Aalto University
School of Electrical
Engineering

Master’s Programme in Automation and Electrical Engineering

Scalable Automated Test System for Pre-Verification of a Non-Invasive Glucose Monitoring Device

---

Anton Eklund

Master’s Thesis
2025

Aalto University
School of Electrical Engineering

Author Anton Eklund

Title Scalable Automated Test System for Pre-Verification of a Non-Invasive Glucose Monitoring Device

Degree programme Automation and Electrical Engineering

Major Control, Robotics and Autonomous Systems

Supervisor Prof. Ivan Vujaklija

Advisors Dr. Alejandro García Pérez, M.Sc. Risto Vänskä

Collaborative partner GlucoModicum Oy

Date 21.8.2025 **Number of pages** 74 **Language** English

Abstract

Diabetes is one of the fastest growing health concerns worldwide, currently affecting more than one-sixth of the world's adult population. Monitoring glucose levels is essential both for individuals diagnosed with diabetes and for those at risk of developing the condition. Emerging non-invasive glucose monitoring (NIGM) technologies offer a promising alternative to traditional monitoring methods, eliminating the need for skin penetration. Developing a novel wearable medical device, such as a NIGM device, is a lengthy process governed by strict regulations. Due to the complex interactions with human skin, especially in vitro testing poses unique challenges for NIGM development.

This thesis presents the design and implementation of an Automated Test System (ATS) intended for pre-verification testing of a wearable NIGM device. The goal is to support the development of the device's electronics and embedded software, by providing a reliable and accurate test environment for in vitro safety and performance testing. An ATS is a useful tool in ensuring regulatory compliance and technical performance of the device, which is essential when applying for market approval. Built on a modular PXI-platform, the ATS was designed with accuracy, repeatability and scalability as primary focus.

Performance of the ATS was assessed through a series of noise analyses and repeatability tests. The results confirm that the implemented ATS is capable of reliably performing repeated current and voltage measurements simulating real-world use of the device. Noise-induced variability was measured to be less than 2 nA, thus satisfying the ±10 nA accuracy requirement. Furthermore, the modular design allows for future scalability to support an increased testing capacity, new test protocols, and modified hardware configurations with need for minimal amount of redesign. Future work could focus on implementing parallel testing capabilities to further enhance efficiency, as this ATS only supports sequential testing.

Overall, the developed ATS demonstrates suitability as an effective pre-verification tool, improving testing efficiency and enabling a more structured, traceable, and confident progression toward regulatory approval and market readiness.

Keywords Diabetes, Medical device, Glucose monitoring, PXI-platform, Test design, Medical device regulations

Aalto-universitetet
Högskolan för
elektroteknik

Författare Anton Eklund

Titel Skalbart automatiserat testsystem för pre-verifiering av en icke-invasiv glukosmätare

Utbildningsprogram Automation and Electrical Engineering

Huvudämne Control, Robotics and Autonomous Systems

Övervakare Prof. Ivan Vujaklija

Handledare Dr. Alejandro García Pérez, M.Sc. Risto Vänskä

Samarbetspartner GlucoModicum Oy

Datum 21.8.2025 **Sidantal** 74 **Språk** engelska

Sammandrag

Diabetes är en av de snabbaste växande globala hälsoproblemen och påverkar i dag mer än en sjättedel av världens vuxna befolkning. Regelbunden uppföljning av glukosnivåer med hjälp av glukosmätare är avgörande både för personer med diagnostiserad diabetes, samt för dem som riskerar att utveckla sjukdomen. Utvecklingen av icke-invasiva glukosmätningsteknologier framskrider fort och erbjuder ett lovande alternativ som eliminerar behovet av att penetrera hud. Att utveckla medicinteknologiska apparater som dessa kräver långa processer som är strikt reglerade av lagar och standarder. De komplexa interaktionerna med mänsklig hud som en icke-invasiv glukosmätare kräver, innebär att in vitro-testande är mycket utmanande att utföra.

Detta diplomarbete presenterar design och implementation av ett automatiserat testsystem (ATS) avsett för pre-verifiering av en icke-invasiv glukosmätare. Syftet är att stödja utvecklingen av apparatens elektronik och programvara genom att erbjuda en pålitlig och noggrann testmiljö för in vitro-testning. Ett ATS är ett användbart redskap för att säkerställa att en apparat uppfyller de krav som definierats. Detta är viktigt vid ansökandet om marknadsgodkännande. Testsystemet är byggt på en PXI-plattform, vilken möjliggör utvecklandet av en noggrann, repeterbar och skalbar lösning.

Testsystemets prestanda utvärderades genom ett antal brusanalyser och upprepningstest för att säkerställa att kraven som ställts uppnås. Resultaten bekräftar att testsystemet är kapabelt att genomföra noggranna spännings- och strömmätningar då de inducerade störningarna konstant är under 2 nA. Den modulära designen möjliggör att testsystemet kan skalas upp eller modifieras i framtiden. Nya testprotokoll kan läggas till, testkapaciteten kan ökas och hårdvarukomponenter kan bytas ut utan stora ändringar. Framtida förbättringar inkluderar stöd för parallelltestande, vilket skulle öka effektiviteten avsevärt. Detta testsystem kan endast genomföra test i sekvens.

Denna implementerade ATS visar sig vara ett lämpligt och användbart redskap för pre-verifiering. Testsystemet möjliggör effektivt testande av elektroniska apparater på ett strukturerat och väldokumenterat sätt, något som är kritiskt för att en ny glukosmätare kan uppnå marknadsberedskap.

Nyckelord Diabetes, Medicinteknisk apparat, Glukosmätning, PXI-plattform, Test design, Standarder

Preface

I began working on my Master's Thesis expecting a smooth sail, however, it turned out to be slightly more challenging than that. Being involved with creating something new, a test system for a novel medical technology, meant that some answers were not at all easy to find, not even with ChatGPT.

Like with all good things, this too must eventually come to an end. Many might have doubted this day would ever arrive – including myself – but here we are. Still, this work would never have been possible to complete without the support of so many people, to whom I am deeply grateful.

Undoubtedly, the greatest thanks go to my university supervisor, Professor Ivan Vujaklija, who still found time and genuine interest in supporting me despite having many other, far more important matters competing for his attention. Ivan consistently motivated me to improve and reminded me of the importance to look at the big picture while focusing on the details.

I am deeply thankful for the support I received from my advisors at GlucoModicum, Alejandro García Pérez and Risto Vänskä. They introduced me to the world of medical technology and taught me the importance of ensuring that every idea is thoroughly thought through. I am especially grateful for spending time reading my thesis during the hottest days of the Finnish summer, when a swim or an ice cream would certainly have been more tempting.

Not only did the rest of my colleagues at GlucoModicum also welcome me with open arms and make me feel as part of their ever-growing family. They also gave me the opportunity to continue my journey here beyond the thesis. There is truly no other office I would rather walk into each morning!

Djupt tack går ut till min familj, släkt och vänner för att ni orkat tro på mej (eller att ni i alla fall inte har medgett att ni tappat hoppet). Tack för att ni orkat lyssna på min optimism och på de tusentals gånger jag nämnt ordet "dippa". Tack för stödet och hjälpen jag fått. Speciellt tack till de härliga vännerna jag fått dela en oförglömlig studietid med. Tack Gäng Gäng, Hutt Burt, Phux16, PhuxK17, PhuxK18 och hela TF! Kiitos rakkaat VTläiset, ITMK18, Aaltoes ja FallUp 2018 jengi.

It may have taken a little bit longer than I first planned, but in that sense this thesis mirrors my entire time as a student. From the beginning I knew that this was never going to be a race, but rather an opportunity to learn new skills, make so many new friends, and create countless unforgettable memories. From "sailing" around Ossinlampi in Otatarhan Ajot to taking a year off to sail around the world, these past nine years have truly been the most adventurous of my life.

As a wise man once said, "it's not the destination, it's the journey". And with those words, this particular journey now comes to an end.

Otnäs, 16 augusti 2025
Anton Eklund

Contents

Abbreviations

Abbreviations

API	Application Programming Interface
ATS	Automated Test System
AUC	Area Under the Curve
BGM	Blood Glucose Monitoring
BLE	Bluetooth Low Energy
CGM	Continuous Glucose Monitoring
CV	Cyclic Voltammetry
DMM	Digital Multimeter
DUT	Device Under Test
ESW	Embedded Software
FDA	U.S. Food and Drug Administration
HALT	Highly Accelerated Life Testing
ISF	Interstitial Fluid
MDR	EU Medical Device Regulation
MHD	Magnetohydrodynamics
MXI	MultiSystem eXtension Interface
NB	Notified Body
NIGM	non-invasive glucose monitoring
OCP	Open-circuit Potential
OQC	Outgoing Quality Control
PCI	Peripheral Component Interconnect
PXI	PCI eXtensions for Instrumentation
QMS	Quality Management System
R&D	Research and Development
SMU	Source Measure Unit
T1D / T2D	Type 1 / Type 2 Diabetes
TUR	Test Uncertainty Ratio
VISA	Virtual Instrument Software Architecture
XML	eXtensible Markup Language

1 Introduction

Personal health monitoring has become increasingly popular over the past few years as wearable devices, including smart rings and smart watches, have become more reliable and available [1]. People, even when completely healthy, actively monitor their well-being more than ever before. Despite this active monitoring, the number of people with health issues of any sort increases rapidly [2][3]. A prevalent global health condition today is diabetes with its number of diagnoses increasing significantly every year. According to the World Health Organization, in 2022 the prevalence of diabetes was 17.6 % of the world population aged over 30 years. This number was 13.5 % in 2012 and 11.2 % in 2002 [4]. Despite notable regional differences in prevalence, the trend of growing numbers is similar throughout the world. In addition to those diagnosed with diabetes, there are estimates that up to an additional half a billion people worldwide have prediabetes and should be monitoring their glucose values to avoid developing diabetes [5].

Glucose monitoring has been available since the 1960's and despite advancements in glucose monitoring technology, the only significant leap forwards occurred in the early 2000's, when continuous glucose monitoring devices became widely available [6]. Traditional blood glucose monitoring requires a blood sample, generally taken from the tip of the subject's finger, which despite being highly accurate [7], is seen as laborious and uncomfortable for the subject [8][9]. Continuous glucose monitoring (CGM) technology enables easier glucose level monitoring as the sensor probe is in constant contact with the body with no actions required by the user. However, CGM technology is still dependent on utilizing a needle to pierce the skin of the user, which can be perceived as uncomfortable for many people [10][7].

Needle fear is a frequent global problem with estimates showing that up to 30 % of people do not seek the medical help they need due to fear of needles[11]. This includes avoiding certain medical procedures, vaccines and the use of medical devices equipped with needles, such as glucose monitoring devices. Non-invasive glucose monitoring technologies that would eliminate the need to pierce skin are being developed, but only one device, the GlucoWatch, has ever been approved for commercial use. The GlucoWatch was quickly withdrawn from the market due to its insufficient performance [12][13][14]. Currently technologies that show promising results are available, [15][14] but the development process of a novel medical device is long and slow due to the numerous regulations that must be followed as well as the need for extensive clinical evidence.

A novel medical device in development needs to go through large-scale clinical trials where it is tested on humans to prove its effectiveness. However, these types of devices also need to be extensively tested in controlled environment without the inclusion of humans, to prove the safety of the device. These tests are labor-intensive and time-consuming, but automation substantially improves efficiency. Automated test systems are widely used in device testing today, with significant advances in test system technology have been made throughout the past decade, making them faster, more affordable, and capable of performing more advanced testing [16][17]. These test systems are useful tools during the development and verification process of the medical device since they can be used to perform repeated large-scale testing of the devices efficiently and accurately. Automating the test procedure ensures the tests being performed in the same way every time while reducing the risk of human error. Scalable systems on the other hand ensure the maximum capacity for testing is continually maintained despite testing demands increasing. Streamlining result processing with the help of automation is another considerable benefit, enabling faster analysis with larger coverage, thus reducing testing time. The challenge, however, especially within the medical device industry with its strict requirements, demands for all automated test systems to be carefully developed to specifically suit the needs of the company in question. No commercial systems without a need for any customization are available currently.

Designing an automated test system, to be utilized in the development and verification process of a non-invasive glucose monitoring device, must carefully follow relevant regulations and standards, as it is a tool within the medical field. The test system must be able to test the operations of the non-invasive glucose monitoring device as it performs when in contact with human skin. As it is utilizing the reactions happening in the skin to calculate the glucose values, full performance testing cannot be performed without the device having contact with human skin. Unfortunately, as no artificial skin mimicking real human skin accurately enough has been developed, testing methods focusing on feature level properties must instead be developed. Since non-invasive glucose monitoring devices are not commercially available, guidelines and examples of their development as well as their verification process are lacking.

Therefore, this thesis aims to present an automated test system designed to be utilized in the pre-verification of a non-invasive glucose monitoring device, supporting its preparation for regulatory approval and eventual market introduction. To accomplish this objective, this thesis will investigate the relevant regulations, requirements, and standards. Furthermore, this thesis designs an automated test system that can accurately test the NIGM device against its requirements with precision and reliability, as well as evaluates the performance and safety of the said system. The developed automated test system will be primarily designed to test signal measurement and output accuracy, as well as reliability in the pre-verification process of the non-invasive glucose monitoring Device currently being developed by GlucoModicum Oy. The main question this thesis aims to answer can, therefore, be summarized in the following sentence.

> "How can a scalable automated test system be developed and utilized in the pre-verification of a non-invasive glucose monitoring device?"

This question can further be divided into three sub-questions, all focusing on the different stages of the development process in focus:

1. What are the key requirements and regulatory considerations for designing an automated test system for pre-verification of a wearable medical device?

2. How can accurate testing of a non-invasive glucose monitoring device, be ensured without the use of real or artificial skin?

3. What are the key factors to consider when developing an automated and scalable test system for pre-verification purposes of a medical device?

The thesis focuses on the general design process of the system in question, as well as the hardware choices made when developing an automated test system. Software design and development are mentioned, but as the software largely builds upon existing software from a previous test system, it is not the primary focus of this thesis and, thus, not fully covered in the work.

The rest of the thesis is structured as follows. Chapter 2 gives an overview of glucose monitoring, medical device testing and relevant regulations and standards. Chapter 3 introduces how the automated test system for GlucoModicum is developed including both hardware and software. The same chapter also presents and motivates the design choices that were made during the development process. Chapter 4 focuses on the validation and verification of the developed test system analyzing its performance. Results are discussed in Chapter 5 and, finally, Chapter 6 presents conclusions and the need for further research.

2 Background

2.1 Diabetes and Glucose Monitoring

Diabetes, a chronic metabolic disorder characterized by sustained elevated blood glucose levels, has become the fastest growing public global health concern [18], responsible for close to a trillion dollars in healthcare expenditure [19] and more than one million deaths [20] globally each year. The reason bodies of individuals with diabetes cannot regulate blood glucose levels is insufficient or inexistent insulin production. Untreated diabetes can lead to severe complications such as cardiovascular disease, kidney failure, neuropathy, vision impairment and lower limb amputation[18][19]. Diabetes affects people of all ages, genders, ethnicity and geographical locations. It is a well-known illness dating far back in time. The first mentions of diabetes have been identified to be from the time of Egyptians around 1500 BC [21].

Diabetes is primarily classified into type 1 diabetes (T1D), where the body destroys insulin-producing pancreatic cells, and type 2 diabetes (T2D) where high glucose levels occur due to insulin resistance and nonfunctioning pancreatic cells [18]. Other types of diabetes include gestational diabetes that can occur during pregnancy, and prediabetes where blood glucose levels are elevated but not high enough to meet the criteria for a diabetes type 1 or type 2 diagnosis. The prevalence of prediabetes is difficult to estimate accurately, as the symptoms of prediabetes are not very noticeable with gradual progression leading to many living in the unknown. Untreated prediabetes runs elevated risk of developing into type 2 diabetes, but identifying symptoms at an early stage and initiating treatment can considerably lower the risk and even reverse it [20].

2.1.1 Diabetes as a worldwide problem

Type 2 diabetes has been recognized to be closely linked with obesity and unhealthy lifestyle [18]. As prevalence of obesity has increased drastically the past decade, so has the prevalence of T2D as well, with a growing number of new diabetes diagnosis every year. The growing trend of diabetes prevalence is similar throughout the whole world, as can be seen in the data by the World Health Organization [4] presented in Figure 1. However, it is visible that the prevalence of diabetes is growing faster in areas considered low- and middle-income, as they are currently experiencing significant lifestyle and dietary changes [22]. Economic growth in areas such as the Middle East, Northern Africa, Southeast Asia and the Pacific Islands have led to improved standards of living and decreased poverty. However, at the same time a shift from more traditional diets to a western-style diet high in processed foods, refined sugars and unhealthy fats can be seen. Increased income has led to larger consumption of fast food, processed snacks and sugary beverages, leading to increased obesity together with increased prevalence in diabetes [23].

Another clear trend globally is the significant increase of diabetes prevalence in younger age groups. This is considered to be a result of cultural changes driven by advances in technology and digital entertainment, leading to reduced physical activity and unhealthy dietary habits [19]. Studies have found a clear link between excessive screen time, obesity and an increased risk for type 2 diabetes, even among younger people [24]. A study [25] found that the average daily screen time among children and adults has nearly doubled during the past two decades, strongly correlating with reduced physical activities.

Insulin revolutionized diabetes treatment when discovered in 1921 by Frederick Banting and Charles Best [26]. This new finding showed that insulin can be used to lower glucose concentration in humans, essentially transforming diabetes from a fatal disease to a manageable chronic condition. Treating diabetes with insulin implies injecting specific doses to maintain a safe glucose level. Active monitoring of blood glucose levels is essential for knowing the size of the insulin dose needed. However, it is argued that active monitoring is also crucial for people with T2D who are not actively using insulin, making it the most important part of diabetes management [27].

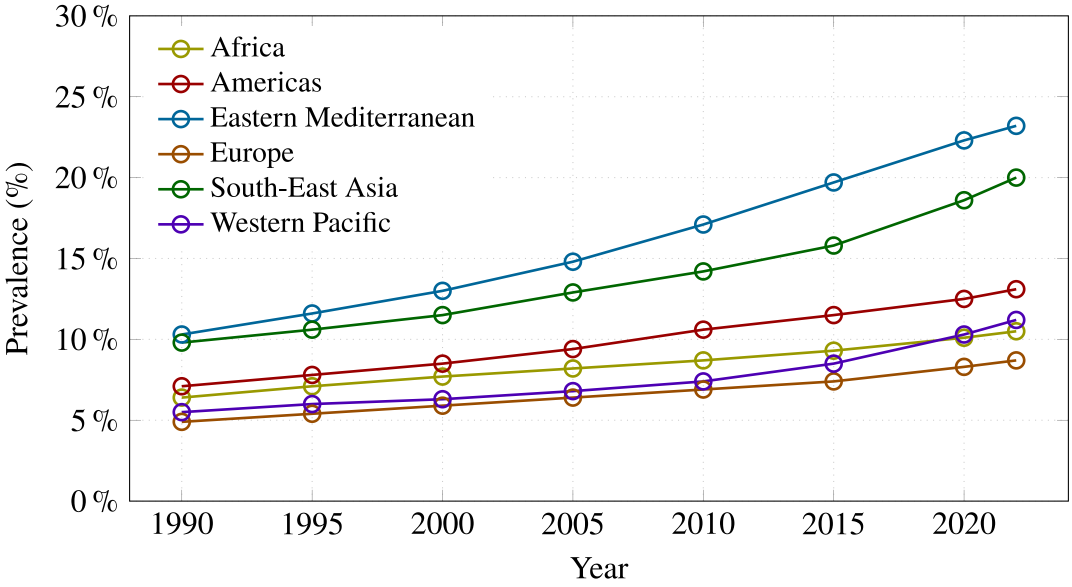


Figure 1: Age-standardized diabetes prevalence by WHO regions of population aged 18+. Source: World Health Organization database [4]

2.1.2 Blood Glucose Monitoring

Blood glucose monitoring (BGM) is conventionally conducted by taking a small blood sample from the tip of the subject's finger. A BGM device is then instantly analyzing the glucose concentration in the blood and expressing the blood glucose level generally either as milligrams per deciliter (mg/dL) or millimoles per liter (mmol/L). Research on glucose level monitoring began in the mid-1800s and throughout the early 1900s the focus was on analyzing glucose levels from urine samples [10]. This method, however, proved to be too unreliable and thus never took off. In the late 1950s and early 1960s the first systems utilizing colorimetric test strips were developed [21]. This proved to be a much more reliable method, but as it relied upon a visual comparison of the colored test strip against a results chart, the results were only qualitative.

Enzyme based electronic blood glucose monitoring became more available during the 1970s and 1980s, however, due to the high cost of such devices they were mainly for professional use in hospitals [21]. The first home use devices reached the market in the late 1970s. Enzyme-based BGM devices function on one of two operational principles. The first being reflective technology, which closely resembles the earlier color strip systems, but instead of visually assessing the color of the strip, it is done optically by measuring the reflection of light [28]. This ensures much more accurate reading compared to if done by the human eye. The other working principle is utilizing electrochemical process where glucose oxidase in the test strip catalyzes the reaction between glucose and oxygen in the blood sample, producing hydrogen peroxide. This oxidation generates a measurable current proportional to the amount of glucose in the sample [29].

Throughout the 1990s and 2000s the BGM devices continued to become smaller, faster and more accurate. The cost of such devices continued to decrease making them more available for home use [30]. However, the need to take a blood sample by piercing the skin with a needle has not been possible to eliminate. This is the reason a large share of people considers BGM to be uncomfortable [31]. A single measurement only gives the current blood glucose level. To be able to understand and predict how the level changes throughout the day needs regular and frequent measurements.

2.1.3 Continuous Glucose Monitoring

By the early 1990s, landmark studies like the Diabetes Control and Complications Trial (DCCT) [32], had firmly established that tight glucose control dramatically reduces the risk of complications. The deeper understanding of diabetes treatment highlighted by the importance of glucose monitoring, together with the rapidly increasing prevalence of diabetes, pushed for development of a novel technology capable of continuous glucose tracking. Despite recent advancements in biomedical technology and a high demand for CGM devices, it was not until 1999 that the first implantable device was commercially approved [6].

This device was the Medtronic Minimed Gold, which worked by collecting data for three consecutive days via a sensor inserted under the skin, after which the data was extracted by a doctor [33][12]. This novel technology enabled users to get information about their glucose level fluctuations over the course of a few days. However, due to the delay between when the measurement was taken and when the data could be extracted and interpreted, the device could not be utilized in decision making about immediate treatment options. Insulin treatment requires fast paced reacting to glucose level changes, which is only possible when measurements can be interpreted within hours and minutes instead of days [34].

Nearly five years passed after the launch of Minimed Gold before Medtronic got the FDA approval for the first real-time CGM device, the Medtronic Guardian RT, in 2005 [35]. This revolutionized glucose monitoring and diabetes treatment especially for people with T2D. The new technology continuously monitored glucose levels and showed the measured glucose value in real-time on a monitor connected to the sensor. But more importantly, the device sounded like an alarm as soon as it measured values considered dangerously high or dangerously low. This enabled the user to immediately take action, for example by taking an insulin injection.

Over the following decade more CGM devices reached the market making the technology more available. The devices continued to get smaller, more accurate and most of all the need for calibration reduced. The Guardian RT model required calibration every 12 hours to be able to provide accurate readings. Accuracy declines in devices from this era mainly occurred due to sensor degradation, leading to measurement drift over time [36]. To counteract this from happening calibration needed to be performed, which is generally made by taking a fingertip blood sample and analyzing it with a separate BGM device to get a reference value. Eliminating the need for calibration reduced the need for manual intervention as well as having a separate BGM device was no longer needed.

The latest remarkable leap forward for glucose monitoring technology happened in the late 2010s when hybrid closed-loop insulin pump systems were introduced [37]. This breakthrough significantly reduced the need for having a human-in-the-loop as glucose level monitoring and diabetes treatment with insulin was close to fully automated.

2.1.4 Non-Invasive Glucose Monitoring

Currently, all commercially available glucose monitoring devices require the subject's skin to be punctured. However, a non-invasive alternative would offer numerous advantages (see Table 1), including reducing the risk of complications such as infections at the measurement site. Additionally, a non-invasive alternative could also be a more attractive alternative for individuals with needle phobia. A fully non-invasive glucose monitoring device would offer a safer option for people suffering from diabetes, who already face increased infection risk due to impaired immune functions [38]. This need has driven multiple companies around the world to strive for the development of the first commercially approved, fully non-invasive glucose monitoring device. The GlucoWatch was the first commercially approved NIGM device, however, it relied on blood sample calibration, thus still requiring the skin to be punctured. To date, no fully non-invasive device has successfully achieved market approval without requiring invasive calibration.

Various methods have been explored to accomplish non-invasive glucose monitoring, generally classified into two categories based on operation style. The first category of NIGM methods consists of NIGM devices that analyze the effect glucose has on the human body. These methods include utilizing spectroscopy or optical polarimetry to measure glucose concentration by optically identifying how light that is propagated through the tissue is affected. Afon Technology is one of the companies developing a NIGM device utilizing spectroscopy, where near infrared light is emitted onto body parts with thin skin and a large amount of blood vessels, like tongue, lips or earlobes. The returning light is analyzed to assess how it has scattered, as a higher glucose concentration will result in a stronger light propagating through the tissue

[15]. Know Labs, another company with promising results developing a NIGM device, is focusing on radio frequency spectroscopy that utilizes the same principles but with radio frequency signals [39].

Optical polarimetry, another approach in this category, relies on a phenomenon where the plane of polarization of visible light is rotated due to the substances it passes through. Studies have identified that glucose molecules rotate the plane of polarization, enabling glucose concentration to be calculated from the rotation angle. This can be done for example via the eye [7]. Despite multiple research efforts, there is limited evidence that any companies are developing systems using this method. In these optical approaches, the device never directly interacts with glucose molecules nor extracts them from the subject.

In contrast, the second category includes methods where glucose is extracted from the subject's body via interstitial fluid (ISF), in a non-invasive way before measurement. This process consists of two distinct steps: extraction and measurement. Methods for extracting ISF include the use of ultrasound, as developed by Echo Therapeutics [40], and reverse iontophoresis, utilized by the device of GlucoTrack Inc [7][40]. GlucoModicum utilizes a method called magnetohydrodynamics (MHD), where ISF is extracted through the skin using a magnetic field and an electric current to induce Lorentz forces on the molecules [41]. Regardless of the extraction method, all approaches essentially utilize the same type of sensing methodology as CGM devices for analyzing the extracted glucose.

A challenge that methods of both categories face is carefully and precisely interacting with human skin and bodily tissues. Factors influencing measurement accuracy include interfering substances on the skin, physiological variability and changes in environmental conditions [15]. Despite differences in technology, it can be stated that all methods of NIGM fundamentally rely on precise measurements of one or more physiological signals. Thus, it can be further argued that the performance of a NIGM device directly correlates with the accuracy of its measurement capability. Unlike BGM and CGM devices, NIGM devices are highly sensitive to external factors, making it exceptionally difficult to replicate exact measurement procedures on artificial skin without partial or total loss of performance.

This complexity underscores the importance of careful planning and execution of in vitro testing methods throughout the NIGM device development process. Robust in vitro testing is essential to validate accurate and safe performance of the device prior to human exposure. Regulatory requirements significantly influence testing methodologies, adding another layer of complexity because they constantly change and evolve over time, often without clear advance indications of future expectations. Anticipating future possibilities and needs remains challenging, emphasizing the necessity of scalability and flexibility in testing methodologies to efficiently adapt to changing requirements.

	BGM	CGM	NIGM
Provides Continuous Real-time Data	No	Yes	Yes
Non-Invasive	No	No	Yes
Trend Analysis Capability	No	Yes	Yes
Operates Without User Intervention	No	Yes	Yes
Low Risk of Infection	No	No	Yes
Commercially Available	Yes	Yes	Not yet

Table 1: Comparison of glucose monitoring technologies

2.2 Regulatory Requirements Steering Medical Device Development

Many devices aiding humans to recover from illnesses and improving their quality of life have been developed during the past hundred years. In the early years of medical device development practically no regulations existed to guide inventors and manufacturers. Numerous new inventions were released on the market claimed to revolutionize healthcare [42]. Quickly, however, it became evident that many of these devices claimed to improve life quality did in fact do the opposite. Famous examples of how inventions were later found to pose a significant risk to health and safety of the user, leading to global disputes and market withdrawal include X-Ray devices intended for hair removal [43] and certain types of electrotherapy systems used for treating psychiatric disorders [44].

When it comes to medical devices, a close association exists between clinical effectiveness, performance, and safety [45]. A medical device should enhance the quality of life of the user while minimizing safety risks and side effects for the patient or the user [45]. To ensure all of these aspects are fulfilled regulations have been put in place to avoid negative surprises later. Three levels of regulations exist, regional, national and international [45]. Medical regulations apply for all products defined as medical devices. According to the EU Medical Device Regulation (MDR) a Medical Device is "any instrument, apparatus, appliance, software, implant, reagent, material or other article" intended to be used for "diagnosis, prevention, monitoring, prediction, prognosis, treatment or alleviation of" an "illness, injury or disability", with a few more definitions for special cases [46]. The MDR sets a framework for all products aimed at being put on the European market, setting restrictions and requirements on all aspects of the product, ensuring effectiveness in combination with safety.

Regulations and technical standards generally evolve slowly, historically often only revised as a response to serious incidents or notable failures [47]. This is also partly true for medical regulations, as many of the biggest changes to medical device regulations and standards have been made following major incidents [48]. The most well-known incident is the PIP breast implant scandal, which is widely regarded as one of the major incidents that sparked the creation of the Medical Device Regulation (MDR) in the EU. In 2009 it was discovered that the French company Poly Implant Prothèse had been manufacturing breast implants using industrial-grade silicone that was not approved for medical use. This led to significant health concerns for thousands of patients worldwide exposing a major flaw in the EU's regulatory framework. Significant changes were introduced with the MDR, including stricter requirements for clinical evidence, improved traceability and better patient safety [49].

A recent major incident which also sparked improvements in the regulations concerning assessment and testing of materials used in medical devices, was the Philips Respironics incident in 2021. Philips was forced to recall millions of CPAP, BiPAP and ventilator devices after it was found that the polyester-based polyurethane sound abatement foam inside the device tended to degrade [50]. This possessed a severe health risk for the users who were exposed to the risk of inhaling toxic particles and gases when using the devices. Philips was forced to revamp their product, manufacturing process and quality control systems to fulfill the new requirements. Another series of major incidents impacting regulations for device design were the Duodenoscope-related infections that occurred in the 2010s. Duodenoscopes were found to be the source of antibiotic-resistant infection outbreaks in patients after medical procedures. Incidents were reported in multiple hospitals globally between 2012-2019, and the cause was found to be design flaws in the devices that made them extremely difficult to thoroughly clean before reuse [51]. This led to new requirements for reprocessing protocols of reusable medical devices issued by the FDA, as well as recommendations to utilize disposable components to ensure sterility.

Major incidents have also influenced the development of stricter medical software regulations. A notable incident was the Therac-25 radiation overdose incident between 1985-1987. Due to software malfunction cancer patients undergoing radiation therapy were severely overdosed resulting in multiple injuries and deaths [52]. This led to recognizing the urgent need for stricter oversight of medical device software with increased emphasis on comprehensive validation, quality processes and safety standards. A common factor, shared by all four previously mentioned incidents, is that they all revealed a serious weakness in how medical devices are monitored after reaching the market. In response, each incident contributed to the growing recognition that more strict post-market surveillance requirements were needed. But at the same time, it can be argued that more extensive testing before market release would most likely have been able to prevent all of the incidents. This emphasizes the importance of extensive testing of all aspects of a medical device during development.

Despite regulatory changes often taking years if not decades to come into effect after the initial need has been identified, companies developing medical devices still struggle to keep up with the constantly evolving requirements [53]. The requirements for extensive clinical evidence force development processes of medical devices to extend over many years. The risk of relevant medical regulations changing in the middle of development increases the longer the development process takes. Average development times for medical devices can be 3-7 years, depending on the type of device, with higher risk devices requiring even longer times [54]. For devices utilizing novel technologies the development time is longer than normal as new clinical validation standards need to be established. Legal issues can complicate the processes even more, especially when it comes to privacy and data handling, as medical data is extremely sensitive type of information [53].

The main factor affecting development times for medical devices is the risk classification. The MDR establishes a risk classification system sorting devices into categories based on factors such as invasiveness, duration of use, amount of interaction with the body and potential risks related to the use of the product. The different classifications are class I, class IIa, class IIb and class III, and they correspond to low risk, low to medium risk, medium to high risk and strictly high risk, respectively. Traditional BGM devices are generally assigned class IIa as they require an invasive capillary blood sample to be taken. CGM and NIGM devices, on the other hand, are assigned class IIb as their duration of use is considerably longer compared to a BGM device [55]. A more severe risk classification means more strict regulations to follow, hence also longer development times [56].

Risk classification IIb requires intense clinical, safety and usability testing of the product before it can be released to the market [57]. A broad amount of performance evidence, proving clinical effectiveness in combination with safety, needs to be collected and provided for conformity assessment by a so-called Notified Body (NB). An NB is an independent, accredited organization designated by an EU member state that is tasked to assess whether a novel medical device complies with the MDR before it can be put on the market in Europe. In essence, a CE marking for a medical device is only possible to achieve once an NB has reviewed all relevant documentation and evidence, conducted audits and tests, and finally deemed everything to be according to the regulations.

EU's MDR contains all the main requirements that need to be considered when designing and developing a medical device. It regulates device design, material choices, operational features, safety systems, documentation needs and traceability processes. The MDR also defines how testing of a medical device should be performed and documented. This has a direct effect on how an automated test system needs to be designed and developed. However, there are other regulations and standards in addition to the MDR that also affect test system design. While using standards is voluntary, unless required by a regulatory authority (RA), regulations on the other hand must be followed [45].

IEC 60601-1 is a widely recognized international standard for safety and performance of medical electrical equipment. It defines a set of requirements covering electrical, mechanical and functional safety aspects to ensure patients and users are not exposed to unacceptable risk. When developing a test system, attention needs to be given to this standard as the test system needs to be able to accurately measure and ensure the device under test (DUT) meets the requirements. Even though the standard does not directly apply to a test system, it is still evident that it indirectly influences test system design and development. Similarly, the ISO 13485 standard for quality management systems of medical devices includes requirements on the extent of documentation and traceability of medical devices, which indirectly poses requirements for a test system as it is used to create these documents. Furthermore, ISO 13485 and ISO 14971 both are standards that define requirements for risk assessment and management for medical devices. Many risks can often be mitigated by thorough testing, again generating needs and requirements for the test system. In addition, the test system itself can be a risk that needs to be assessed. A malfunctioning test system passing DUTs with false positive results could have serious consequences. ISO 62304, an international standard regarding software lifecycle of medical devices, puts special requirements on extensive testing of medical device software as well.

For a medical device company based in Finland, the primary regulatory framework to follow are the ones enforced by the EU, as they are legally binding to most of Europe and govern access to the European market. Without compliance with MDR a medical device cannot obtain CE marking and thus not be sold in Europe. However, the regulations of the United States, that are primarily governed by the U.S. Food and Drug Administration (FDA), are highly relevant as well. This is because the U.S. is the largest single market in the medical technology field, accounting for a significant share of global sales [58]. EU and U.S. regulations, however, are closely aligned with each other making it easy to ensure compliance with both markets. Areas especially aligned are risk categorization, QMS requirements and the need for clinical evidence and documentation. Other markets, such as Canada, Japan and China, have their own regulatory frameworks which, while partially aligned with those in the EU and the U.S., are not fully compatible and can differ significantly in specific requirements. This regulatory difference makes it challenging for a medical device startup to ensure compliance across all potential target markets [59]. As a result, companies often focus their efforts only on a few of the markets with highest observed potential. From the perspective of diabetes, regions like the Middle East, North Africa and South Asia face world's highest and fastest growing diabetes prevalence rates, yet misaligned regulatory frameworks in these areas makes it more difficult to strive for those markets early on, ultimately leading to reduced access to novel medical technologies [18][22].

2.3 Medical Device Testing and Verification

Testing is a key part of any development process, whether it involves an active device, a static object, or software. In medical device development, testing is more important than ever, as such devices directly affect health and safety of humans [60]. Before obtaining approval to release a new medical device onto the market, extensive verification and validation testing must be completed. However, testing should start already at an earlier stage to demonstrate compliance and prevent unforeseen issues. Furthermore, as mentioned in the previous chapter, medical device development must follow strict regulations, which require that testing data from an early stage of the development is documented. In addition, ensuring that a new device meets the requirements when entering verification testing, reduces the risk of unwanted surprises which could delay getting market approval.

Regular and consistent testing throughout the development process ensures that the project continues moving in the correct direction, ensuring that progress is constantly made. Regulations typically do not specify how testing should be performed, thus it is up to the company to design and implement a suitable testing plan. Many different testing methodologies exist, and usually multiple methods is needed to achieve an extensive testing coverage assessing all aspects of the device.

Medical device companies need to prove that their medical device is reliable before it is allowed onto the market. Reliability, in the context of medical device testing, refers to the consistency and repeatability of measurements. Specifically, a reliable device consistently produces stable measurement results under identical test conditions. Thus, given identical test parameters a device should always produce identical test results with minimal deviation or variation [61].

One of the primary factors influencing reliability is measurement accuracy, defined as the degree to which measured values conform with their true reference values [12]. Factors impacting measurement accuracy include systematic deviation, variation due to interference and noise, offset error and measurement delays. Thorough and repeated testing provides the most effective means to assess the accuracy of a device.

Ensuring reliability and accuracy requires that all measurement results are metrologically traceable [62]. Metrological traceability means that each measurement can be related to an international reference standard through an unbroken chain of documented calibrations [62]. This type of traceability is essential for ensuring comparability of test results and regulatory compliance. The international standard ISO/IEC 17025 specifies general requirements for calibration of instruments and is one of the cornerstones for achieving metrological traceability.

2.3.1 Challenges in Wearable Medical Device Testing

A significant challenge in wearable medical device testing arises from handling sensitive health information [53]. Clinical trials, which are essential for device validation, involve collecting data that can directly or indirectly be linked to individuals. Such information is classified as protected health data under regulations like GDPR in the EU and the Health Insurance Portability and Accountability Act (HIPAA) in the US [57]. Consequently, special procedures for collecting, processing and storing the data, including data anonymization, must be implemented [57]. Managing these data protection requirements adds complexity and risk, slowing down development cycles and restricting availability of quick feedback on design choices.

Other factors also lead to clinical trials being challenging, resource-intensive, and strictly regulated. Recruiting subjects, ensuring ethical compliance, applying for permissions from the governing authorities, and navigating practical complexities increases development timelines and costs significantly. Therefore, early-stage testing conducted without human subjects is highly beneficial whenever possible.

However, in vitro testing of wearable medical devices introduces its own set of challenges, one of the largest being how to realistically simulate the human factor. Devices designed for direct skin contact must perform consistently and safely across diverse conditions. Skin properties can differ significantly between two individuals, and even within the same individual over time. Skin properties such as roughness, moisture level, and hair coverage can differ substantially and directly influence device accuracy and performance [63][64]. Additionally, everyday substances, such as sunscreen or skin lotion, can further affect device operation.

Beyond physiological variability, wearable devices must withstand diverse environmental conditions encountered during real-world use, and these are difficult to replicate accurately in laboratory settings. Temperature fluctuations, vibrations, humidity changes, and movement all significantly impact device functionality. Effective in vitro testing must therefore incorporate methods to simulate these dynamic conditions as realistically as possible.

Lastly, as mentioned in the previous chapter (Chapter 2.2), despite regulations often evolving slowly, in certain situations development of novel wearable medical technologies face the risk of regulatory standards changing too rapidly. This is especially relevant for novel technologies as product development cycles can extend longer than regulatory framework revision processes. In situations like these, compliant devices might suddenly become non-compliant before reaching market approval [65]. To minimize this risk, companies need to stay alert and incorporate flexible testing processes and modular test systems that can quickly be adapted to new regulatory needs.

2.3.2 Success Factors of Medical Device Testing

Multiple factors play a crucial role in achieving successful testing of a medical device. One important factor is how tests are designed. A highly recommended approach is to use a comprehensive risk-based test planning approach [66]. This involves designing test strategies based primarily on potential risks that have been identified to be associated with the use of the device. As mentioned in 2.2, a hazard identification must be conducted for every medical device according to ISO 14971. Potentially hazardous cases identified and included in the hazard assessment can then be utilized to design test scenarios.

Another important consideration is that testing must confirm compliance also with regulatory and technical requirements of the device. Simply confirming device safety is insufficient, the device must also prove to be effective and fulfill its intended purpose. Verification testing is critical, especially for a medical device that is prepared to be deployed to the market. IEC 60601 covers extensive requirements that a medical device needs to comply with and serves as a useful tool when designing tests.

The extent of testing should be proportional to the criticality of each test case. Insufficient testing might undermine confidence, while overly extensive testing with every possible parameter combination is impractical. It is recommended that critical safety tests and regulatory compliance tests are conducted more thoroughly, as failure could lead to severe consequences, including significant design changes [66]. It is also important to test devices at the end of their lifetime, to ensure that they continue to meet the requirements even after long-term use, and that no sudden issues arise from repeated use. This type of testing is referred to as HALT testing, and includes subjecting the devices to extreme environmental stresses beyond normal conditions [67].

Successful testing also demands that test systems themselves are sufficiently accurate to provide reliable assessment of the devices. This is especially important when evaluating measurement capabilities of a device. A common practice is to use the Test Uncertainty Ratio (TUR) to determine the requirements for measurement accuracy of a test system. Historically TUR was considered to be 10:1, however, nowadays a TUR of 4:1 is widely accepted as being sufficient [68]. This shift is recognized to have occurred due to advancements in measurement technologies, leading to commercial-off-the-shelf devices for consumers inhabiting similar levels of accuracy as laboratory-grade devices used by professionals.

2.3.3 Testing of Non-Invasive Glucose Monitoring Devices

NIGM devices, like any medical devices, need to undergo thorough testing during development before they could potentially be introduced to the market. However, a significant challenge that sets NIGM devices apart from many other products, is the complex operation principle that makes in vitro testing difficult. As mentioned in Chapter 2.1.4, NIGM devices rely heavily on analyzing the effects glucose has on the human body or actively interacts with the body to access glucose. This difference in operation style compared to traditional BGM or CGM devices means that the same testing practices cannot be used for NIGM devices. While BGM and CGM devices face their own challenges related to measurement accuracy, sensor drift, and sensor degradation, they can be tested in vitro by exposing them directly to a reference liquid with a known glucose concentration. However, NIGM devices cannot be tested in the same way, as the measurements are based on interaction with human skin. The high sensitivity to skin conditions, environmental factors and individual differences all affect measurements, leading to significant challenges in accurately performing system-level in vitro testing.

Despite these challenges, testing of NIGM devices remains essential to ensure safety and effectiveness before they are approved for market. As a first step toward formal verification, many companies rely on pre-verification testing. Pre-verification refers to testing activities performed internally prior to initiating the official verification process. The aim of pre-verification testing is to steer the later part of the development phase in the right direction, ensuring that verification testing will be successfully passed. Accuracy, repeatability and robustness, as well as safety is tested against set regulations, applicable standards or earlier defined requirements. Any potential design flaws or safety issues can be identified and corrected while it still is easier to make design changes. Especially for novel technology, such as NIGM technology, pre-verification is immensely important to build up the confidence needed and technical evidence to move closer to regulatory compliance.

To overcome the challenges associated with system-level in vitro testing of NIGM devices, a pre-verification approach can be applied by focusing on individual subsystems separately. In GlucoModicum's case the NIGM device consists of two distinct parts: a reusable electrical device and a replaceable biosensor. Thus, in vitro testing is logically divided into two primary categories: biosensor testing and electronics and ESW testing.

As the biosensor is detachable it can easily be evaluated independently. The biosensor generates electrical signals through electrochemical reactions that are relative to the glucose concentration it is exposed to. By exposing the detached biosensor to various substances and analyzing the resulting electrochemical reactions with the help of instruments, a clear correlation between exposure to the specific substances and the electrical signals produced by them can be established. These tests do not directly determine the accuracy of glucose measurement but rather aim to characterize the sensor's behavior under controlled conditions.

In parallel, the electronics and ESW are tested independently. Reference signals simulating the behavior of the biosensor are generated using calibrated lab instruments to reduce variability. This enables assessment of measurement capabilities of the electronic components, and accuracy of the calculations performed by the ESW. The ATS developed in this thesis is designed specifically for this second type of testing, focusing on the electronics and ESW rather than biosensor evaluation.

A typical test protocol for electronics and ESW testing of such a device may include the following steps:

1. Hardware bring-up to ensure correct manufacturing of the device
2. Functionality testing to assess that all operations perform as intended
3. Accuracy and reliability testing with a test bench to assess performance of the device
4. Environmental stress testing to assess performance with different external factors
5. EMC, ESD and other safety related testing to assess that the device handles special circumstances
6. Long term stability testing (including HALT testing) to confirm that the device is reliable and safe throughout its whole life cycle

For each test type, acceptance criteria are typically derived from relevant standards, applicable regulatory requirements, or technical specifications defined during development. This structured way of testing a device in development helps ensure that the final product meets both the technical needs, as well as the regulatory requirements. This also enables collection of structured documentation to serve as evidence that can be referenced in future verification stages, during audits, or when communicating with regulatory bodies. Establishing a traceable documentation history that can be used to accurately recreate the testing procedures later is particularly important for novel medical technologies. Comprehensive documentation can be used to clearly justify design decisions and to support performance claims at a later stage, and is often required by regulatory bodies.

2.3.4 Verification of a Medical Device

Verification is a critical step in the development of a new medical device. The intent of thorough verification is to ensure that the device has been designed correctly aligning with intended specifications and requirements. Medical devices, especially those of a higher risk classification, are required to undergo an official verification process, as they must receive approval from relevant local authorities before entering the market. For devices targeted to be placed on the EU market it means that it needs to get CE marking. The CE marking is received by getting the approval from a Notified Body (NB). The MDR states the requirements needed to get the approval from a NB, and they include passing safety tests, sufficient documentation, implemented risk management and evidence of clinical performance. For the US market the governing authority is the FDA that sets the requirements and approve new devices. The procedures and requirements are, however, highly aligned between the EU and the US.

The verification process of a medical device entering the EU market can be divided into five distinct steps. First the applicable directives need to be identified, and the device needs to be classified. A device classified as a medical device will be affected by different regulations compared to a device classified as an in vitro diagnostic device. Furthermore, the device needs to be given a risk classification. As mentioned in Chapter 2.2, non-invasive glucose monitoring devices are generally assigned Class IIb. These classifications will later define many of the requirements that the device must adhere to.

The second step in the verification process is ensuring that the developed device meets essential requirements. The company developing the device will define many of the technical requirements. However, there are numerous other requirements, such as the safety related requirements listed in 2.2, that are defined by the governing authorities. It is useful to utilize harmonized standards to guarantee that the devices pass any audits performed by governing authorities at a later stage.

Once the design has been confirmed to be compliant with all requirements, an official conformity assessment will be conducted by the NB for EU approval, or FDA for U.S. approval. This is only mandatory for high-risk devices, such as those of class IIa, class IIb and class III. In a conformity assessment the focus lies on four key areas: reviewing technical documentation, auditing the quality management system, assessing the clinical evidence and testing of the device itself.

Especially the quality management system is thoroughly examined to assess compliance with ISO 13485. This assessment goes beyond the product itself and also ensures the company has all required processes up to standard. One such process is the post-market surveillance, obligating companies to track performance of the devices throughout their whole lifetime [69]. The clinical evidence is closely examined to confirm that the device performs in a safe manner and is capable of producing the claimed results. Safety testing is also repeated by the NB to ensure compliance with IEC 60601 and other relevant standards.

If the NB determines that the device meets all regulatory requirements, the manufacturer is authorized to affix the CE mark onto the device. A CE-marked device can subsequently be registered and legally introduced into the EU market. However, the verification process does not conclude at this point. As previously mentioned, continuous quality assurance must be continued throughout the lifetime of the device. Ongoing vigilance and systematic post-market surveillance are essential to secure continued compliance. Any deviation or breach of the requirements could result in the withdrawal of the CE marking. Such a withdrawal would have serious implications for the manufacturer, potentially requiring immediate termination of device sales and, in the worst case, forcing previously distributed devices to be recalled.

2.4 Scalability in Automated Testing

Scalability is a fundamental requirement in automated test systems, ensuring that a system can efficiently adapt to evolving demands. These demands may include increases in the number of devices under test (DUTs), expanded test requirements, or changes in hardware and software specifications. Scalability is particularly crucial in medical device testing due to strict regulatory requirements and long development processes that require extensive testing documentation. A truly scalable ATS can integrate new measurement instruments, accommodate multiple DUTs, and support new test cases with minimal redesign efforts.

Automated test systems can scale in many ways, including hardware, software, and process scalability. Hardware scalability encompasses the system's ability to easily integrate additional instruments, adding switching modules, reconfiguration of the test circuit, and handling an increasing number of DUT volumes tested either in parallel or sequentially. Developing a large-scale test system immediately might not be practical due to complexity, increased cost, and evolving testing needs. Thus, designing a system with built-in scalability enables efficient future expansion without extensive redesign or financial burden.

Software scalability focuses on creating modular and reusable software components, and designing tests that can be easily modified, extended or replicated to address future testing scenarios. Designing software that supports multiple device variations with minimal modifications helps future-proof the test system. Similarly, as with hardware, starting with overly complicated software developed beyond current needs might unnecessarily increase initial development time. Therefore, a modular design facilitates easier modifications later and promotes reuse of existing components, ultimately reducing cost and time required.

Process scalability addresses the system's ability to efficiently manage increased testing throughput without compromising efficiency or compliance as requirements evolve. Early in development, testing volumes are often smaller and test scenarios simpler. As product development progresses, new unforeseen testing needs may emerge. A scalable ATS can quickly adapt to these new requirements without significant downtime due to system redevelopment.

However, scaling an ATS to support more DUTs, new test scenarios, and incorporate with additional features, increase complexity exponentially, raises the potential for errors and increases areas requiring validation. Hence, to ensure an ATS can meet the needs of the future, it needs to be developed with scalability in mind right from the start. Another significant challenge to consider when scaling up an ATS is managing the exponential growth of generated data. Ensuring reliable data processing and efficient storage becomes progressively more challenging with increased test volumes. Nonetheless, the effectiveness of an ATS relies on precise data handling, as unprocessed data will limit the practical value of the entire system.

To successfully incorporate scalability into an ATS, utilizing interchangeable modules and instruments connected through industry standard interfaces is highly recommended. Choosing off-the-shelf components instead of custom-built alternatives further ensures that the hardware can be upgraded and expanded with ease. While industry standard communication interfaces, such as IVI, enable that instruments can easily be interchanged, they can sometimes limit instrument functionalities. Vendor specific APIs, on the other hand, provide taking full advantage of the instruments, but may constrain interchangeability. Thus, deciding between flexibility and full functionality is critical early in the ATS design phase.

An additional critical decision point in designing a scalable ATS concerns simultaneous testing capabilities. An ATS accommodating multiple DUTs can adopt either parallel or sequential testing methods. Parallel testing, where multiple DUTs are tested simultaneously, is significantly more time-efficient, but substantially increases system complexity. Parallel systems require duplication of instruments, signal paths, and computational resources for each DUT. This results in significantly higher initial costs and increased validation requirements. Sequential test systems, on the other hand, only utilize a single set of testing resources to evaluate DUTs individually. Although sequential testing extends total testing time, it significantly reduces complexity, cost, and risk associated with maintaining and validating more complex parallel systems. Therefore, choosing between a parallel and a sequential testing alternative should reflect the testing needs, available resources, and the acceptable level of system complexity. Given these considerations, careful planning and decision-making are important to design a truly scalable ATS that effectively balances immediate testing requirements with potential future needs.

3 Design and methods

A company striving to develop the first non-invasive glucose monitoring device in the world will start off purely with research of novel technology in mind. Despite the uncertainty of not knowing whether a functioning and market approved device will ever be reachable the company needs to start focusing on other aspects of product development early on. As stated in Chapter 2.2 regulations are strict within the medical field, making development processes long [70].

GlucoModicum, a Finnish medical technology company, is in that situation right now. To ensure readiness for launching a regulatory-compliant NIGM device onto the market once the novel technology under research is validated, the company has already advanced significantly in the product development process. Guaranteeing that the device will be regulatory compliant is a key focus. This has led to an increased demand for product testing, opening up a need for a reliable automated test system. With the help of such a system the testing capacity could be increased, as well as ensuring reliable repeated testing over both short and long timeframes. An ATS would be a useful tool in evaluating that the technical requirements of the device design are met. This is a critical step in the path of preparing a NIGM device for the market.

The following chapter will introduce and describe the development of a scalable automated test system designed for pre-verification of the NIGM device that GlucoModicum is developing. The design of the system is driven by the need for a highly reliable and efficient test system that increases test capacity at the same time as it minimizes room for variance, such as variance caused by human error. Another key factor that will be in focus is the need to develop a system that is easily scalable. This is important to be able to accommodate increasingly growing and possibly changing future needs. A high precision test system is a considerate investment for a small company not making any revenue yet, thus it is necessary to ensure the lifetime of the test system will be maximized.

3.1 Development Process

For the development of any product to be successful the development process itself needs to be well planned and executed. The development process for the development of this ATS was decided to be split into four distinct phases. Despite slightly overlapping, each phase incorporated a distinct start and end point, defined by clear actions. The four phases were the Design Phase, the Development Phase, the System Testing Phase and the Deployment Phase.

The project began with the Design Phase, during which the design principles and a general high-level design were defined. The specific needs, requirements and restrictions were also defined in this phase by the project group, laying the cornerstones of the work. A project plan was produced stating all necessary details about the automated test system to be built. Once a clear implementation plan was constructed and the technical requirements were clearly defined, the project could transition into the next phase, the Development Phase. Without clearly defining needs and requirements there is a significant risk that the ATS, once completed, will not be able to provide the necessary results. If pre-verification is not possible to conduct due to limitations of the test system, it will render the development of the test system completely pointless, as that is the main need.

With clear requirements defined the development could be initiated. In this project, the Development Phase was defined to begin once all critical decisions necessary for acquiring hardware or initiating the implementation of software were confirmed. The main aim of this phase was to produce a system that was believed to fulfill all the needs and requirements defined in the Design Phase. Often the development phase is divided into multiple sub-phases to maintain a clearer structure of a long and complex phase. Here, however, it was decided not to, as the scale of the project was deemed small enough.

The third phase of the project was the testing phase, where the developed ATS was thoroughly tested, calibrated and verified to ensure it meets all the pre-defined requirements. Test results were documented and stored to serve as evidence in the future. Once the ATS was successfully verified and validated, and all the appropriate documents were approved it could officially be deployed for use. Readiness for deployment indicated that the project could move into the final phase, the Deployment Phase.

The Deployment Phase includes, as the name states, the deployment of the ATS, but also everything surrounding taking it into use. From this stage onward, all possible changes made to the system or related processes need to follow a configuration control process. Once the ATS was successfully in use and documentation had been completed, the deployment phase, as well as the project in whole, could be deemed to be concluded.

3.2 Design Principles

During the Design Phase a set of design principles were defined. This was set to ensure that the finalized test system would meet the needs of GlucoModicum. Design principles are essentially fundamental guidelines steering the development towards fulfilling the requirements. Requirements, on the other hand, can in themselves be seen as clearly defined detailed end goals. While requirements are measurable, design principles are more subjective. This creates a clear distinction between the two, as requirements can thus be verified, while design principles cannot.

The main design principle was defined as "the design and development of the test system must follow a top-down design". A top-down design is a systematic approach where a complex system is broken down into smaller subsystems before focusing on specific implementation details [71]. Before starting with designing the test system, a proper understanding of the specific needs that the test system is intended to fulfill, is essential. The primary need of GlucoModicum was to have a test system capable of performing safety and performance testing on their device in vitro, to be able to collect evidence data for the verification process of the product. This was further split into multiple subsets, all contributing towards the goal of collecting verification evidence. The different subsets corresponded towards a distinct feature or property of the device. Each of the subset could then be further divided into specific requirements contributing to a common list of requirements for the test system specified in Chapter 3.4.

A second design principle was set to require the test system to be designed with reliability and repeatability in mind. As stated in Chapter 2.3, the DUT needs to be proven to be reliable. This is only possible if the test system is able to assess the accuracy of the DUT with high reliability. Accuracy and measurement integrity is one of the main properties that the test system must possess. Multiple requirements have been defined to ensure accuracy and measurement integrity. Users of the test bench need to be able to trust that every test result produced by the system is accurate and reflects the properties of the device. Tests need to be possible to repeat on multiple devices multiple times. Tests might be performed at various locations, so it must be ensured that no room is left for speculations that the test location impacts test results. Results produced by the test system will be used as references when additional testing is performed later. This is useful when there is a need to ensure no unwanted changes have occurred because of an ESW update, or to prove shelf-life requirements are met. In these cases, the user needs to be able to trust the testing is identical and the only variable is the DUT. Thus, the test system needs to be designed and developed with this important need in mind.

The third general design principle is that the test system should be automated and efficient. Manual verification testing is possible to perform on devices like the NIGM device of GlucoModicum. To ensure the usefulness of the test system once completed, it must be more efficient than a human performing the testing manually. The test system can achieve this by actively utilizing automation together with a system design that minimizes the need for human input. This does not only increase efficiency, but it also minimizes the risk for human error ensuring accuracy and measurement integrity.

The final design principle is that the test system needs to be regulatory compliant. It affects the whole development process, not only just the end product, and must thus be defined as a design principle. Despite the test system not being a medical device on its own, according to EU's MDR, it is part of the development of a medical device and thus is regulated. Software development, hardware components, verification procedures, and most importantly documentation needs to be fully compliant with regulations as a non-compliant test system could have severe consequences for a medical device in development.

These design principles have laid the path for the design and development of the automated test system that will be used for pre-verification of the NIGM device. However, it is important to note that the test system developed in this Master's Thesis serves as a prototype intended to demonstrate the feasibility of such a system. Consequently, the specific system will not be used in official verification process of the product. While this prototype might not be fully compliant in itself, the design and methodology of the system should fully align with all relevant regulations.

3.3 Device Under Test

This specific automated test system is designed and developed to function specifically as a tool for pre-verification of GlucoModicum's NIGM device. To understand the requirements for an automated test system it is crucial to understand the device it will be testing. The device, its properties and how it operates set most of the requirements for the automated test system. To be able to perform reliable and repeated testing requirements like that must be met.

The NIGM device consists of two components, a reusable electronics module and a single-use biosensor. ESW, within the electronics module, controls the device and handles communication with a connected smartphone via Bluetooth. The two components can be snapped together, and an electrical connection is thus established by pogo pins ensuring that signals can be transmitted between the biosensor and the device. When in use, the device is placed on the upper arm of the user with help of medical adhesives on the biosensor, where it operates automatically.

The system utilizes MHD technology to non-invasively measure glucose levels of the user [72][13][41]. The process is a finely tuned combination of electromagnetic fields, electrochemical reactions, electric signal processing and data analysis. Every step of the operation occurs within the device and only one glucose measurement is relayed to an external controlling device like a smartphone.

Four electrodes make up the working surface of the biosensor. The electrodes are covered in specific metal compounds and surrounded by hydrogel to accommodate the electrochemical reactions. The four electrodes have all different purposes and are named Extraction Electrode (EE), Counter Electrode (CE), Working Electrode (WE) and Reference Electrode (RE), respectively.

The device performs repeated measurement cycles consecutively where each successfully completed cycle results in one glucose value data point. One cycle consists of three different operation types: extraction, measurement and preservation, performed in respective order.

The extraction operation is named after the goal of extracting ISF from the skin to the biosensor. During extraction an electric current is applied from the EE to the CE through the skin of the user. This current, in combination with the magnetic field produced by the neodymium magnets inside the device, creates an outwards pointing force on the ISF in the skin. This force occurs due to the Lorentz force law and causes ISF to seep through the skin towards the device's biosensor.

When extraction is performed and ISF reaches the hydrogel around the WE, the glucose extracted to the hydrogel reacts with glucose oxidase enzyme and oxygen, producing hydrogen peroxide. The hydrogen peroxide is consumed in a redox-process at the WE, producing a current signal proportional to the amount of extracted glucose.

Currents produced by the electrochemical reactions are measured by a potentiostat circuit, after which the signals are digitized and processed by the ESW to get a glucose concentration estimate. After successful measurement operation the device performs a preservation operation, which is similar to an extraction operation, but with a current in the opposite direction. When applying a current from the CE to the EE the force affecting the ISF and the glucose molecules switches direction effectively pushing particles away from the WE. This procedure ensures that the biosensor lasts longer without material decay.

As shown in Figure 2, it can be generalized that the objective of the biosensor is to transform the glucose concentration in the ISF into measurable electric currents. These

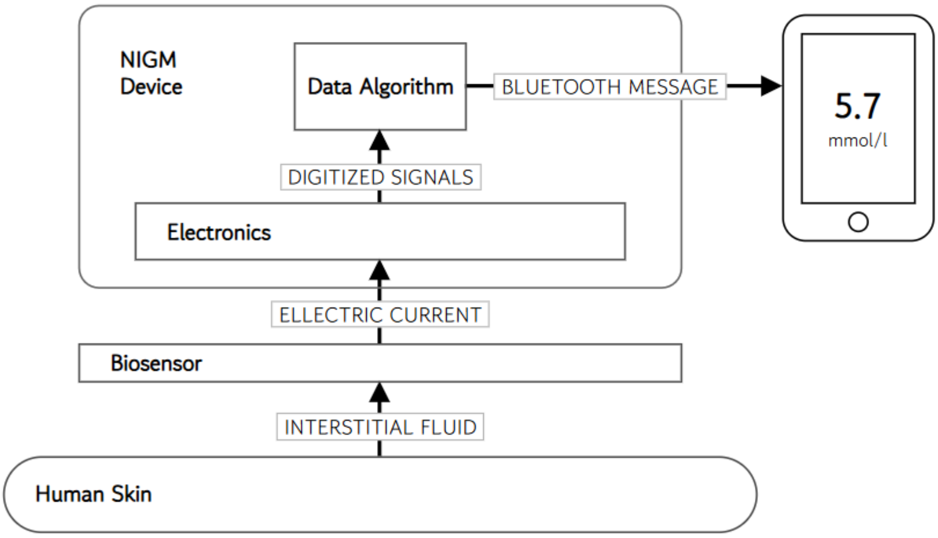


Figure 2: Signal diagram over GlucoModicum's NIGM device

currents are then received and processed by the reusable electronics module ("NIGM device" in Figure 2). Therefore, repeatable and detailed testing can be performed on the reusable electronics module by substituting the biosensor for measurement instruments. This eliminates significant variance created by multiple factors, such as the variance in biosensors and the variance in skin properties. Testing in this manner is also significantly faster, as there is no need to measure glucose concentration as a reference by analyzing a blood sample, instead reference values are instantly available from the current measuring instruments. Performing system testing in this manner evaluates the accuracy and reliability of the measurement and output capabilities of the device with a focus on the electronics and ESW. The biosensor undergoes separate testing and the product in completeness is tested and evaluated in clinical studies. These are not within the scope of this thesis and thus will not be addressed.

Testing the device with the help of measurement instruments enables testing with specific user specified parameters, including parameters that are not considered to be part of normal operation. This is especially useful when it comes to safety testing of potential single fault situations. Such tests include the electrical safety tests defined by the IEC 60601-1 standard, mentioned in Chapter 2.2. These tests need to be performed on all electrical devices. Thus, as the device operates on a 4-volt battery, thorough electrical safety testing remains essential, even for devices generally classified as low-powered, like this one. Using an ATS it is possible to simulate and test different single fault states and ensure that the safety mechanisms of the device operate as they should.

3.4 Test System Requirements

Ensuring that the final product meets previously defined needs is a critical aspect of developing any technical device or system. A fundamental practice to achieve this alignment is through the explicit definition of specific, measurable, and verifiable requirements that the finalized product must fulfill. Requirements serve to translate abstract goals into concrete criteria. Fundamentally, requirements facilitate systematic evaluation of the system's performance, ensuring that the development aligns consistently with the initial objectives and intended use.

In the context of developing an ATS for verification testing of a medical device, clearly defined requirements play an exceptionally important role. As mentioned in Chapter 2.2, medical device development requires following strict regulations and compliance with numerous internationally recognized standards to ensure device effectiveness together with patient safety. Consequently, these regulatory and standard-based requirements for the device can directly be translated into requirements for the ATS. Essentially, the requirements of the device under test form the basic guidelines and set clear expectations that must be thoroughly met by the test system.

There are also regulatory requirements directly addressing test equipment used for medical device testing, however, these requirements are generally less extensive than those applied to commercial medical systems intended as end products. For instance, ISO 13485, the standard for Quality Management Systems, defines requirements for test method validation and result traceability that a test system used for device verification must comply with [73].

In addition to requirements established by regulations and standards, further requirements emerge from existing software or hardware components and systems with which the ATS must integrate. The operational environment itself also defines specific requirements for the features and specifications of the test system. These requirements consider both the conditions in which the ATS operates, but also those in which the DUT is tested. To accurately simulate and evaluate DUT performance across all potential usage scenarios, the test system must be capable of producing, accommodating, or withstand these varied environmental conditions.

Moreover, internal company-defined requirements are essential for guiding the development process to remain consistent with strategic goals and initial design intentions. Such requirements protect the development effort from scope drift or over-engineering. They also aid with optimizing resource utilization and indicating what to priorities if design compromises must be made. Although these requirements may not directly originate from external specifications, they are nevertheless well-founded and carefully motivated to increase clarity throughout the development process.

Detailed requirements documentation serves multiple critical purposes. It provides essential support for regulatory compliance by clearly demonstrating conformity with relevant standards and regulations. Additionally, comprehensive requirements definitions function as a communication tool between developers, regulatory authorities, and management. By clearly outlining development goals and expectations, all stakeholders are kept aligned. Requirements documentation also provides crucial insights into the reasoning behind specific design decisions, thereby justifying technical choices made during development. This significantly improves transparency, simultaneously reducing the risk of repeating past mistakes.

Ultimately, requirements form the fundamental basis for the verification and validation process, setting clear benchmarks against which the system's final performance and success are assessed. Effective requirements must, therefore, be sufficiently detailed to minimize ambiguity, rigorously strict to ensure correct and safe functionality, yet flexible enough to remain realistically attainable. In this thesis, the non-functional, hardware and software requirements outlined in the following subchapters have been defined to support the primary aim of developing a reliable, accurate, scalable, and regulatory-compliant ATS for pre-verification of a NIGM device.

3.4.1 Non-functional requirements

Non-functional requirements guide the overall properties and features of the test system, shaping behavior, usability, and quality of the ATS. These are not directly tied to specific functions but are crucial for the effectiveness, robustness and maintainability of the system. The non-functional requirements listed here are derived from best practices of system design, internal discussions and meeting notes at GlucoModicum.

1. The system shall require no manual intervention during testing other than preparing the DUTs for testing, initiating the tests and ending the test session. The test system shall be designed in a way that minimizes the need for manual labor. There shall be no need for the user to re-plugging cables, instrument and test circuit configuration shall be set automatically by the software. The motivation for this requirement is to decrease the risk of human errors, to ensure repeatability, as well as to improve efficiency of the system.

2. The hardware setup shall be mechanically and electrically robust, with no loose or interference-prone components that could cause a decrease in signal quality. All paths shall be electrically isolated where necessary to ensure signal integrity. This is essential to increase reliability in the test results over multiple test cycles. It is particularly important when measuring nanoampere-level currents.

3. The system shall be intuitive and easy to operate. It shall include appropriate labels and visual indicators, as well as written instructions to guide the user to correct use. This requirement is highly important to minimize the risk of human error, as well as improving the overall efficiency of the system, while reducing training needs.

4. The test system shall be portable and capable of being duplicated with minimal variation in performance. This is necessary so that testing can be conducted in other locations (e.g. manufacturing facilities or regulatory labs). Additionally, the system design and documentation shall support replication or repair in the event of failure, to ensure continuity of test procedures and comparison with historical test data.

5. The test system shall be designed to require minimal maintenance and calibration no more than once per year. The motivation behind this requirement is increased reliability in the test results, as well as higher availability of the system. Instruments should have long calibration intervals and minimal moving parts.

3.4.2 Hardware Requirements

The hardware requirements for the ATS have been derived from three primary sources: regulatory requirements applicable to medical device verification, the requirements of the DUT, and the non-functional requirements previously defined. These requirements for the hardware of the system are set to ensure that the ATS meets the precision, scalability and flexibility demands that are necessary to support the ongoing development of the NIGM device at GlucoModicum.

1. The test system shall be capable of measuring currents in the range of 0-1 mA with an accuracy of ±10 nA.

2. The test system shall be capable of measuring voltages in the range of -12 V to 12 V with an accuracy of ±0.01 V.

3. The test system shall be capable of supplying constant current, and an arbitrary waveform in the range of 0-1 mA into the test circuit, with an accuracy of ±10 nA.

4. The test system shall be capable of supplying constant voltage in the range of 0-4 V into the test circuit, with an accuracy of ±0.001 V.

5. The test system shall support the application of a programmable load in the range of 0-100 kΩ, with a resolution of 10 Ω.

6. The test system shall be capable of measuring both current and voltage between any combination of electrodes of the DUT, including device ground.

7. The hardware architecture of the test system shall be designed to support an increased number of DUTs per test session without requiring significant hardware reconfiguration.

8. The test circuit and switching shall be modular, allowing easy controlling and reconfiguration, to accommodate expansion.

Requirements 1-4 stem directly from the system requirements of the NIGM device. In order to confidently verify that the DUT meets its measurement accuracy requirements, the ATS must comply with a Test Uncertainty Ratio (TUR) of at least 4:1. This means that the ATS needs to measure at minimum four times more precisely compared to the DUT's specified tolerance, to provide reliable results [68]. In addition to an accurate measurement instrument, a stable and precise current source is also essential for generating accurate reference signals to use in testing.

Requirement 5 specifies the range needed for the programmable resistive load to enable the ATS to simulate a variety of physiological and environmental conditions. This is essential to ensure that DUT performance is validated under both ideal and edge case scenarios. The requirements 6-8 are set to ensure that the test system will be easily adaptable to changing needs. As the development of the NIGM device is still ongoing, the possibility for modifications to the design, addition of new features, as well as changes to testing needs are possible. This might cause needs for reconfiguration of test logic, addition of new measurement channels and running of completely new test protocols. Consequently, the ATS must be scalable and modular to ensure long-term usefulness without the need for long maintenance periods or redevelopment efforts.

3.4.3 Software Requirements

Software requirements include those defined for the application used to control the test system. These have been derived from good software design principles as well as the design principles.

1. The software must be designed to be fully automated from the start of the first test sequence until the end of the last test sequence.
2. Instruments, fixtures, DUTs and tests must be configurable for all test sequences at once before the start of the test sequences.
3. The software must support running multiple tests in sequence without human intervention.
4. The software must support running tests for multiple DUTs in sequence without human intervention.
5. The software must be able to handle any errors, either by handling them independently or by notifying the user, after which operations must be possible to continue.
6. After all test sequences have finished the test results must be processed and a test report summarizing the test results must be possible to automatically create and save into a specified folder.
7. The UI of the software shall be simple and intuitive, possible to be used by a new user with minimal training.
8. The tests shall be easy to reconfigure, and new tests shall be easy to add to the software.
9. The software shall be designed to be modular and scalable to enable easy adaptation to changed needs.

3.5 Automated Test System Architecture

A previous version of a custom-made Automated Test System, that has served as a testing tool for GlucoModicum's R&D purposes, serves as inspiration and as a foundation for the Automated Test System designed and developed as part of this thesis. Development of this previous ATS is described in Alexander Björklund's Master's Thesis "Modular automated evaluation system for electrical safety and performance testing of a non-invasive glucose monitoring device" [74]. Especially the software controlling the ATS is widely reused, with only individual features improved, some new features added, and modifications made to ensure compatibility with the new hardware.

As highlighted in the previous chapter, the key requirements for the new ATS include accuracy, reliability, and scalability. Accuracy is achieved through high-precision instruments and low-noise test circuits, ensuring precise measurements. Reliability involves minimizing noise and variability in the test results, maximizing consistency. Scalability refers to the ease of modifying, expanding, and further developing the system as needs evolve over time. Several types of automated test systems exist, each characterized by different technological foundations, structural design and feature sets. During the Design Phase, a thorough analysis of different alternatives was conducted to determine whether the ATS should be fully custom-developed or built using commercially available components of a well-established instrumentation platform.

The advantages of a fully custom-developed system were found to be a significantly lower initial cost due to a wider range of instrument options and the possibility of utilizing existing instruments. Therefore, expenses would primarily be limited to custom cabling or PCB development for the test circuit. Additionally, it provides maximum flexibility in meeting precise design requirements without compromise. However, this approach has considerable disadvantages, including significantly longer development times due to the need to design each component of the system individually. Furthermore, there is a greater risk of encountering unforeseen technical challenges that could potentially delay development progress. Although maximum flexibility addresses initial requirements, it does not guarantee scalability for future modifications. Developing a scalable custom ATS is feasible but adds complexity to both the system design and implementation process. Moreover, custom-designed systems pose greater challenges regarding regulatory-compliant calibration and validation due to the necessity for developing custom validation protocols.

In contrast, an off-the-shelf approach using commercially available components significantly reduces development effort and enables a shorter project timeline. Although commercial components typically have a higher initial cost, they offer a much more scalable option and require less complexity in the validation process as they come with certified calibration documents. Additional benefits include extensive technical support and comprehensive documentation simplifying development and troubleshooting. The decision depends on balancing factors such as flexibility, cost, development efficiency, and regulatory compliance. Ultimately, it was found that rapid development on a tight schedule ensuring regulatory compliance was of higher

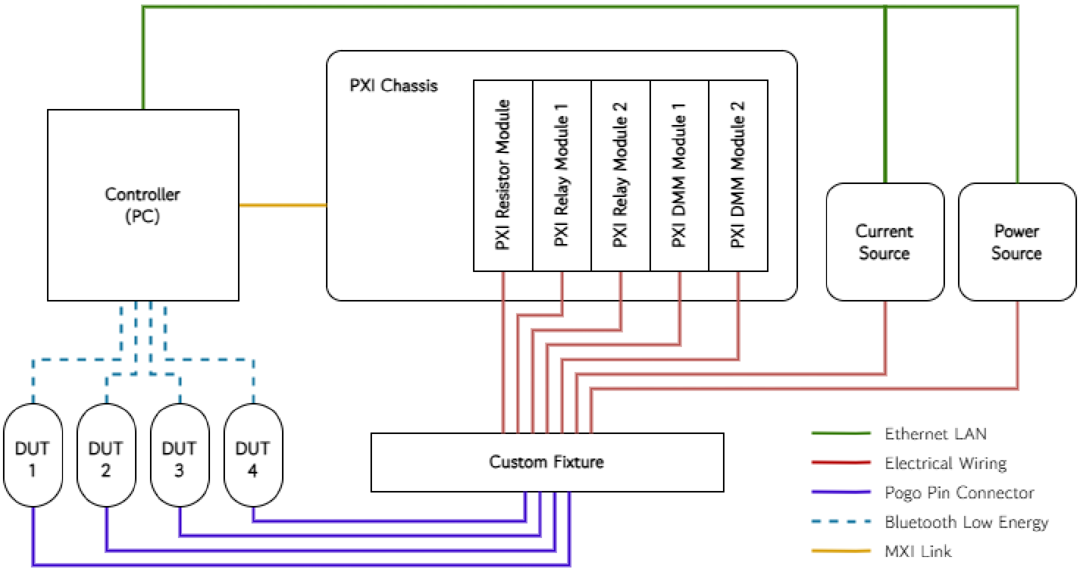


Figure 3: ATS architecture

priority, than absolute flexibility and minimized costs. It was also determined that due to the relative simplicity of the test scenarios, an off-the-shelf approach would be sufficient to fulfill all essential flexibility requirements.

The ATS developed in this project, illustrated in Figure 3, is based on the PCI eXtensions for Instrumentation (PXI) platform, originally introduced by National Instruments in 1997. The PXI-platform provides a modular and high-performance solution, combining flexibility through interchangeable modules, with proven reliability and accuracy from calibrated and certified instruments. The ATS consists of a chassis housing multiple test modules, including digital multimeters, relay modules and resistor modules, all interconnected within a custom-made fixture that serves as an interface for the DUT. The chassis is connected to a computer that controls the system by managing and executing the test sequences. Additionally, an external power source and an external current source are integrated into the ATS, connected to the controlling computer via Ethernet cable, as well as wired directly to the fixture. Although PXI-based options for these sources were available and represented a more optimal solution, external instruments were selected instead to reduce the costs by leveraging existing resources.

The custom fixture accommodates four separate DUTs, each mounted onto the fixture with custom designed mounts ensuring a secure connection through the pogo pins, minimizing external interference and signal loss. Additional features of the ATS include a connector designed for an external dummy battery, and dedicated ports to all the electrode terminals, either for signal extraction or connecting external mounts. The different modules and components are described in more detail in the following chapters.

3.5.1 PXI Instruments and Modules

The decision to adopt a PXI-based platform was motivated by the need to develop a modular system capable of easy modification and future expansion. The PXI-platform offers extensive flexibility in combining numerous types of modules to achieve a test system that is capable of performing accurate and precise testing. Its modular nature also facilitated building a prototype from cost-effective modules, providing proof-of-concept validation with minimal adjustment needed when transitioning to higher-end modules for an official version.

PXI modules are available from several vendors, all compatible through a standardized communication protocol. In addition to the original developer of the PXI-platform, National Instruments, other prominent vendors include Pickering Instruments and Keysight Technologies. Modules from these vendors were evaluated during the design phase, considering factors such as meeting specified requirements, cost efficiency, scalability for future needs, and module availability.

The 5-slot PXI-1033 chassis by National Instruments was selected as the primary platform due to its robustness, wide availability, and affordability. Despite being an older generation PXI-chassis, it provides more than sufficient functionality for this application. Its five module slots comfortably accommodate two digital multimeter (DMM) modules, two reed relay modules, and one resistor module. However, the limited number of slots means that an external power source and an external current source were needed for the prototype. These external instruments were chosen as they were available and previously integrated into GlucoModicum's legacy ATS, avoiding additional costs. In the official system, an integrated Source Measure Unit (SMU) module will replace the external sources. Furthermore, the prototype system is controlled by an external PC connected via MXI, while the official version is planned to be controlled by an integrated controller unit. Consequently, the chassis of the next iteration will require additional module slots compared to the PXI-1033.

Two separate DMM modules were utilized to allow simultaneous current and voltage measurements necessary for certain DUT test scenarios. The primary DMM chosen was the 7.5-digit PXI-4071, optimized for low-current measurements and offering the required precision needed. The secondary DMM module, a 6.5-digit PXI-4070, was selected for voltage measurements, which in this case have slightly less strict accuracy requirements, and the PXI-4070 provides a cost-effective alternative compared to an additional PXI-4071. Both DMM modules are of older generation, making them an optimal choice for the prototype as their long history ensures proven reliability, affordability, and widespread availability, especially as second-hand modules.

Test circuit configuration is managed by two 32 SPST Reed Relay modules (Pickering PXI 40-115-121), providing a total of 64 switches. Reed Relays were specifically selected due to their exceptionally low leakage currents, typically in the femtoampere (fA) or picoampere (pA) range, significantly enhancing measurement accuracy compared to solid-state switches used in the previous iterations of the ATS. A programmable resistor module (Pickering PXI 40-290-021), capable of providing a load ranging from 1 Ω to 131 kΩ, with a resolution of 1 Ω, and bypass functionality providing a close to non-existing load, was also incorporated. Together, these relay and resistor modules enable a wide range of test circuit configurations can be achieved. Additionally, four capacitors (10 µF, 100 µF, 1000 µF, and 7500 µF) were integrated into the test fixture. Activating the appropriate switches incorporates these capacitors into the circuit. These capacitor values were selected based on previous experimental data, ensuring the DUT is tested under realistic and varied capacitance conditions.

To be able to execute testing protocols capable of assessing the accuracy of the DUT, stable current and voltage signals are needed as reference. The external Keysight E36103B Power Source provides voltage outputs with a resolution of 1 mV and current resolution of 0.5 mA. The external Keithley 6221 Current Source outputs currents from 100 fA to 100 mA with a resolution of 100 fA and supports arbitrary waveform generation. This capability is essential for accurately recreating and simulating biosensor signals observed in real-world applications. This is a prerequisite for the current source to fulfill the requirements of the ATS, detailed in the third hardware requirement (Chapter 3.4.2). Both external instruments are connected and controlled via an Ethernet connection to the PC serving as the ATS control unit.

3.5.2 Test Fixture and DUT Interface

To enable a secure easy-to-use interface connection between the DUT and the test system, a custom fixture with four separate DUT slots has been developed. The fixture facilitates the wiring between the four DUT slots, the relay modules, the resistor module and the DMM modules, as well as to and from the external current and power sources. Wiring connections have been made utilizing Push-In Distribution Blocks by Phoenix Contact. This provided an easy and reliable solution for a test system prototype as wires can easily be rerouted and modified. For the official version, however, the wiring will instead be constructed with connectors attached to a custom PCB.

A DUT is connected to the test system by placing it in one of the slots, creating an electrical connection to all the pogo pins on the DUT. This enables easily repeatable testing between any of the electrodes on the DUT. The ATS has been designed to require minimal action when setting up a DUT for testing or switching one for the next. Still signal integrity had been maximized together with minimizing potential variables that could introduce errors. Encapsulating all wiring and connectors into a sealed 3D-printed fixture structure, except for the exposed pogo-pin connectors, ensures both safety for the user as well as minimizes the risk for unintentional modification to the fixture and circuit, while keeping it easy and quick to place and remove DUTs.

The fixture is freely standing only connected to the rest of the ATS via the cables, which can all be easily disconnected from the chassis. This makes the ATS portable, in case it needs to be used at other facilities for example. Scalability is also ensured by having a modular design for the ATS, ensuring that it can be easily expanded in the future if needed. More slots can be added by adding just a few more wire connections and 3D-printed DUT mounts, and the fixture itself can be sized up if new additions make the fixture container too small volume-wise.

To not have to hard-code all fixture configuration possibilities into the program software itself, each configuration possibility is defined in an XML file. This also ensures that the test system can be adapted to future needs much more easily. The XML file includes an element for each test configuration, where each configuration is given a name, an identifier and a list of the relevant relays with their IDs and their state to be set to. The list can be manually updated, or an XML editing program can be used. This also enables the same software to be used with multiple different hardware setups as different XML files can be created for each setup.

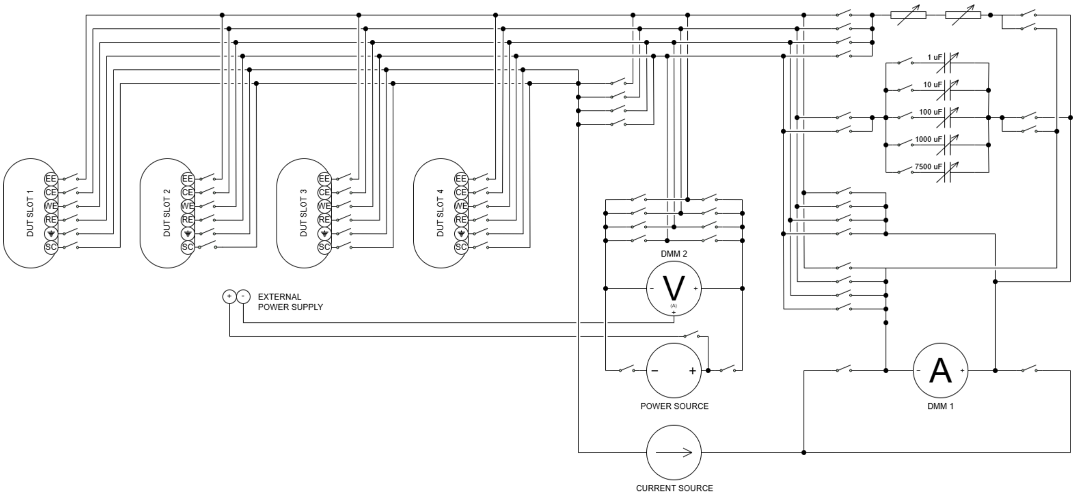


Figure 4: Wiring diagram of ATS

3.5.3 Control System

The control system is the computer running custom test software. It handles communication between the instruments and the DUTs. It also configures the instruments and the DUT, requests needed operations from them, and after the measurements are completed, it processes the measured data and creates automatic test reports. The control system has a wired connection to all the instruments and a wireless connection via Bluetooth to DUTs. The user interface (UI), through which the user controls the test system, is also a part of the control system.

For the ATS prototype the decision was made to utilize an external PC as a controller unit. This choice was mainly driven by the goal of keeping costs down, as well as due to the PXI-1033 Chassis having all its five slots occupied by the other modules, meaning a bigger chassis would have been needed to fit the integrated control module. A Lenovo ThinkStation P2 Tower PC was selected to control the ATS. It was fitted with a PCI-MXI Interface Card from National Instruments enabling a wired connection with an MXI cable between the PC and the PXI system. This link essentially makes the PXI components become extensions of the PC, and they can be controlled directly through vendor-provided software or with custom developed code that utilizes API calls.

The software controlling the ATS is a custom program written in C#. As stated earlier, it is widely based on the custom software developed for GlucoModicum's previous test system. Several components of the software were possible to reuse thanks to its modular design. The mainframe, test structures and DUT communication handling barely needed any modification, and the communication to the current and power sources were reused in completeness. New components for handling communications to the PXI modules were developed following the same modular design. The communication between the PC and the modules occurs directly, thus due to the system utilizing modules from both NI and Pickering, two different sets of API libraries had to be used.

The PXI modules do support the standardized Virtual Instrument Software Architecture (VISA) interface, which is a widely used industry standard API used for communication between computer and measurement instruments. The ATS uses VISA when communicating with the external instruments, however, it was decided not to use VISA for communication with the PXI modules. Despite being a possible option, it was deemed more suitable to go for the vendor-specific APIs from NI and Pickering. This choice was mainly motivated by the fact that VISA is not able to support all properties of the PXI modules, meaning that leveraging the full potential would not be possible. Using VISA to communicate can also be slightly slower than using a vendor-specific API, as every message needs to be translated to and from VISA format.

Despite the main program being developed in C#, the module responsible for automatically generating test reports based on the completed tests, is written in Python. When it comes to creating data plots and handling text documents, the libraries used with Python are far superior to what exists for C#.

3.6 Test Protocols

To be able to execute standardized testing in a repeatable manner, test protocols are created. Different test protocols are used for different types of testing, such as for testing against technical requirements, for quality and maintenance testing, and testing for debugging and development purposes. Each protocol consists of a set number of tests. Each test focuses on evaluating the performance and compliance of a specific property of the DUT. While tests are widely different in content, each test is built up of the same objects, which are a test circuit configuration, test parameters, steps to conduct the test, and one or multiple acceptance criteria.

This thesis focuses on a test protocol for pre-verification testing of the DUT. This means that the tests have been designed to evaluate compliance with the technical requirements, essentially testing whether the DUT functions as intended by the company. These tests can be utilized as evidence for device performance, and compared against, for example after mechanical safety testing or device modification. This way essential performance can be shown to not have been altered.

Each test in a test protocol has a short descriptive name providing the context of what the test is evaluating. The focus of the example test in Table 2 is "Extraction Current Output Accuracy". This means that the test is evaluating whether the DUT can output current during extraction with sufficient accuracy. The electrodes EE and CE are connected with a load in-between. This simulates the connection the electrodes of the biosensor would have when the device is placed on human subject, with the load being the natural resistance existing in human skin. Test parameters are defined so that the test can be repeated multiple times and that results can be compared, ensuring all controllable variables are identical. One test can be conducted with a set of different variable values to ensure the DUT is compliant in different situations. In the example test three different values for the load R has been selected as it is known that human skin can have different amounts of resistance due to facts like skin moisture level and contact pressure. Test procedure steps are crucial to be well defined. These instructions define the actions that either an ATS is set to perform automatically, or that a human can perform manually with separate instruments. Finally, clearly defined acceptance criteria indicate how it is determined whether a test passes or fails. Every

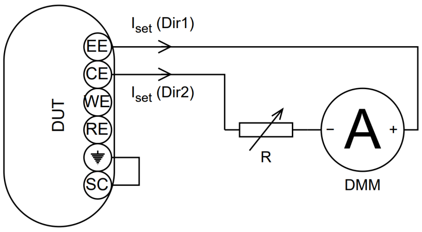
Description	Extraction current output accuracy
Configuration	
Parameters	$t = 5$ s $I_{out} = 191\ \mu$A $R = [0, 10, 100]$ kΩ $fs = 100$ S/s
Steps	1. For each value of R, initialize instruments and test circuit, and set load to R 2. Request extraction operation from DUT with duration t and current I_{out} 3. Measure current with the DMM at a sampling rate fs 4. Calculate an average of the current measured by the DMM over the time period t
Acceptance Criteria	For each value of R, the average current measured by DMM shall be I_{out} ±2 %

Table 2: Example of a test conducted by the ATS

individual test is configured in the software of the ATS. When running a test protocol, the associated tests are selected and set up as sequences within the protocol. Each sequence corresponds to a specific test, with one or multiple procedures, corresponding to each execution of the test, depending on the amount of test parameter combinations. Each test procedure then consists of one or multiple test operations. An operation is one action performed by the test system in the form of a message request to one of the test system components. Operations can be configuration of the test circuit,

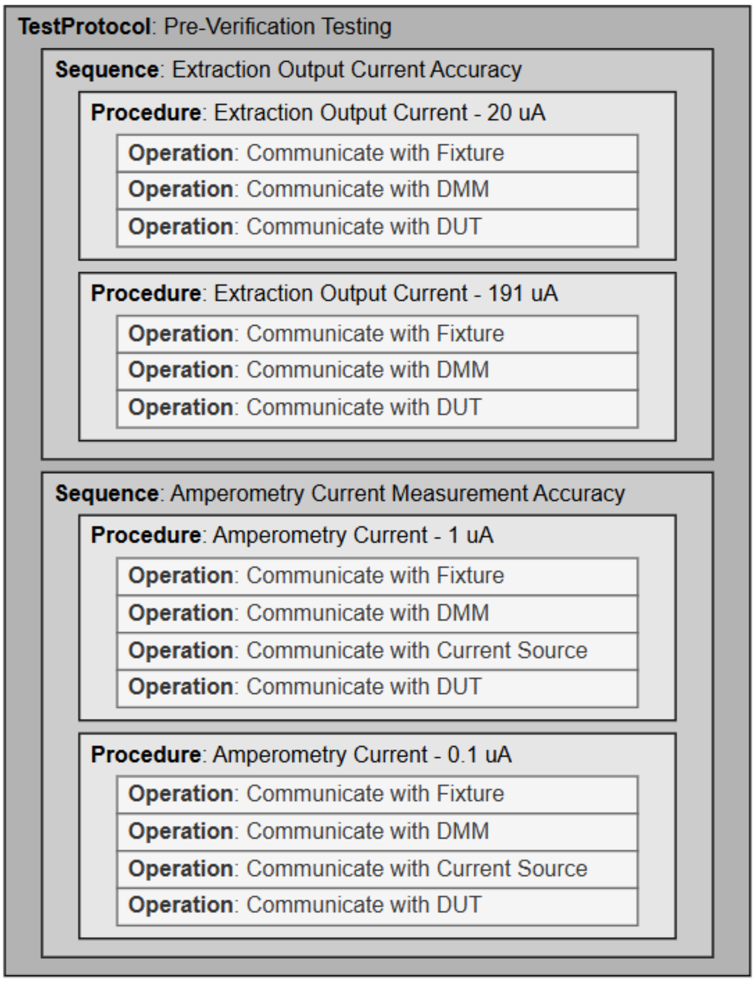


Figure 5: Hierarchy of sequences, procedures and operations of a test

initialization of a measurement by the DMM, or a request to the DUT to perform a specific DUT operation. Figure 5 shows the structure of how sequences, procedures and operations exist within one another inside the test protocol.

From a software point of view, once the test protocol has been selected, each sequence is set up by initializing procedures with operations. Then for each procedure one operation is executed at a time. When running the test in Table 2 the first operation is configuring the fixture. This begins with getting the correct relay configuration from the configurations XML file based on the configuration ID defined as a test parameter. The relay configuration is then translated into two commands, one for each relay module. These commands are then sent to the relay modules, as well as a command for setting the needed load to the resistor module based on the test parameter. A response is awaited from the modules to confirm that the operation was successful.

Next operation is initializing the DMM preparing it for measurement. The parameters defined in the test are sent to the DMM module and a trigger operation is configured. The trigger operation ensures that the measurement is not initiated before the DUT operation commences. For tests where DUT operation is not needed, setting up the trigger can be omitted and instead measurement is initiated immediately. Initiating measurement also triggers a task that will wait for the measurement to be completed, after which it will get the measurement data from the DMM.

When the DMM has been initialized and is ready to measure the current in the test circuit, an operation sending a request to the DUT is processed. This request includes parameters specified in the test, such as DUT operation type, operation length in seconds, and for example output current. In the example case the DUT operation is an extraction operation with the output current 191 µA and an operation time of t = 5 s.

Other operation types include commands to start and stop the current source and the power source, or a so-called "wait operation" that creates a time delay between two other operations. This is useful when a current signal needs time to stabilize before it should be measured.

After all operations have been processed, the system then waits for the event `OnProcedureCompleted` to be raised before it proceeds to the next procedure. This event is raised when it is confirmed that all measurements have been completed and measurement data has been retrieved. The system runs through all the procedures within the sequence before it processes the measurement data into a test result. When the last procedure of the sequence finishes, the event `DriverFinished` is raised initiating results processing. The data for each procedure is processed and compared to the acceptance criteria. Even if just one procedure result is a FAIL, the whole test will produce a FAIL result.

3.6.1 Protocol for Reliability Testing

To comprehensively assess the reliability of the ATS a series of nine distinct tests were selected. Each test was carefully chosen to highlight and evaluate different capabilities of the ATS by targeting various essential properties of the DUT. The selected tests collectively ensure thorough coverage of the DUT's performance characteristics, and the ATS needs to be capable of producing consistent results for each of the types of tests.

The specific details for each test, including test circuit configuration, parameters and acceptance criteria are documented and presented in Tables 3–11. These parameters are based directly on the Outgoing Quality Control (OQC) tests currently used by GlucoModicum, ensuring alignment with established quality procedures. In most of the tests a measurement period of t =5 s is used, as it has been determined via previous testing to be long enough for the device to produce stable values. While many different currents, voltages and loads are used to ensure accurate results in different scenarios, an extraction current specifically of size 191 µA is often used. This is simply because it has been found to work well with the recent version of the biosensor. Similarly, loads of size 1 kΩ and 10 kΩ are more often used as they have been found to simulate the resistance in human skin. The structure of the tests follows the same format as described in the previous chapter (Chapter 3.6).

Tests T-1 through T-4 (presented in Tables 3–6) aim to verify that the stability and accuracy of the current source within the DUT by ensuring consistent output of

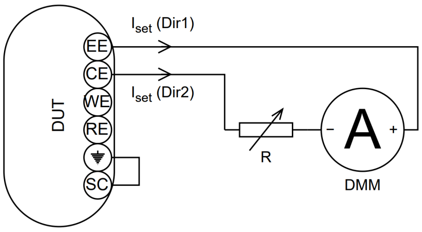
Test ID and Name	**T-1:** Extraction Output Accuracy with Varying Currents
Description	Assessing the accuracy of the output current during an extraction operation with varying set currents
Configuration	
Parameters	$t = 5\,\mathrm{s}$ $I_{out} = [20, 191, 400, 600, 1000]\ \mu\mathrm{A}$ $R = 10\,\mathrm{k}\Omega$
Acceptance Criteria	For each value of I_{out}, the average current measured by DMM over the measurement period t shall be $I_{out} \pm 2\,\%$

Table 3: Test T-1 of the test protocol for reliability testing

extraction currents within the specified limits. Additionally, these tests are designed to assess the voltage measurement accuracy of the potentiostat integrated into the DUT. Including these tests in the repeatability test protocol demonstrates the ATS's ability to configure the resistive loads in the test circuit and to execute precise measurements of currents ranging from 50-1000 µA and voltages within 11-13 V range.

To further demonstrate the ATS's capability at lower measurement scales, tests T-5 (Table 7) and T-6 (Table 8) have been included in the protocol. These tests evaluate the ATS's accuracy when measuring small currents ranging from 1 µA down to 100 nA and voltages at millivolt levels. Test T-5 (Table 7) specifically demonstrates that the current source implemented into the ATS is able to deliver accurate current outputs on microampere and nanoampere ranges.

Many test cases within the OQC test protocol require measurements of stable current and voltage signals over extended durations, allowing for longer sampling periods to get more precise data acquisitions. However, it is essential for the ATS to also be able to measure dynamic signals, such as the triangular voltage waveform used in cyclic voltammetry. Tests T-7 (Table 9) and T-8 (Table 10) specifically assesses this capability of the ATS. It is done utilizing an area-under-curve analysis to compare measurements from the DUT and ATS. As with earlier tests, resistive and capacitive loads are used to simulate skin impedance characteristics, ensuring that the DUT consistently performs under realistic operating conditions.

Test T-9 (Table 11) demonstrates low voltage measurement accuracy of the ATS, a critical aspect when testing the DUT's open-circuit potential operations. At the same time Test T-9 confirms the ATS is capable of producing stable reference voltages on

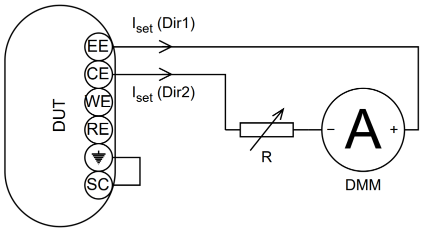
Test ID and Name	**T-2:** Extraction Output Accuracy with Varying Loads
Description	Assessing the accuracy of the output current during an extraction operation with varying loads in the test circuit
Configuration	
Parameters	$t = 5\,\mathrm{s}$ $I_{out} = 191\,\mu\mathrm{A}$ $R = [0, 10, 100]\,\mathrm{k}\Omega$
Acceptance Criteria	For each value of R, the average current measured by DMM over the measurement period t shall be $I_{out} \pm 2\,\%$

Table 4: Test T-2 of the test protocol for reliability testing

the millivolt scale.

Preliminary expectations suggest that all tests will meet acceptance criteria with clear margin, as the OQC tests, that these tests build upon, have been designed to pass with the legacy ATS found to inherit a higher level of identifiable noise and inaccuracies. Given the precision and accuracy specification of the PXI-based ATS, the acceptance criteria for the tests are expected to be met much easier and with greater consistency. Minor variations and deviations within the acceptance threshold are expected, however, these should remain well below the defined acceptance limits and be clearly lower in comparison with results produced by the legacy ATS.

Specifically, standard deviation across repeated tests is anticipated to be extremely low, indicating high consistency and reliable repeatability of the system. Test T-3 (Table 5) is expected to yield particularly stable and consistent results with minimal variability, given the relatively large voltage measurements typically exhibit reduced noise sensitivity. Test T-8 (Table 10), on the other hand, is anticipated to show slightly higher variability due to the capacitive load involved in the test circuit.

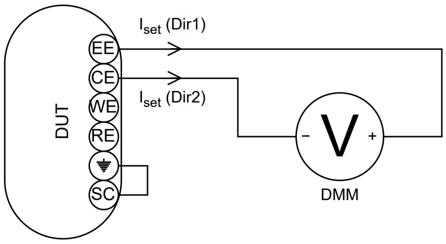
Test ID and Name	**T-3:** Extraction Output Voltage Compliance
Description	Assessing that the compliance voltage during an extraction operation is 12 V and that the DUT measures it correctly
Configuration	
Parameters	$t = 5$ s $I_{out} = 191\,\mu A$
Acceptance Criteria	The average voltage measured by the DMM and the average voltage measured by the DUT over the measurement period t is 12 V ±1 V

Table 5: Test T-3 of the test protocol for reliability testing

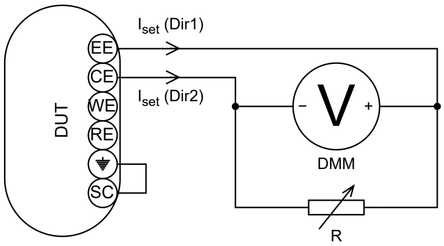
Test ID and Name	**T-4:** Extraction Voltage Measurement Accuracy
Description	Assessing voltage measurement accuracy during an extraction operation with varying set currents
Configuration	
Parameters	$t = 5$ s $I_{out} = 191\,\mu A$ $R = [1, 10, 25, 38, 60]\,k\Omega$
Acceptance Criteria	For each value of R, the average voltage outputted and measured by the DUT over period t, is equal to the average voltage measured by the DMM over the same period ±0.1 V

Table 6: Test T-4 of the test protocol for reliability testing

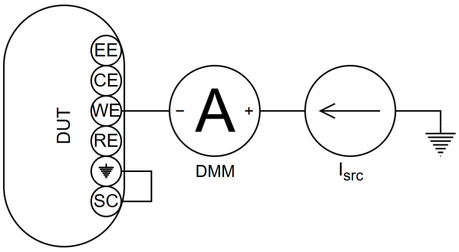
Test ID and Name	**T-5:** Amperometry Current Measurement Accuracy
Description	Assessing the accuracy of the current measurement during an amperometry operation with varying reference currents
Configuration	
Parameters	$t = 5\,\mathrm{s}$ $V_{bias} = 100\,\mathrm{mV}$ $I_{src} = [10, 1, 0.1, 0.01]\ \mu\mathrm{A}$
Acceptance Criteria	For each value of I_{src}, the average current measured by the DUT over the period t is equal to the average current measured by the DMM over the same period ±3 % or ±5 nA

Table 7: Test T-5 of the test protocol for reliability testing

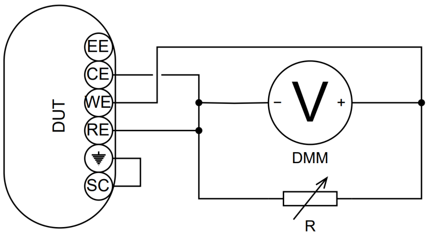
Test ID and Name	**T-6:** Amperometry Bias Voltage Accuracy
Description	Assessing that the set Bias Voltage during an amperometry operation remains within a specified range
Configuration	
Parameters	$t = 5\,\mathrm{s}$ $V_{bias} = [+550, +100, +5, 0, -5, -100, -550]\ \mathrm{mV}$ $R = 10\,\mathrm{k\Omega}$
Acceptance Criteria	For each value of V_{bias}, the average voltage measured by the DMM over the measurement period t, is equal to V_{bias} ±2 % or ±1 mV

Table 8: Test T-6 of the test protocol for reliability testing

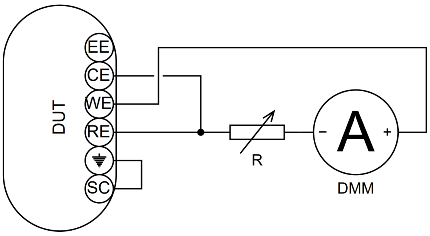
Test ID and Name	**T-7:** CV Current Measurement Accuracy with Resistive Loads
Description	Assessing current measurement accuracy during a cyclic voltammetry operation with resistive loads in the test circuit
Configuration	
Parameters	$V_0 = 200\,\text{mV}$ $V_A = -600\,\text{mV}$ $V_B = 600\,\text{mV}$ $n = 1$ $R = [1, 10]\ \text{k}\Omega$
Acceptance Criteria	For each value of R, the Area Under Curve (AUC) of the current measurements by the DUT during the CV operation (V_0 as start voltage, V_A as vertex A, V_B as vertex B and n as number of cycles) is equal to the AUC of the current measurements by the DMM ±8 %

Table 9: Test T-7 of the test protocol for reliability testing

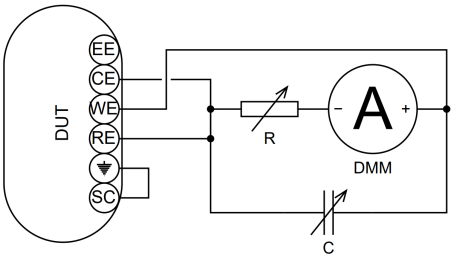
Test ID and Name	**T-8:** CV Current Measurement Accuracy with Resistive and Capacitive Loads
Description	Assessing current measurement accuracy during a cyclic voltammetry operation with both resistive and capacitive loads in the test circuit
Configuration	
Parameters	$V_0 = 0\,\text{mV}$ $V_A = -600\,\text{mV}$ $V_B = 600\,\text{mV}$ $n = 1$ $R = [1, 10]\ \text{k}\Omega$ $C = 100\,\mu\text{F}$
Acceptance Criteria	For each value of R, the Area Under Curve (AUC) of the current measurements by the DUT during the CV operation (V_0 as start voltage, V_A as vertex A, V_B as vertex B and n as number of cycles) is equal to the AUC of the current measurements by the DMM ±8 %

Table 10: Test T-8 of the test protocol for reliability testing

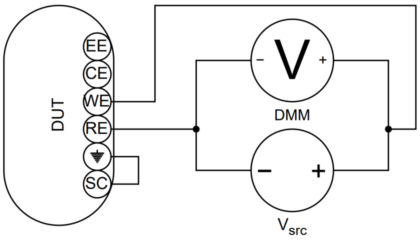
Test ID and Name	**T-9:** OCP Voltage Measurement Accuracy
Description	Assessing the voltage measurement accuracy during an open-circuit potential operation with varying reference voltages
Configuration	
Parameters	$t = 5$ s $V_{src} =$ $[+500, +100, +10, +1, -1, -10, -100, -500]$ mV
Acceptance Criteria	For each value of V_{src}, the average voltage measured by the DUT over the period t is equal to the average voltage measured by the DMM over the same period ±0.2 % or ±0.1 mV

Table 11: Test T-9 of the test protocol for reliability testing

4 Results

This chapter presents the results obtained from tests conducted on the ATS. The primary objective of these tests was to validate if the developed test system fulfills the criteria of being a reliable test system. Two different focus areas for system evaluation were chosen. First, the signal integrity and accuracy of measurements within the test system were analyzed as part of a "noise analysis". This aimed at showing that the system can perform measurements accurate enough to meet the requirements. Secondly, a "repeatability test" was performed to assess the rigidity of the test system ensuring that it is capable of performing repeated testing and able to deliver consistent test results. Both metrics together can be argued to display the reliability of the test system.

4.1 Measurement Accuracy

A key requirement for the ATS is the capability to accurately evaluate the measurement accuracy of the NIGM device under test. To effectively determine the accuracy and performance of the DUTs, the measurement accuracy of the ATS itself must be significantly superior, ensuring that any errors observed during testing originate from the DUT rather than the testing equipment. Thus, understanding the intrinsic noise and baseline measurement accuracy of the ATS is crucial to confidently interpret test results and verify DUT performance.

To characterize the baseline measurement accuracy and the level of induced noise, a comprehensive noise analysis was conducted. Specifically, short-circuit current measurements were performed to establish baseline readings for leakage current, inherent system noise, and measurement stability of the ATS. By performing measurements in a short-circuited configuration, any current detected by the digital multimeter can be attributed primarily to the internal noise and leakage of the ATS. This provides a clear indication of the test system's minimum detectable signal and accuracy limits.

The testing procedure involved short-circuiting the measurement inputs and subsequently capturing the output data using a high-precision DMM. For reference, baseline measurements were also conducted with the DMM disconnected from the test circuit, ensuring that the recorded data accurately reflected only the internal instrument noise and not external circuit contributions. Several test sequences were executed, varying both the measurement duration and load to thoroughly evaluate the ATS across a range of realistic test scenarios. The parameters and conditions used for these sequences are detailed in Table 12. The data collected from the ten measurement sequences that were performed demonstrated consistently low noise levels and minimal variation across different test conditions. Based on this consistency, further testing was deemed at this point unnecessary.

The measurement durations were carefully selected to cover various testing scenarios: short measurements (45 seconds), operational duration measurements (7 minutes), and extended measurements (45 minutes). Short-duration tests are particularly valuable as they produce stable results quickly, facilitating frequent test

	Parameters		Results		
Sequence	R_{Set}	Measurement Time	Mean Current	Peak-to-Peak	Standard Deviation
1	0 Ω	45 s	0.775 nA	0.106 nA	0.025 nA
2	0 Ω	45 s	0.754 nA	0.112 nA	0.027 nA
3	1 kΩ	45 s	0.929 nA	0.101 nA	0.019 nA
4	1 kΩ	45 s	0.864 nA	0.063 nA	0.014 nA
5	10 kΩ	45 s	1.815 nA	0.082 nA	0.022 nA
6	10 kΩ	45 s	1.613 nA	0.308 nA	0.083 nA
7	0 Ω	7 min	0.811 nA	0.275 nA	0.064 nA
8	1 kΩ	7 min	0.976 nA	0.403 nA	0.089 nA
9	10 kΩ	7 min	1.635 nA	0.379 nA	0.080 nA
10	10 kΩ	45 min	1.632 nA	0.862 nA	0.161 nA
reference	N/A	45 min	0.008 nA	0.151 nA	0.005 nA

Table 12: Parameters and results of short-circuit noise analysis

repetitions and rapid diagnostics. Operational duration tests correspond closely to typical operational cycles of the DUT under realistic conditions. Matching test durations with real-life operational cycles enhances the reliability and relevance of the verification results. Extended-duration tests are essential for detecting slow drifts or cumulative effects that shorter tests might overlook. Although tests of this duration are not commonly performed due to their time-consuming nature, they provide insight into the overall stability and performance of the ATS.

The test fixture was evaluated using a variety of resistive loads to simulate diverse test scenarios. In addition to a 0 Ω load the load values 1 kΩ and 10 kΩ were chosen, as they reflect realistic resistance levels representative of human skin impedance. These specific loads were selected as they are industry-standard values based on prior empirical research [75] commonly used in biomedical device testing. The 0 Ω load represents an idealized short-circuit scenario actively used in numerous test scenarios during verification of the DUT when no resistive load is needed. The 1 kΩ load represents typical low-resistance human skin conditions encountered during optimal operation of the DUT, while the 10 kΩ load simulates higher impedance scenarios which are also common during normal operation of the DUT, and might be caused by variations in skin condition, such as hydration levels. Conducting evaluation with a variety of loads offers insight into how the ATS performs under a broader range of practical conditions, ensuring comprehensive verification of its performance. This is crucial to confirm the capability of the ATS to reliably assess the DUT across multiple realistic operational scenarios.

Simple data analysis was conducted on the measurement data that was collected from the short-circuit noise analysis to evaluate the performance of the ATS. All measured currents across the different measurement sequences, presented in Table 12 are also visualized in Figure 6. It is clearly visible that the measured current, in other words, the identified noise, stays consistently under 2 nA for all sequences. A reference value was measured with the DMM disconnected from the fixture to set a baseline of the noise-induced by the DMM. This baseline was found to be 0.008 nA. It is clearly visible that there is a great correlation between the size of the resistive load and the size of the noise, where a higher load equals a higher noise level.

The statistical range of the measured values, also known as the peak-to-peak value, is found to be less than 0.5 nA for nine out of the ten sequences. For the last sequence, where measurement time was significantly longer, the values fluctuated within a range of 0.86 nA. Clear correlation between measurement time and size of the peak-to-peak range is identified. This is expected, as if measurement periods are longer the chance of more diverse values increases. This indicates that, even though measurements further from the mean can be identified when performing long measurements, such are uncommon and despite that the ATS is able to measure with great accuracy.

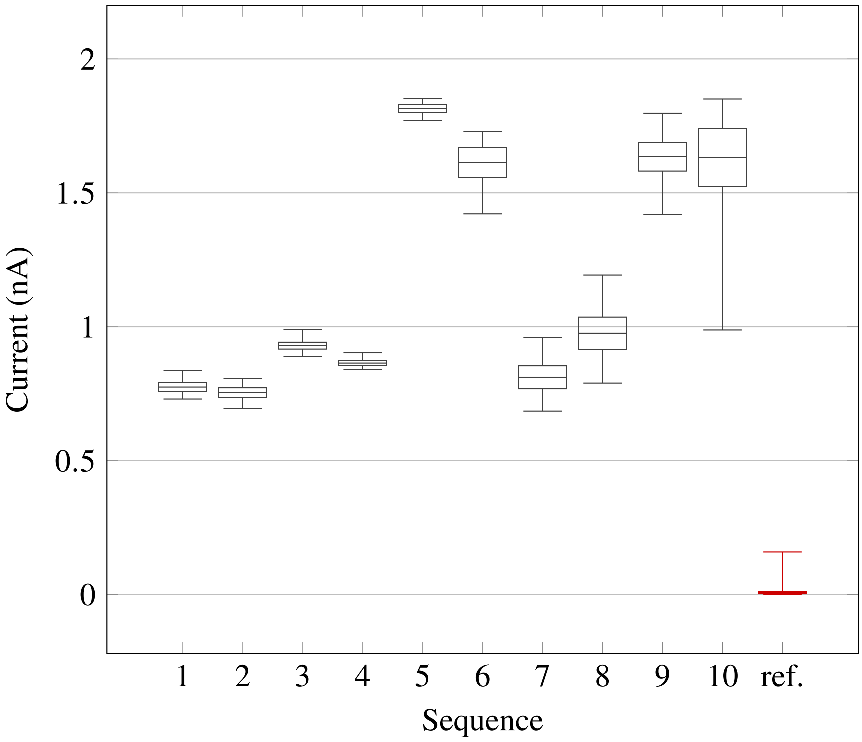


Figure 6: Short-circuit noise analysis results

Standard Deviation of the measured values is measured to be relatively low. For measurements of less than a minute the average standard deviation is 0.02 nA, increasing to 0.08 nA for measurements of around 7 minutes. For the measurement lasting 45 minutes the standard deviation was 0.16 nA. However, as only one singular measurement of this length was performed, no conclusions can be drawn confidently.

In addition to the short-circuit analysis, an input-output current comparison test was also performed. This type of test is a useful tool to evaluate how much a current signal is distorted when passing through a system. The current source was used to input different currents into the test system circuit and the signal was measured with the DMM when it exited the system. A set of measurements were also collected by measuring current flow between the current source and the test system circuit. This data set was used as reference to set the baseline for what the accuracy level of the current source is.

Sequence	I_{set}	R_{Set}	Statistical Range (IN)	Statistical Range (OUT)	Standard Deviation (IN)	Standard Deviation (OUT)
1	0.01 µA	0 Ω	0.295 %	0.834 %	0.005 nA	0.008 nA
2	0.01 µA	1 kΩ	0.298 %	0.554 %	0.005 nA	0.006 nA
3	0.01 µA	10 kΩ	0.291 %	0.327 %	0.005 nA	0.006 nA
4	0.1 µA	0 Ω	0.026 %	0.038 %	0.005 nA	0.006 nA
5	0.1 µA	1 kΩ	0.027 %	0.061 %	0.005 nA	0.007 nA
6	0.1 µA	10 kΩ	0.033 %	0.045 %	0.004 nA	0.007 nA
7	1 µA	0 Ω	0.004 %	0.004 %	0.006 nA	0.007 nA
8	1 µA	1 kΩ	0.006 %	0.008 %	0.007 nA	0.008 nA
9	1 µA	10 kΩ	0.003 %	0.007 %	0.006 nA	0.008 nA
10	20 µA	0 Ω	0.010 %	0.010 %	0.381 nA	0.409 nA
11	20 µA	1 kΩ	0.012 %	0.010 %	0.393 nA	0.441 nA
12	20 µA	10 kΩ	0.009 %	0.010 %	0.404 nA	0.386 nA
13	191 µA	0 Ω	0.011 %	0.010 %	3.698 nA	4.004 nA
14	191 µA	1 kΩ	0.012 %	0.012 %	3.819 nA	4.130 nA
15	191 µA	10 kΩ	0.013 %	0.009 %	3.889 nA	3.797 nA
16	641 µA	0 Ω	0.005 %	0.005 %	4.931 nA	4.671 nA
17	641 µA	1 kΩ	0.003 %	0.003 %	4.426 nA	4.491 nA
18	641 µA	10 kΩ	0.004 %	0.005 %	4.457 nA	4.943 nA
19	1000 µA	0 Ω	0.024 %	0.021 %	36.02 nA	41.12 nA
20	1000 µA	1 kΩ	0.018 %	0.019 %	34.02 nA	34.09 nA
21	1000 µA	10 kΩ	0.019 %	0.018 %	38.98 nA	35.81 nA

Table 13: Parameters and results of current input-output analysis

Comparing the two data sets made it possible to analyze how much the quality of the signal is degraded by the test system circuit. The two different measurements (before passing through the test fixture and after it) were executed separately. This was done to be able to utilize the same DMM for both measurements, eliminating device calibration differences as potential source of error. Furthermore, performing the test in this way ensured that the second current measurement was not affected by the first current measurement.

Similarly to the short-circuit test, a set of varying current and load values were chosen to represent the whole scale of the usage area. This way it can be confirmed that the ATS is able to produce accurate and stable results in all possible scenarios. As the range of current stretches from nanoamperes up to milliamperes, the errors and measured noises are given both as absolute values and relative values. Table 13 shows the parameters that were used, as well as the measured results.

4.2 Repeatability Testing

The intended use of the ATS involves executing multiple test sequences repeatedly over extended periods, requiring stable and reliable results each time. As described in Chapter 2.3, reliability in testing requires that repeated testing produces consistent results. Test results from different sequences must be comparable and free from any variability introduced by the ATS itself. To empirically confirm the repeatability of the ATS, a structured repeatability analysis was conducted, focusing on both short-term and long-term measurement stability.

The overall repeatability testing involved executing a predefined test sequence, as detailed in Chapter 3.6.1, a total of 20 times. These repetitions were performed in two distinct sessions on separate days to effectively evaluate both intra-day (short-term) and inter-day (long-term) repeatability. This number of repetitions was carefully selected as a balance between statistical evidence and practical feasibility. Performing 10 repetitions per session ensures that the measurement consistency is clearly established and allowing observation of short-term variability. Conducting tests across two separate days, on the other hand, confirms stability against potential session-to-session variation. All tests were performed in similar environmental conditions to minimize it being the source of possible external variation. This approach with running 20 repetitions over two days was considered sufficient to clearly demonstrate repeatability of the ATS while avoiding unnecessary complexity.

Test ID	Test Description	Number of Passed	Number of Failed
T-1	Extraction Output Accuracy with Varying Currents	20	0
T-2	Extraction Output Accuracy with Varying Loads	20	0
T-3	Extraction Output Voltage Compliance	20	0
T-4	Extraction Voltage Measurement Accuracy	20	0
T-5	Amperometry Current Measurement Accuracy	20	0
T-6	Amperometry Bias Voltage Accuracy	20	0
T-7	CV Current Measurement Accuracy with Resistive Loads	20	0
T-8	CV Current Measurement Accuracy with Resistive and Capacitive Loads	20	0
T-9	OCP Voltage Measurement Accuracy	20	0

Table 14: Test results of repeatability analysis

Table 14 summarizes the results of the tests conducted on a DUT that was previously confirmed to function as intended and meeting all its requirements. It clearly demonstrates that all measurement results produced by the ATS consistently met the predefined acceptance criteria for the DUT. From analyzing the results, it is clear that the variability in test results across multiple sessions is negligible compared to the precision that is expected from the ATS. This underscores the system's strong repeatability.

	Absolute Error			Relative Error		
Test ID	Mean	Peak-to-Peak	Std Dev	Mean	Peak-to-Peak	Std Dev
T-1	0.251 µA	0.013 µA	0.001 µA	0.10 %	0.02 %	0.01 %
T-2	-	-	-	0.10 %	0.03 %	0.01 %
T-3	0.090 V	< 0.001 V	< 0.001 V	-	-	-
T-4	0.024 V	0.031 V	0.001 V	-	-	-
T-5	0.006 µA	0.005 µA	0.001 µA	1.11 %	1.75 %	0.51 %
T-6	0.668 mV	0.066 mV	0.012 mV	4.69 %	0.37 %	0.12 %
T-7	-	-	-	3.96 %	0.10 %	< 0.01 %
T-8	-	-	-	4.12 %	2.85 %	0.82 %
T-9	0.080 mV	0.034 mV	0.001 mV	0.30 %	0.48 %	0.15 %

Table 15: Errors in measurements during repeatability analysis

5 Discussion

The primary objective of this thesis was to design and develop a scalable, reliable and highly accurate automated test system for pre-verification purposes of a NIGM device. As stated in Chapter 3.4, achieving this objective required careful definition of a series of system requirements, and well-designed test protocols (Chapter 3.6) to validate these requirements.

The results obtained from extensive testing of the ATS, presented in detail in the previous chapter, clearly demonstrate that the ATS meets the specified needs and requirements. When closely analyzing the results, it is clearly visible that the test system is able to perform measurements exceeding the DUT's accuracy requirements, enabling confident assessment of the DUT's performance. The achieved nanoampere-level standard deviation is more than sufficient for assessing the accuracy of a DUT, surpassing the TUR guideline of 4:1, commonly accepted as the best practice in measurement accuracy assessments.

The observed results, depicted in Figure 6 reveal an expected pattern, where the range of measured values increases proportionally with longer measurement durations. This pattern was anticipated, as any measured analog signal inherently incorporates some amounts of randomness, meaning that longer measurement intervals naturally accumulate more random fluctuations. Nonetheless, even during extended measurement intervals, measurement accuracy consistently remained well within acceptable limits for assessing the NIGM device with confidence.

Significant improvements in measurement accuracy and reduced noise levels were evident compared to the previous generation ATS. Primary contributors to this enhanced performance include employing a more accurate DMM, selecting relays with lower current leakage, and implementing a more optimized overall design of the system. These improvements underscore the effectiveness of the PXI-based design choice, which provides robust modularity and scalability in addition to increased measurement precision.

Another key factor of the ATS assessed through testing was the ability of the ATS to produce reliable results over multiple repeated test cycles. The repeatability evaluation involved conducting nine distinctive tests repeated across 20 test cycles. The data indicates a strong ability to produce reliable test results repeatedly, with all 180 tests successfully passing with consistent results. Certain tests, such as T-4 and T-7, generated nearly identical outcomes every cycle. Even the least coherent test, T-8, maintained a relative peak-to-peak of below 3 % and a relative standard deviation below 1 %, which falls within acceptable standards for reliability.

Addressing the research question "How can accurate testing of a non-invasive glucose monitoring device, be ensured without the use of real or artificial skin?", the findings of this thesis provide clear evidence that utilizing an automated test system designed to simulate a biosensor effectively addresses this challenge. By carefully designing the test circuit to replicate the electrical characteristics and responses of human skin through resistive and capacitive elements, the ATS enables precise, consistent, and reliable evaluation of individual subsystems and specific properties of the DUT. Thus, the ATS serves as a robust, effective solution for accurately assessing the performance and safety of non-invasive glucose monitoring devices in an in vitro manner without utilizing human or artificial skin.

5.1 Scalability, Modularity and Maintainability

In addition to accuracy and reliability, scalability, modularity and maintainability also played a significant role in the design of the ATS. To ensure long term usefulness of such a test system, and to ensure the ability for it to be adapted to changing needs, scalability, modularity and maintainability has been a top priority throughout the development of it. Scalability, as mentioned in Chapter 2.4 means that the system can handle growth in all aspects, whether it is increased number of DUTs, test cases, test cycles or instruments. The biggest contribution to scalability is a modular design approach both regarding software as well as hardware.

The ATS prototype was developed to include four separate DUT slots to be able to perform serialized testing. Four slots were selected as it was the maximum number of slots that was possible to include with the number of relays available. Thanks to the design of the test system being modular, the number of slots can easily be increased in the future. As the testing of each DUT is performed in series there is essentially no upper limit to how many slots could be implemented. An ideal number of DUT slots for an ATS used to conduct serialized pre-verification testing is challenging to determine, as it depends on processes and procedures of the company using it. For GlucoModicum the main benefit of serialized multi-DUT testing is to be able to run extensive testing overnight, utilizing the quiet hours for data gathering. How many DUTs can be tested during a night is inversely proportional to the length of the test protocol used. The length of the OQC Protocol that has been used when planning the requirements of the ATS is approximately 30 min per DUT. The length of the protocol used for the Repeatability Test, on the other hand, is only around five minutes. Four individual slots, however, were found to be a suitable number of slots as it does not add substantially to the complexity of the ATS architecture but still is enough to prove the multi-device capabilities of the ATS. At the same time, four slots are considered to be a widely popular number of slots for parallel testing as it can decrease test cost per-device with up to 50 % while keeping system complexity manageable [76]. With four slots the ATS is ready to be further developed at a later stage into supporting parallel testing.

The steps to increase the number of DUTs the system can support include adding an additional relay module or upgrading the existing ones to modules with more relays. Six additional relays are needed for each new slot added. Alternatively, a multiplexer module could be utilized to handle the switching between the fixture and the DUTs. In addition to the hardware needed for the upgrade, a small modification is needed to the software. The number of slots available needs to be updated within the Hardware Configuration class. With these changes the test system can easily be upgraded from supporting four DUTs to supporting more devices.

Modifying the test system to support parallel testing is slightly more complex as it includes adding more instruments, as well as modifying the software classes that handle communication to the instruments. Due to almost all test cases needing a DMM to measure a signal, and due to the significant cost of DMM modules, the benefits of parallel testing were deemed not to be worth the cost. New modules can, however, be added to the system by wiring them up to the fixture, adding them to the XML instrument configuration file and updating the relay configuration XML file. Existing instrument modules can also easily be replaced by higher-performing units just by updating the device parameters — including the module address, type and identifier — in the instrument specification XML file `InstrumentDefinitions.xml`.

As vendor-specific API libraries were utilized in communication with the modules, adding or exchanging a module of the same vendor as an existing module requires substantially less modification to the software in comparison to if a module of a different vendor would be used. The operation principles of PXI modules seem, however, to be similar in kind. This ensures that, despite different libraries being used, the class structures can be maintained.

New tests and test protocols can easily be added to the system by utilizing the modular design of the test system software. All tests are hardcoded into the software and include the test procedures and operations, tested parameters, results processing functions, acceptance criteria, as well as metadata for the test reports. Future additions to the test system include a more flexible way to define and design tests either by adding test controls to the UI or by utilizing an XML-based test file design.

Maintainability is ensured by extensive documentation, extensive testing and by a modular design. If some part of the test system is suspected of not performing as intended, system testing can be performed to provide more insight into the potential issues. The software can be tested with the help of unit tests to ensure correct logic, while hardware can be tested by performing a set of measurement tests that assess the accuracy and correct operations of the instruments and fixture. If a fault is found in the hardware, that component can easily be exchanged thanks to the modular design of the test system. PXI modules can be obtained with a short lead time from a vendor or distributor, while cables and the fixture PCB will have a slightly longer lead time due to them being custom made. It has been found that a PXI-based system has a lower overall life cycle cost, when compared to a legacy system thanks to its modular nature [77].

Software maintainability is ensured by extensive testing before a new release. Changes are tracked with the help of Git, making it possible to revert back to a stable release if a software problem is found later. Modular software design ensures that only specific parts of the software are affected when making changes.

Overall, a modular codebase and a modular hardware platform option enabled that overall modularity was easily achievable. The design choices made will make it easier to modify the system in the future if the need for such arises.

5.2 Traceability of Results

As previously highlighted, documentation and traceability hold exceptional importance in the medical device sector. Within the EU, the MDR outlines extensive requirements on documentation to ensure safety and effectiveness. When developing a medical device detailed documentation of all development steps, design decisions and technical specifications is mandatory and must be consistently maintained. This applies to medical devices in development, but also to any tools, instruments, and systems that may be used in the development process.

Specifically, for the ATS these regulatory requirements imply that all instruments must be properly calibrated and certified according to recognized standards such as ISO 17025. Furthermore, the operating software must adhere to software development standards relevant for medical device systems. Additionally, as mentioned in Chapter 2.2, a comprehensive risk assessment, typically documented as a structured hazard analysis per ISO 14971, is required to capture any risks potentially introduced by the ATS to the DUT. Compliance with established electrical safety standards, particularly IEC 60601-1, must also be documented to demonstrate that the ATS itself does not introduce safety risks during testing.

A vast amount of documented DUT performance evidence is essential for both verification purposes, as well as for possible future needs. When modifications are made to a medical device on the market, regulatory framework requires that essential performance remains unchanged, shown by directly comparing new test results to historical test data. Therefore, it is crucial that all test results are metrologically traceable, meaning each step, parameter setting, measurement outcome, and procedural detail must be systematically documented and interlinked to ensure complete reproducibility. Documentation of this scale can be time consuming, but when integrated as part of automated systems, it is efficient.

The developed ATS addresses these documentation and traceability requirements through automatic data logging and systematic digital record-keeping of test details and parameters. Test protocols, generated data logs, and structured test reports collectively create a cohesive documentation system, facilitating complete and straightforward traceability. This aligns well with industry standard procedures, which have been found to provide clear benefits during validation and verification procedures [78]. Table 16 shows how the three key documents utilized during testing are interconnected together by different data ('x' indicates that the given information is recorded in the corresponding document), guaranteeing complete traceability of every performed test.

	Test Protocol	Test Report	Log File
Test Protocol ID	x	x	
Test Software Version		x	x
Date of testing		x	x
Test ID	x	x	x
Test Procedure	x		x
Test Parameters	x	x	
Acceptance Criteria	x	x	x
Test Result		x	x
Equipment ID		x	x
DUT ID		x	x

Table 16: Information included across test documents

To assess the effectiveness and robustness of the implemented traceability measures, an extensive audit of recorded test sequences was conducted. This audit specifically evaluated the extent of data logging, accuracy of timestamps and comprehensiveness of the stored data, to enable full reconstruction of test procedures. The audit confirmed that sufficient amount of data is automatically stored by the ATS to be able to recreate all performed tests exclusively from the recorded information, thus meeting regulatory and operational requirements comprehensively.

5.3 Practical Implications

As the developed ATS has proven to meet both the functional and performance requirements, it is clear that it provides a reliable, scalable and compliant tool well suited for use in pre-verification of the non-invasive glucose monitoring device that GlucoModicum is currently developing. This ATS will be utilized to gather important performance data of the NIGM device, ensuring that the development of the device continues in the right direction towards a successful verification process. New hardware and ESW changes for the NIGM device will be evaluated to confirm that all properties and functions of the device have either remained unchanged or improved. Quality testing to ensure safety and performance of NIGM devices used in Clinical Studies will be conducted with the new PXI-based ATS. Furthermore, the ATS will also work as proof-of-concept for a new and improved iteration of a PXI-based ATS, that will then be used in the official verification process of the NIGM device. The prototype itself will remain in active R&D use, functioning as a testing platform for hardware and ESW improvements of the NIGM. This PXI-based ATS will replace the previous ATS version. New test cases might be implemented supporting the R&D work, possibly accelerating the development cycles and improving product reliability prior to market release.

5.4 Challenges and Limitations

While the developed ATS successfully fulfills its purpose as a prototype displaying the functionality of the concept and meeting all set requirements, several challenges were encountered during the development process. Therefore, the ATS inherits some known limitations that were not possible nor feasible to correct at this stage. These limitations do not compromise the system's suitability for its intended purpose, but they highlight areas where future development would be required.

Firstly, the number of relay switches on the Pickering relay modules limit the complexity of the test circuit and the number of available DUT slots. Creating a fixture where every possible test circuit configuration is implemented would have required significantly more relays than the available 64 relays. It was decided that two 32-relay modules would be used for the prototype, as other alternatives with more relays would have increased the cost of the prototype and the availability would have become a challenge. At the same time, it was deemed that all possible test configurations were not necessary to be available, as long as the main configurations would be possible to create. The number of relays also limited the amount of available DUT slots. As mentioned in Chapter 5.1, each new slot added to the system requires an additional relay for each of the six pins.

Another limitation caused by the choice of components is the restricted range of the programmable resistor module. It is limited to only being able to set loads of up to 131 kΩ, meaning that testing requiring higher loads are not possible without the usage of an external load module. PXI Resistor Modules with a wider range exists and were considered, however, to save on expenses it was decided to use this specific module for the prototype ATS. The most common loads used when testing the NIGM device are 0 Ω, 1 kΩ and 10 kΩ, which are all possible to produce with this module. They reflect the realistic working environment that the NIGM device is expected to operate in. Yet, it could be useful to conduct certain types of tests with loads up to 1000 kΩ or even 10 000 kΩ.

Despite being a scalable ATS that supports multiple DUTs, a notable limitation of the design is that it does not support parallel testing. Up to four DUTs can be prepared and configured simultaneously, but test measurements are conducted on one DUT at a time, without any overlap. This means that if the desire to be able to perform parallel testing in the future grows, a substantial amount of work is required redeveloping both the software and the hardware components. For GlucoModicum's purpose it is not feasible to invest in the additional hardware needed for enabling full parallel testing, as R&D and pre-verification testing capacity are not bottlenecks. However, certain tests are possible to reconfigure to allow parallel execution without the need for hardware scaling. These tests include those where the voltage measurement accuracy is evaluated, and the reference voltage is not affected by multiple DUTs being connected to the same test circuit at the same time. Similarly, tests evaluating operation length and timing, where essentially no reference signal nor test circuit is needed, are possible to execute at the same time as the current design.

A known limitation that was the focus of one of the major design decisions made during the development of the ATS was the choice of the communication protocol to be used between the controlling software and the instrument modules. This design choice was presented in Chapter 3.5.3. The choice was made to utilize vendor specific control libraries. As the test system is designed to utilize DMM Modules from National Instruments and Relay and Resistor Modules from Pickering, switching the modules to alternatives from other vendors will require certain software reconfiguration, despite the hardware fully supporting immediate change. Despite the fact that developing the ATS to utilize VISA for communication would have made the ATS more scalable, the decision was made to utilize vendor specific communication as the other benefits outweighed the scalability support. Therefore, this was a conscious limitation assessed to be acceptable.

The error handling capabilities of the software can also be further improved. Basic error conditions are currently well-handled and logged by the system. However, certain disconnection scenarios, where, for example, communication with an instrument or a DUT is lost, can cause the software to enter an unrecoverable error state, requiring manual intervention. Further development of the software could improve auto-recovery in these types of situations allowing the ATS to automatically self-correct. This would improve system autonomy, reducing the need for user supervision. Refining the software was not put as one of the focus areas, as the project only aimed to test the feasibility of a PXI-based ATS.

The UI of the ATS incorporates only very limited features and is thus a possible focus area for further improvement. Currently the test circuit cannot be manually configured, only by choosing predefined configurations detailed in an XML file. Full manual control of the circuit would be especially beneficial in R&D use where unique test circuit configurations might be needed on short notice. The UI also only displays the result of a conducted test as PASS or FAIL, without giving the user any deeper insight into the data. Displaying mean or standard deviation values, giving a near real-time visibility into the results would improve debugging and much quicker highlight serious issues with the test sequence. Currently, deeper insight into the results requires generation of a test report or exporting of the raw data, which is only available after the test protocol is completed.

It is evident that further development is required to realize the full potential of the ATS presented here. Nevertheless, a simpler yet more reliable system often provides greater value than a feature-rich alternative whose functionality cannot be fully trusted. In practice, increased system complexity requires significantly more rigorous testing efforts to reach comparable levels of reliability [79].

6 Conclusion

This thesis aimed to develop and evaluate a scalable automated test system to enable efficient and accurate pre-verification testing of a non-invasive glucose monitoring device. The prime objectives were to ensure regulatory compliance, measurement accuracy and test repeatability. Additionally, this work sought to illustrate key design choices and critical decisions required during the development process, emphasizing important considerations.

The motivation behind developing an automated test system for pre-verification purposes stems from the necessity to perform accurate and comprehensive testing of specific device components, operational features, and subsystems. Such testing confirms that the DUT fulfills set technical requirements to ensure performance. In addition, certain safety-related tests, such as compliance with electrical safety standards like IEC 60601-1, need accurate in vitro testing methodologies to assess compliance. The ATS allows these critical tests to be performed and documented efficiently and reliably without involving human subjects. This significantly mitigates ethical and practical challenges.

Although, in vitro testing can be conducted manually, the ATS clearly demonstrated to be a more efficient and accurate option. Utilizing a tool like an ATS ensures that each test cycle is executed under identical conditions, eliminating variability introduced by manual operation. Furthermore, as measured data is directly handled by a machine, the risk of human-induced errors is significantly minimized, providing consistent and dependable results. Primarily, however, the ATS offers substantial advantage by conducting repeated testing at a faster pace, occupying fewer human resources.

Two distinct approaches were considered when determining the optimal design of the ATS. The first alternative involved creating a fully custom designed test system from individual instruments and components, tailored precisely according to the needs and requirements of this specific task. However, this alternative presented multiple challenges, including requiring a longer development timeline, due to needing to design each component individually. Other challenges included less available technical support, and increased complexity when ensuring scalability. The other alternative was to leverage a commercially available PXI-based platform, developing the ATS utilizing off-the-shelf modules, only custom-developing essential components.

Selecting the PXI-platform has proved to be a successful choice. The ATS demonstrated excellent scalability, as the individual modules are interchangeable and the test circuit can flexibly be modified without the need for extensive redesign. The finalized design of the ATS effectively delivers highly accurate and consistently repeatable test results, while remaining easily reconfigurable if needs evolve over time. The ATS successfully met the defined requirements specified in Chapter 3.4. Verification tests demonstrated that noise measurement stayed constantly within acceptable levels. Maximum measured noise was found to be less than 2 nA, which indicates that accuracy is more than sufficient to meet the required ±10 nA, falling inside the TUR 4:1 margin. Furthermore, The ATS proved capable of producing accurate results consistently over multiple test cycles, indicating strong repeatability. The standard deviation of the relative error across all 20 test cycles remained below

1 %. Certain tests with less variability achieved a standard deviation of the relative error of less than 0.001 %.

The ATS serves its intended purpose by demonstrating that a scalable ATS built on a PXI-base is a highly efficient tool for pre-verification testing. From an academic standpoint, this thesis contributes by showcasing a practical and effective implementation of an automated test methodology for wearable medical device development, potentially guiding future research and system implementations. From the company's perspective, implementing the ATS into GlucoModicum's design & development project as a testing tool, shows promise of ensuring a quicker and smoother process towards regulatory and market release.

At the same time the scalable design enables flexible modification of the test system without the need for significant redesign and with minimal downtime. The modularity of the design enables other companies to adapt the findings in this thesis and develop a similar test system for a different type of device. An ATS like this could be adapted to testing blood pressure monitors, smart watches, digital hearing aids, and other electronic devices that output or measure small current and voltage signals.

While the ATS demonstrates high performance and promising results, room for improvement still remains. Tackling some of the known limitations in future iterations will increase the performance and efficiency of the ATS even further. Future research should focus on integrating parallel capabilities to increase scalability and testing capacity even more. Another valuable long-term improvement to consider would be to transform the ATS into utilizing a result database model, where new measurements can be compared to historical results, enabling comparison, trend analysis, and automatic alerts of sudden changes in behavior.

In conclusion, the developed ATS demonstrates promising potential for becoming a highly useful tool for GlucoModicum, serving an important role in ensuring the NIGM device will be ready for official verification. Introducing a robust non-invasive glucose monitoring device to the market could mark a significant leap forward in glucose monitoring and treatment of diabetes. This will potentially help reverse the current global diabetes prevalence trends and positively impact both healthcare spending and patient health outcomes.

References

[1] R. K. Kher and D. M. Patel, “A comprehensive review on wearable health monitoring systems,” *The Open Biomed. Eng. J.*, vol. 15, no. 1, pp. 213–225, 2021. doi: 10.2174/1874120702115010213.

[2] C. J. L. Murray, “The global burden of disease study at 30 years,” *Nat. Med.*, vol. 28, no. 10, pp. 2019–2026, 2022. doi: 10.1038/s41591-022-01990-1.

[3] S. E. Vollset *et al.*, “Burden of disease scenarios for 204 countries and territories, 2022–2050: a forecasting analysis for the global burden of disease study 2021,” *Lancet*, vol. 403, no. 10440, pp. 2204–2256, 2024. doi: 10.1016/S0140-6736(24)00685-8.

[4] World Health Organization, “Prevalence of diabetes,” World Health Organization, December 2024, Accessed: Apr. 13, 2025. [Online]. Available: https://www.who.int/data/gho/data/indicators/indicator-details/GHO/prevalence-of-diabetes-age-standardized

[5] M. R. Rooney *et al.*, “Global prevalence of prediabetes,” *Diabetes Care*, vol. 46, no. 7, pp. 1388–1394, 2023. doi: 10.2337/dc22-2376.

[6] J. Kesavadev *et al.*, “The historical evolution of continuous glucose monitoring - the story of 25 years,” *Int. J. Diabetes Technol*, vol. 2, no. 4, pp. 129–136, 2023. doi: 10.4103/ijdt.ijdt_16_24.

[7] W. Villena Gonzales, A. T. Mobashsher, and A. Abbosh, “The progress of glucose monitoring-a review of invasive to minimally and non-invasive techniques, devices and sensors,” *Sensors*, vol. 19, no. 4, p. 800, 2019. doi: 10.3390/s19040800.

[8] S. Vaddiraju, D. J. Burgess, I. Tomazos, F. C. Jain, and F. Papadimitrakopoulos, “Technologies for continuous glucose monitoring: current problems and future promises,” *J. Diabetes Sci. Technol.*, vol. 4, no. 6, pp. 1540–1562, 2010. doi: 10.1177/193229681000400632.

[9] P. P. Pai, P. K. Sanki, S. Sarangi, and S. Banerjee, “Modelling, verification, and calibration of a photoacoustics based continuous non-invasive blood glucose monitoring system,” *Rev. Sci. Instrum.*, vol. 86, no. 6, p. 064901, 2015. doi: 10.1063/1.4922416.

[10] J. Smith, *The Pursuit of Noninvasive Glucose: Hunting the Deceitful Turkey*, 9th ed., 12 2023.

[11] J. McLenon and M. A. M. Rogers, “The fear of needles: A systematic review and meta-analysis,” *J. Adv. Nurs.*, vol. 75, no. 1, pp. 30–42, 2019. doi: 10.1111/jan.13818.

[12] I. Lee, D. Probst, D. Klonoff, and K. Sode, “Continuous glucose monitoring systems - current status and future perspectives of the flagship technologies in biosensor research,” *Biosens. Bioelectron.*, vol. 181, p. 113054, 2021. doi: 10.1016/j.bios.2021.113054.

[13] T. A. Hakala *et al.*, “Sampling of fluid through skin with magnetohydrodynamics for noninvasive glucose monitoring,” *Sci. Rep.*, vol. 11, no. 1, p. 7609, 2021. doi: 10.1038/s41598-021-86931-7.

[14] J. Kim, A. S. Campbell, and J. Wang, “Wearable non-invasive epidermal glucose sensors: A review,” *Talanta*, vol. 177, pp. 163–170, 2018. doi: 10.1016/j.talanta.2017.08.077.

[15] J. Wu, Y. Liu, H. Yin, and M. Guo, “A new generation of sensors for non-invasive blood glucose monitoring,” *Am J Transl Res*, vol. 15, no. 6, pp. 3825–3837, 2023, Accessed: Dec. 8. 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/37434817

[16] S. Ulewicz and B. Vogel-Heuser, “Increasing system test coverage in production automation systems,” *Control Eng. Pract.*, vol. 73, pp. 171–185, 2018. doi: 10.1016/j.conengprac.2018.01.010.

[17] E. Alégroth, R. Feldt, and P. Kolström, “Maintenance of automated test suites in industry: An empirical study on visual gui testing,” *Inf. Softw. Technol.*, vol. 73, pp. 66–80, 2016. doi: 10.1016/j.infsof.2016.01.012.

[18] M. J. Hossain, M. Al-Mamun, and M. R. Islam, “Diabetes mellitus, the fastest growing global public health concern: Early detection should be focused,” *Health Sci. Rep.*, vol. 7, no. 3, p. e2004, 2024. doi: 10.1002/hsr2.2004.

[19] International Diabetes Federation, *IDF Diabetes Atlas*, 10th ed. Brussels, Belgium: International Diabetes Federation, 2021.

[20] X. Lin *et al.*, “Global, regional, and national burden and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025,” *Sci. Rep.*, vol. 10, no. 1, p. 14790, 2020. doi: 10.1038/s41598-020-71908-9.

[21] S. F. Clarke and J. R. Foster, “A history of blood glucose meters and their role in self-monitoring of diabetes mellitus,” *Br. J. Biomed. Sci.*, vol. 69, no. 2, pp. 83–93, 2012, Accessed: Jan. 26, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/22872934

[22] N. C. D. R. F. Collaboration, “Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants,” *Lancet*, vol. 404, no. 10467, pp. 2077–2093, 2024. doi: 10.1016/S0140-6736(24)02317-1.

[23] B. M. Popkin, “The nutrition transition: an overview of world patterns of change,” *Nutr. Rev.*, vol. 62, no. 7 Pt 2, pp. S140–S143, 2004. doi: 10.1111/j.1753-4887.2004.tb00084.x.

[24] C. M. Nightingale *et al.*, “Screen time is associated with adiposity and insulin resistance in children,” *Arch. Dis. Child.*, vol. 102, no. 7, pp. 612–616, 2017. doi: 10.1136/archdischild-2016-312016.

[25] L. Krist *et al.*, “Determinants of physical activity and screen time trajectories in 7th to 9th grade adolescents-a longitudinal study,” *Int. J. Environ. Res. Public Health*, vol. 17, no. 4, 2020. doi: 10.3390/ijerph17041401.

[26] F. G. Banting and C. H. Best, “The internal secretion of the pancreas,” *J Lab Clin Med*, vol. 7, pp. 251–266, 1922.

[27] M. Lind *et al.*, “Continuous glucose monitoring vs conventional therapy for glycemic control in adults with type 1 diabetes treated with multiple daily insulin injections: The gold randomized clinical trial,” *JAMA*, vol. 317, no. 4, pp. 379–387, 2017. doi: 10.1001/jama.2016.19976.

[28] M. M. F. Choi, “Progress in enzyme-based biosensors using optical transducers,” *Microchim. Acta*, vol. 148, no. 3-4, pp. 107–132, 2004. doi: 10.1007/s00604-004-0273-8.

[29] H. Teymourian, A. Barfidokht, and J. Wang, “Electrochemical glucose sensors in diabetes management: an updated review (2010-2020),” *Chem. Soc. Rev.*, vol. 49, no. 21, pp. 7671–7709, 2020. doi: 10.1039/d0cs00304b.

[30] S. K. Vashist, D. Zheng, K. Al-Rubeaan, J. H. Luong, and F. S. Sheu, “Technology behind commercial devices for blood glucose monitoring in diabetes management: a review,” *Anal. Chim. Acta*, vol. 703, no. 2, pp. 124–36, 2011. doi: 10.1016/j.aca.2011.07.024.

[31] S. Cradock and J. Hawthorn, “Pain, distress and blood glucose monitoring,” *J. of Diabetes Nurs.*, vol. 6, no. 6, pp. 188–191, 2002.

[32] The DCCT Research Group, “Diabetes control and complications trial (DCCT): results of feasibility study,” *Diabetes Care*, vol. 10, no. 1, pp. 1–19, 1987. doi: 10.2337/diacare.10.1.1.

[33] J. Sato, T. Hirose, and H. Watada, “Continuous glucose monitoring system: Is it really accurate, safe and clinically useful?” *J Diabetes Investig*, vol. 3, no. 3, pp. 225–30, 2012. doi: 10.1111/j.2040-1124.2012.00197.x.

[34] J. L. Kelly, “Continuous insulin infusion: When, where, and how?” *Diabetes Spectr.*, vol. 27, no. 3, pp. 218–23, 2014. doi: 10.2337/diaspect.27.3.218.

[35] U.S. Food and Drug Administration, “Premarket Approval (PMA) Report: P980022S011 – Guardian RT,” July 18, 2005 2005, Accessed: Jul. 25, 2025. [Online]. Available: https://fda.report/PMA/P980022S011

[36] F. Thomas, C. Pretty, J. Dickson, M. Signal, G. Shaw, and J. G. Chase, “A simple method to model a continuous glucose monitoring signal,” *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 8775–8780, 2017. doi: 10.1016/j.ifacol.2017.08.1736.

[37] S. Templer, “Closed-loop insulin delivery systems: Past, present, and future directions,” *Front. Endocrinol.*, vol. 13, p. 919942, 2022. doi: 10.3389/fendo.2022.919942.

[38] I. B. Bender and A. B. Bender, “Diabetes mellitus and the dental pulp,” *J. Endod.*, vol. 29, no. 6, pp. 383–9, 2003. doi: 10.1097/00004770-200306000-00001.

[39] D. Klyve, S. Lowe, K. Currie, J. H. Anderson, C. Ward, and B. Shelton, “Noninvasive blood glucose measurement using RF spectroscopy and a lightgbm ai model,” *IEEE Sensors Journal*, vol. 24, no. 17, pp. 28 049–28 055, 2024. doi: 10.1109/jsen.2024.3405800.

[40] D. B. Keenan, J. J. Mastrototaro, G. Voskanyan, and G. M. Steil, “Delays in minimally invasive continuous glucose monitoring devices: a review of current technology,” *J. Diabetes Sci. Technol.*, vol. 3, no. 5, pp. 1207–14, 2009. doi: 10.1177/193229680900300528.

[41] T. A. Hakala *et al.*, “Pilot study in human healthy volunteers on the use of magnetohydrodynamics in needle-free continuous glucose monitoring,” *Sci. Rep.*, vol. 12, no. 1, p. 18318, 2022. doi: 10.1038/s41598-022-21424-9.

[42] C. Rados, “Medical device and radiological health regulations come of age,” *FDA Consum.*, vol. 40, no. 1, pp. 58–65, 2006. [Online]. Available: https://www.ncbi.nlm.nih.gov/pubmed/16528829

[43] A. C. Cipollaro and M. B. Einhorn, “The use of X-rays for the treatment of hypertrichosis is dangerous,” *J. Am. Med. Assoc.*, vol. 135, no. 6, pp. 349–53, 1947. doi: 10.1001/jama.1947.62890060007008.

[44] H. Steinberg, “Electrotherapeutic disputes: the ’Frankfurt Council’ of 1891,” *Brain*, vol. 134, no. Pt 4, pp. 1229–1243, 2011. doi: 10.1093/brain/awr040.

[45] S. S. Altayyar, “The essential principles of safety and effectiveness for medical devices and the role of standards,” *Med Devices (Auckl)*, vol. 13, pp. 49–55, 2020. doi: 10.2147/MDER.S235467.

[46] Council of European Union, “Regulation (EU) 2017/745 of the European Parliament and of the Council of 5 April 2017 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC,”

pp. 1–175, 2017/05/05 2017, Accessed: Aug. 5, 2025. [Online]. Available: http://data.europa.eu/eli/reg/2017/745/oj

[47] M. Fenwick, W. A. Kaal, and E. P. Vermeulen, "Regulation tomorrow: what happens when technology is faster than the law," *Am. U. Bus. L. Rev.*, vol. 6, no. 3, pp. 561–594, 2016.

[48] M. Fink and B. Akra, "Comparison of the international regulations for medical devices-USA versus Europe," *Injury*, vol. 54 Suppl 5, p. 110908, 2023. doi: 10.1016/j.injury.2023.110908.

[49] P. Verbruggen and B. van Leeuwen, "The liability of notified bodies under the EU's new approach: The implications of the PIP breast implants case," *Eur. Law Rev.*, vol. 43, no. 3, pp. 394–409, 2018.

[50] R. L. Owens, K. C. Wilson, I. Gurubhagavatula, and R. Mehra, "Philips respironics recall of positive airway pressure and noninvasive ventilation devices: A brief statement to inform response efforts and identify key steps forward," *Am. J. Respir. Crit. Care Med.*, vol. 204, no. 8, pp. 887–890, 2021. doi: 10.1164/rccm.202107-1666ED.

[51] G. G. Balan, C. V. Sfarti, S. A. Chiriac, C. Stanciu, and A. Trifan, "Duodenoscope-associated infections: a review," *Eur. J. Clin. Microbiol. Infect. Dis.*, vol. 38, no. 12, pp. 2205–2213, 2019. doi: 10.1007/s10096-019-03671-3.

[52] N. G. Leveson and C. S. Turner, "An investigation of the Therac-25 accidents," *Computer*, vol. 26, no. 7, pp. 18–41, 1993. doi: 10.1109/mc.1993.274940.

[53] J. B. Bronneke, J. Muller, K. Mouratis, J. Hagen, and A. D. Stern, "Regulatory, legal, and market aspects of smart wearables for cardiac monitoring," *Sensors*, vol. 21, no. 14, p. 4937, 2021. doi: 10.3390/s21144937.

[54] G. A. Van Norman, "Drugs, devices, and the FDA: Part 2: An overview of approval processes: FDA approval of medical devices," *JACC Basic Transl. Sci.*, vol. 1, no. 4, pp. 277–287, 2016. doi: 10.1016/j.jacbts.2016.03.009.

[55] A. Ravizza, C. De Maria, L. Di Pietro, F. Sternini, A. L. Audenino, and C. Bignardi, "Comprehensive review on current and future regulatory requirements on wearable sensors in preclinical and clinical testing," *Front. Bioeng. Biotechnol.*, vol. 7, p. 313, 2019. doi: 10.3389/fbioe.2019.00313.

[56] L. Peter, L. Hajek, P. Maresova, M. Augustynek, and M. Penhaker, "Medical devices: Regulation, risk classification, and open innovation," *J. Open Innov. Technol. Market Complex.*, vol. 6, no. 2, p. 42, 2020. doi: 10.3390/joitmc6020042.

[57] J. Shahid, R. Ahmad, A. K. Kiani, T. Ahmad, S. Saeed, and A. M. Almuhaideb, "Data protection and privacy of the internet of healthcare things (IoHTs)," *Appl. Sci.*, vol. 12, no. 4, p. 1927, 2022. doi: 10.3390/app12041927.

[58] C. Amaral, M. Paiva, A. R. Rodrigues, F. Veiga, and V. Bell, “Global regulatory challenges for medical devices: Impact on innovation and market access,” *Appl. Sci.*, vol. 14, no. 20, 2024. doi: 10.3390/app14209304.

[59] Y. J. Chen, C. M. Chiou, Y. W. Huang, P. W. Tu, Y. C. Lee, and C. H. Chien, “A comparative study of medical device regulations: US, Europe, Canada, and Taiwan,” *Ther. Innov. Regul. Sci.*, vol. 52, no. 1, pp. 62–69, 2018. doi: 10.1177/2168479017716712.

[60] R. Grau, P. Ens, K. Bause, and T. Düser, “Medical device validation: State of the art and benefits of X-in-the-loop framework implementation,” *Volume 4: Biomedical and Biotechnology Engineering*, 2025. doi: 10.1115/IMECE2024-145076.

[61] N. H. Abd Rahman, A. K. Ibrahim, K. Hasikin, and N. A. Abd Razak, “Critical device reliability assessment in healthcare services,” *J. Healthc. Eng.*, vol. 2023, no. 1, p. 3136511, 2023. doi: 10.1155/2023/3136511.

[62] V. L. Goncalves, L. C. Mendes, A. C. Rosa, A. A. Pereira, and S. T. Milagre, *Systematic Review: Metrology in Health- Failures, Limitations and Challenges*, ser. IFMBE Proceedings. Springer Singapore, 2019, book section Chapter 132, pp. 847–850. doi: 10.1007/978-981-13-2119-1_132.

[63] B. B. Murphy, B. H. Scheid, Q. Hendricks, N. V. Apollo, B. Litt, and F. Vitale, “Time evolution of the skin-electrode interface impedance under different skin treatments,” *Sensors*, vol. 21, no. 15, 2021. doi: 10.3390/s21155210.

[64] J. Lee, M. Kim, H. K. Park, and I. Y. Kim, “Motion artifact reduction in wearable photoplethysmography based on multi-channel sensors with multiple wavelengths,” *Sensors*, vol. 20, no. 5, 2020. doi: 10.3390/s20051493.

[65] A. D. Stern, “Innovation under regulatory uncertainty: Evidence from medical technology,” *J. Public Econ.*, vol. 145, pp. 181–200, 2017. doi: 10.1016/j.jpubeco.2016.11.010.

[66] F. Redmill, “Theory and practice of risk-based testing,” *Softw. Test. Verif. Rel.*, vol. 15, no. 1, pp. 3–20, 2005. doi: 10.1002/stvr.310.

[67] D. H. Collins, A. V. Huzurbazar, and R. L. Warr, “Highly accelerated life testing (HALT): A review from a statistical perspective,” *WIREs Comput. Stat.*, vol. 16, no. 4, 2024. doi: 10.1002/wics.70000.

[68] S. M. Mimbs, “Measurement decision risk - the importance of definitions,” in *Measurement Science Conference*. National Aeronautics and Space Administration (NASA), 2008, Conference Proceedings.

[69] M. Bretthauer, S. Gerke, C. Hassan, O. F. Ahmad, and Y. Mori, “The new European Medical Device Regulation: Balancing innovation and patient safety,” *Ann. Intern. Med.*, vol. 176, no. 6, pp. 844–848, 2023. doi: 10.7326/M23-0454.

[70] J. B. Pietzsch, L. A. Shluzas, M. E. Paté-Cornell, P. G. Yock, and J. H. Linehan, "Stage-gate process for the development of medical devices," *J. Med. Devices*, vol. 3, no. 2, p. 021004, 2009. doi: 10.1115/1.3148836.

[71] A. Bahill and R. Botta, "Fundamental principles of good system design," *Eng. Manag. J.*, vol. 20, no. 4, pp. 9–17, 2008.

[72] E. Kemp *et al.*, "Influence of enzyme immobilization and skin-sensor interface on non-invasive glucose determination from interstitial fluid obtained by magnetohydrodynamic extraction," *Biosens. Bioelectron.*, vol. 206, p. 114123, 2022. doi: 10.1016/j.bios.2022.114123.

[73] International Organization for Standardization, "ISO 13485:2016 – Medical devices – Quality management systems – Requirements for regulatory purposes," Geneva, Switzerland, 2016.

[74] A. Björklund, "Modular automated evaluation system for electrical safety and performance testing of a non-invasive glucose monitoring device," Master's thesis, Aalto University, Espoo, Finland, 2023.

[75] M. S. Ghoneim, A. Mohammaden, L. A. Said, A. H. Madian, A. G. Radwan, and A. M. Eltawil, "A comparative study of different human skin impedance models," in *Proc. 38th Nat. Radio Sci. Conf. (NRSC)*, vol. 1, Cairo, Egypt, 2021, Conference Proceedings, pp. 271–277. doi: 10.1109/nrsc52299.2021.9509823.

[76] T. Adachi, A. Pramanick, and M. Elston, "Parallel, multi-DUT testing in an open architecture test system," in *Proc. IEEE Int. Test Conf.*, 2005, Conference Proceedings. doi: 10.1109/TEST.2005.1584053.

[77] J. Ginn and M. Dewey, "Next generation test system architectures for depot and O-level test," in *Proc. IEEE Autotestcon*, 2011, Conference Proceedings, pp. 316–320. doi: 10.1109/AUTEST.2011.6058758.

[78] O. McDermott, N. Conroy, M. Thenarasu, and S. Duarte, "Digitalised validation systems as an enabler for quality 4.0 within a medical device manufacturer," *Sustain. Futures*, vol. 8, p. 100383, 2024. doi: 10.1016/j.sftr.2024.100383.

[79] K. Kevrekidis, S. Albers, P. J. M. Sonnemans, and G. M. Stollman, "Software complexity and testing effectiveness: An empirical study," in *Annu. Rel. Maintainability Symp. (RAMS)*, 2009, Conference Proceedings. doi: 10.1109/rams.2009.4914733.