction delay as
        there are no specific
        features of LBBB or
        RBBB mentioned. The
        notched R wave in lead
        III may also represent
        a localized conduction
        delay."
    },
    "HYP": {
        "Result": false,
        "Explanation": "There is
        no evidence of left or
        right ventricular
        hypertrophy such as
        tall R waves in I, V5-
        V6, deep S waves in V1-
        V2, or increased R wave
        in V1 with deep S wave
        in V6. The P wave
        morphology also does
        not meet criteria for
        atrial enlargement."
    }
}

```

#### **Note of Author:**

Although there is disagreement between the two models regarding the result for MI, from the content of the Explanation, it can be observed that both models have acknowledged the potential for MI. Therefore, their overall conclusions can be considered similar.

## Appendix 2. Example 2 of Prompts and Outputs

### Start Prompts:

Same as Appendix 1

### Diagnosis Guidance:

Same as Appendix 1

### Feature Prompts:

sinus rhythm. voltages are high in chest. st segments are depressed in i, ii, avf, v4, v5, v6.

### Augmenting Information:

Same format as Appendix 1

### Format Prompts:

Same as Appendix 1

### Result of GPT 4:

```
{
  {
    "NORM": {
      "Result": false,
      "Explanation": "The ECG
        shows abnormalities
        that exclude a normal
        interpretation."
    },
    "MI": {
      "Result": false,
      "Explanation": "No clear
        evidence of ST
        elevation myocardial
        infarction provided,
        but ST segment
        depressions and
        potential for ischemia
        suggest further
        evaluation is required
        ."
    },
    "STTC": {
      "Result": true,
      "Explanation": "ST segment
        depressions in leads I
        , II, aVF, V4, V5, V6
        indicate ischemia."
    },
    "CD": {
      "Result": false,
      "Explanation": "No direct
        indication of bundle
        branch blocks or
        atrioventricular
```

conduction delays are
mentioned from the
provided ECG
descriptors."

```
},
"HYP": {
  "Result": true,
  "Explanation": "High
    voltages in chest leads
    may suggest
    ventricular hypertrophy
    ."
}
}
}
```

### Result of Claude 3:

```
{
  "NORM": {
    "Result": false,
    "Explanation": "The ECG is
      not normal due to the
      presence of ST segment
      depression in multiple
      leads and high voltages
      in the chest leads."
  },
  "MI": {
    "Result": false,
    "Explanation": "There are
      no clear signs of
      myocardial infarction
      such as ST segment
      elevation, pathological
      Q waves, or regional T
      wave abnormalities
      suggestive of injury or
      infarct."
  },
  "STTC": {
    "Result": true,
    "Explanation": "There is
      significant ST segment
      depression noted in
      leads I, II, aVF, V4,
      V5, and V6, which is
      indicative of ischemia.
      However, no other
      arrhythmia is mentioned
      , making the ST/T
      changes the primary
      finding."
  },
  "CD": {
    "Result": false,
    "Explanation": "The ECG
      does not show any
```

```

        conduction disturbances
        such as a widened QRS
        complex (>0.12s), RSR'
        or rSR' patterns in V1
        (RBBB), wide S wave or
        notched R wave in V6 (
        LBBB), or other
        intraventricular
        conduction delays."
    },
    "HYP": {
        "Result": true,
        "Explanation": "The high
        voltages noted in the
        chest leads are
        suggestive of
        ventricular hypertrophy
        , likely left
        ventricular hypertrophy
        given the leads
        affected. However,
        specific voltage
        criteria are not
        provided to
        definitively diagnose
        LVH."
    }
}

```

**Report of human phycian:** “sinus rhythm. probable left atrial enlargement. voltages are high in chest leads suggesting lvh<sup>6</sup>. st segments are depressed in i, ii, avf, v4,5,6. this may be due to lv<sup>7</sup> strain or ischaemia.”

**Note of Author:** LLM should report true in HYP and STTC, false in the others. The 2 models used, GPT 4 and claude 3, both gave relatively accurate returns.

---

<sup>6</sup>left ventricular hypertrophy, a kind of HYP

<sup>7</sup>left ventricular