DTU Health Tech
Department of Health Technology

Mobile Health Application for Self-Administered Evaluation of Neuropathy

Master Thesis

Mobile Health Application for Self-Administered Evaluation of Neuropathy

Master Thesis
August, 2023

By
Julia Gabriela Makulec, Mads Westermann

Cover photo: Vibeke Hempler, 2012
Published by: DTU, Department of Health Technology, Richard Petersens Plads, Building 324, 2800 Kgs. Lyngby Denmark
www.healthtech.dtu.dk
ISSN: [0000-0000] (electronic version)
ISBN: [000-00-0000-000-0] (electronic version)

ISSN: [0000-0000] (printed version)
ISBN: [000-00-0000-000-0] (printed version)

Approval

This thesis has been prepared over a semester at Department of Health Technology, at the Technical University of Denmark, DTU, in partial fulfilment for the Master of Science in Engineering degree, MSc Eng.

Julia Gabriela Makulec s213041	Mads Westermann s174508
.....	.....
Signature	*Signature*
.....	.....
Date	*Date*

Abstract

Peripheral neuropathies are disorders of peripheral nerves, producing symptoms like pain, muscle weakness and sensory loss. It has multiple causes and can affect a sizeable percentage of the population, especially diabetics and people exposed to toxic treatment like chemotherapy. The diagnostic process requires specialized tools and trained clinicians. There is few tools available for patients to conduct an evaluation of their symptoms by themselves.

A Mobile Health (mHealth) application was designed and developed during this project, implementing a self-assessment questionnaire inspired by the Utah Early Neuropathy Scale (UENS). During the design User-Centered Design (UCD) methods were employed to create a set of instructions, tests and questions accessible for users with no medical training.

Additionally to usability validations over the iterations of design, a study with 17 participants was carried out to evaluate the new tool and compare it to a state-of-art measure. In the study patients used the implemented mHealth application to conduct a self-assessment, and then the score was compared with results of Total Neuropathy Score clinical (TNSc) conducted by the authors. The Pearson correlation was 0.861, implying a high positive correlation between the two measures.

Acknowledgements

Jakob Eyvind Bardram, PhD, Professor in Computer Science at Technical University of Denmark (DTU)
The authors would like to thank Jakob E. Bardram for supervision during the thesis project, as well as providing guidance on implementation and writing the manuscript.

Martin Ballegaard, Associate professor, consultant neurophysiologist, Sjællands Universitetshospital Roskilde
Martin Ballegaard as the clinical supervisor provided invaluable insight and directions during the project. The authors would like to express gratitude for assistance, facilitating the study with patients and overall excellent collaboration.

Nanna Brix Finnerup, Professor, consultant neurologist, Danish Pain Research Center at Aarhus University
The authors wish to thank Nanna Brix Finnerup for providing feedback on the project design, helping to bridge the domains of medicine and technology.

Contents

1 Introduction

Peripheral neuropathies are disorders of peripheral nerves, which are located in the extremities. They can be symmetrical across the body axis, asymmetrical or localized. Polyneuropathy describes involvement of multiple peripheral nerves simultaneously, bilaterally and symmetrically. In most polyneuropathies, small sensory nerve fibers are affected earlier and more severely than large motor fibers. Symptoms progress from the extremities towards the body, typically starting in toes and gradually affecting larger parts of legs. Eventually, the disease appears in hands and then arms. Clinical examination findings, also referred to as ”signs” of polyneuropathy include absent or reduced deep tendon reflexes, paresthesias, sensory loss and distal muscle weakness in extremities [1][2].

One of the most common sub-types of peripheral neuropathy is Distal Symmetrical Axonal Polyneuropathy (DSP). There is multiple causes of DSP, including chemotherapeutic agents, heavy, chronic alcohol use and most prominent - diabetes. Patients may experience numbness, tingling, decrease in motor function, pain or a mixture of these. Symptoms are present in a distal-to-proximal gradient pattern, meaning severity increases with distance from body center. Usually, they will reach knee level before appearing in hands [3][4][5][6]. Early detection of DSP is crucial for improving morbidity and mortality, such as preventing diabetic foot-ulcers [7][8][9]. For judging certainty of the presence of DSP, Tesfaye et al. [10] define the following levels of certainty:

- **Possible**: The patient has symptoms or signs of DSP.
- **Probable**: The patient has symptoms and signs of DSP.
- **Definite**: The patient has symptoms and signs of DSP and an abnormal result from a Nerve Conduction Study (NCS) (or different specialized study if NCS is normal).

In order to achieve a definite diagnosis of DSP, an examination consisting of identifying symptoms, gathering findings and a NCS of the patient is required [11][10]. That examination is complex and consists of several parts, requiring a trained professional as well as special clinical equipment. Wait time from referral to a neurology consultation can take multiple weeks, and patients with a suspicion of a peripheral nerve disorder on average wait longer for their appointment [12].

In settings where the specialized equipment is not available, like oncology wards or general practitioner’s office, a field study can be used to evaluate symptoms and findings of DSP [11]. A field study is a method of gathering data or research outside of a clinical setting. Since it does not employ NCS, the highest achievable level of certainty is probable [10].

There is a variety of measures that can be used by a clinician when making a DSP diagnosis [5]. They vary in procedures used and the areas of examination. Grading (score) of the results is also specific to the measure, as points assigned for particular symptoms or findings are unique to each tool. This means the result point scales for measures are different, and not all of them provide a diagnosis threshold.

Patient-Reported Outcome (PRO) is a report on a patient’s health condition sourced directly from the patient, not interpreted by a clinician or anyone else [13]. It is considered to be growing in importance in the patient-centered healthcare system [14]. Self-assessment tools and scores can be meaningful in assessing a condition’s extent, an example being

the Self Assessment Vitiligo Extent Score [15]. While there are many types of PROs, electronic PROs are found to be more beneficial than paper-based ones, reducing the risk of incorrect data entry, increasing patient's willingness to disclose sensitive information, and enabling immediate access to data [14].

Currently, few self-assessment tools for DSP have been developed (see chapter 2.3) and most are focused on Chemotherapy-Induced Peripheral Neuropathy (CIPN). One of the available solutions - Neuropad [16] is a physical self-assessment tool that can be used for early detection of DSP in diabetes patients [17]. Currently, the product is only available for purchase in the UK and Ireland. Thus there is a lack of a generally available self-assessment tool for early assessment of DSP.

mHealth applications can offer a remote way of monitoring a patient, allowing for early detection of complications and avoiding unnecessary hospitalizations [18]. They could be applied as a medium to obtain PROs or to perform field studies. Using such applications for detecting diseases and disorders has also been a topic of research, yielding promising results in multiple fields, such as Parkinson's disease detection, automated detection of strabismus and monitoring sleep disorders [19][20][21].

A field study method - like an mHealth application cannot provide a definite level of diagnosis, however, a self-assessment measure could be beneficial in improving and accelerating the diagnostic process of polyneuropathy and DSP. A self-administered electronic tool grading neuropathy symptoms and findings would be an accessible form enabling early detection and intervention by clinicians. The fundamental challenge in this is enabling the patients to be the user of the state-of-art measures, instead of the clinician. It requires careful and precise instructions and questions, to ensure the patient can perform the examination as accurate as possible, without any medical training.

1.1 Research Question

The research question of this project is as follows: What is the design of an mHealth self-assessment tool for grading symptoms and findings in neuropathy patients, meant to assist in the early detection and diagnosis of polyneuropathy?

In order to answer this question, a number of sub-questions were formulated. Firstly, what are the state-of-art neuropathy diagnostic tools and methods? Then, having obtained the necessary background knowledge, the next question to ask is what are the tasks and questions that the patients can perform and answer themselves with accuracy and consistency? Leading from that, what is the design of a mobile PRO tool that will allow for patients to carry out a neuropathy exam via self-assessment? What is the grading score for this tool, for its result to provide similar value to other state-of-art options? Having answered the previously stated questions, lastly, does it perform comparably to existing clinical measures?

1.2 Research goals and methods

The goals of this project are:

- Review existing work in the field of neuropathy diagnosis and assessment, as well as Patient-Reported Outcome (PRO) related to polyneuropathy.
- Design a new grading tool for peripheral neuropathy inspired by state-of-art measures, that can be used by patients.
- Employ user-centered methods for the design and usability evaluation of the mHealth application containing the new tool.

- Implement a mobile application for self-assessment of polyneuropathy with the new tool.
- Run a small validation study involving neuropathy patients.

Reviewing the existing work will include researching the methods and tests used when making a neuropathy diagnosis. The research will be carried out using the forward snowball method, starting from materials provided by the clinical supervisor. The signs and symptoms of peripheral neuropathy will be described, providing the necessary medical background. We will present the existing measures for neuropathy and their application.

We will explore the existing self-assessment tools in the project domain. The research will be focused on tools that use mHealth applications, but other measures like questionnaires will also be investigated. The search will be carried out using scholar search engines - DTU Findit and Google Scholar, using combinations of the following key words: ”neuropathy”, ”self-assessment”, ”mobile application”, ”measure” and ”PRO”.

The new grading tool meant to be used by patients will be designed based on a state-of-art measure. We will consult with the clinical supervisor and assess which tests can be carried out and which should be altered or removed. Then, we will design questions and instructions for the patients to be able to perform the examination in a manner as close as possible to how clinicians would perform it. The new tool will be formed as a questionnaire.

We will design an mHealth mobile application for the new grading tool. In this process, UCD methods will be used to ensure usability of the tool. The design will be carried out in an iterative manner, using prototyping software and then implemented application. The research methods will be described in detail, as they are the focal part of the project. We will present iterations and important learnings that come with shifting the user of an evaluation tool from the clinician to the patient.

The designed mobile application will be implemented using Flutter. The grading tool will be available both in English and Danish. The structure and implementation will be driven by the Copenhagen Research Platform (CARP) Research Package (RP).

A validation study will be performed to observe and evaluate the new grading tool compared to a different state-of-art measure. It will be performed in a clinical setting, and will involve patients of the neurology department, including neuropathy patients. The patients will use the implemented application by themselves, without assistance. We will observe the process and then perform an examination using a different measure, to compare the results. We will discuss the findings and the potential of the application to be used by patients outside of the clinical setting.

1.3 Thesis overview

Chapter 2: Background and Prior Work - The medical background of neuropathy and some of the measures used in its diagnosis are explained. Other tools that have been developed to allow self-assessment of neuropathy are briefly reviewed as well.

Chapter 3: Research Methods - This chapter explains how the problem of designing an app suitable for self-assessment of neuropathy was handled, including how the questionnaire was designed and how potential users were involved in ensuring that it could be completed without the help of specialists.

Chapter 4: Mobile Health Application for Self-Assessment of Neuropathy - Here, the design, flow and implementation of the self-assessment tool as a Flutter app is presented.

Chapter 5: Validation Study - A study was carried out to validate the scoring system by comparing it to a commonly used measure of neuropathy. This chapter explains how the study was carried out and displays the results.

Chapter 6: Discussion - The research sub-questions, study results and other important topics of the self-assessment tool are reflected upon, putting the project into perspective.

Chapter 7: Conclusion - This chapter summarizes the project, results, and learnings.

2 Background and Prior Work

To provide an overview of the project and the domain, medical background on polyneuropathy is presented in this chapter. To understand the procedures and examinations used for polyneuropathy diagnosis, it briefly introduces the disease and the elements of examinations that are performed in various diagnostic measures.

Then, a summary of diagnostic tools for polyneuropathy is given, followed by a section reporting existing tools for self-assessment that are applicable in the neuropathy field.

2.1 Medical background

The peripheral nerves are responsible for communication between the central nervous system and peripheral receptors or effectors - for example skin and muscles. A nerve fiber is the long part of the nerve cell, called an axon. Nerve fibers are protected by Schwann cells and are surrounded by connective tissue. Each peripheral nerve contains multiple nerve fibers, both nonmyelinated and myelinated. Myelinated fibers are enclosed in multiple myelin sheets, which increases the fiber isolation and increases signal transportation efficiency [22].

DSP, polyneuropathy in the axon, is a result of interrupted function of peripheral nerves. It can be caused by a toxic or metabolic factor, such as diabetes, alcoholism, chemotherapy and more [3][23]. Both myelinated and nonmyelinated fibers can be affected [24][25].

Figure 2.1 shows the possible findings of DSP in diabetes. The myelin sheets and Shwann cells break down, which disrupts signal transportation. In nonmyelinated fibers Shwann cells can still detach. Eventually, the disease causes axonal degeneration that progresses from the ends of extremities proximally [24][23]. In drug-induced neuropathy - caused by for example chemotherapeutical agents, similar findings are also observed, with or without axonal degeneration. Nerve fibers can regenerate the myelin sheets, but the prognosis of recovery and disease management depends on the sustained damage [23].

In the United States, conservative estimates reckon that DSP is affecting over 20 million people, more than 10% of the population [5]. The disorder is more prevalent and burdensome with older age, and depends on the duration of disease and its management [26][27]. Symptoms range in severity, and usually progress from the end of extremities towards the body. Patients report experiences such as numbness, tingling, pins and needles, and pain. Since the neural transmission is disturbed, loss in motor functions, diminished reflexes and abnormal nerve conduction can occur [3][4][5][6].

Medical terms and various procedures that are used in diagnostic tools and examinations are described in detail in this section, for better understanding of the domain of the project.

2.1.1 Pin-prick Sensation

Examining the pin-prick sensation is used to determine sensory loss in different areas of the body. A clinician demonstrates the test to the patient, and then gently pricks multiple areas of patient's skin with a sharp object. The patient is then asked if the sensation was sharp or dull [28].

The sharp object used can for example be a safety pin or a sharp wooden pin. Oftentimes, a reference area is used so the patient can compare the sensation. Since polyneuropathy affects distal nerves first, the comparison is made against sections of skin closer to the

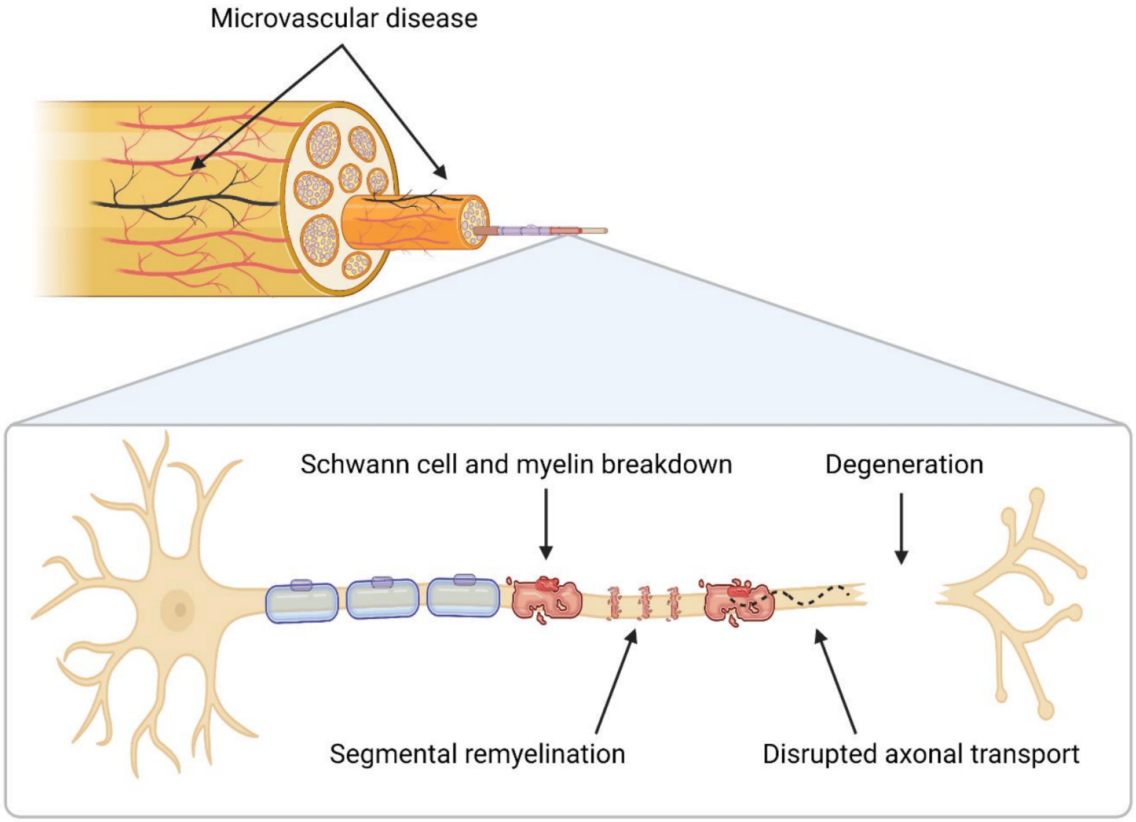


Figure 2.1: Neuropathological findings of distal symmetrical polyneuropathy in diabetes [24].

central nervous system, for example the clavicle or neck. If the patient is affected by DSP, they may feel reduced sharpness in extremities, or not feel the prick at all.

2.1.2 Large Fiber Sensation

Large fibers are responsible for both cutaneous touch sensation and sense of movement and position [29]. Cutaneous senses include touch, pressure, vibration, temperature and pain reception [30]. Abnormalities in this sensation are measured with joint position and vibration tests. The most distal joints are tested first, as they are affected before more proximal ones. If testing on the distal joint finds an abnormality, more proximal joints are tested successively until reaching a joint with normal sensation [31].

For the joint position test, the patient closes their eyes. The clinician moves the joint up and down, while the patient reports every movement. An alternative method is to examine when the movement is perceived, by starting with moving the joint slightly and gradually extending the range of movement [31].

The vibration sensation test should be done over joints or areas where bones are close to the skin. A tuning fork is used and the patient reports if they can feel the vibration - often described as a buzzing sensation, and when they can no longer feel it [31]. The tuning fork used is usually 128Hz, less often 256Hz [32].

2.1.3 Nerve Conduction Study (NCS)

A Nerve Conduction Study (NCS) is a tool broadly used in evaluation of the peripheral nervous system, assessing large myelinated nerve fibre function [33]. It measures various parameters of nerve and muscle action potentials. They are used to ascertain the number

of functioning nerve fibers and speed of conduction. Electrodiagnostic patterns involving those measures assist in discerning the type of pathophysiology affecting the nerves [34].

Two types of NCS are usually carried out: motor and sensory. They are performed with an external handheld stimulator that is stimulating the nerve, while surface recording electrodes record the potentials [34].

2.1.4 Deep Tendon Reflexes

Peripheral neuropathy is the most common cause of absent reflexes [35]. When a muscle tendon is tapped, the muscle immediately contracts. Presence of these reflexes can be influenced by a variety of factors including patient's age, sex and examiner's technique [36]. In a screening examination of deep tendon reflexes, multiple points can be tested. The commonly examined reflexes are in the ankle, knee, finger, wrist, biceps, triceps and jaw [35]. The tendon of muscles in these points are briskly tapped with a tool such as a reflex hammer.

The best results are acquired when the patient is relaxed, and not thinking about the examination. There are multiple techniques the clinician can employ if the reflex is hard to obtain - for example changing the limb position or having the patient put slight tension into the muscle. When a reflex is present on one side, the other side is examined immediately, as asymmetrical reflexes can suggest abnormality [35].

2.1.5 Motor Examination

Muscle weakness in extremities can be a sign of polyneuropathy, and a standard neurological examination contains manual muscle testing [37]. Motor strength can be tested by instructing the patient to oppose the clinician's force on part of the body [31]. For example, to measure great toe extension - dorsiflexion, the examiner will press down on the toe, while the patient tries to overcome that pressure and extend the toe upwards. This test can be performed on various points on the extremities, such as ankles, knees, fingers and wrists, to assess the degree and areas of the weakness.

2.1.6 Other indicators

Paresthesia is a term describing an unexpected sensation of burning or pricking. It usually appears in the extremities. It can also be described by patients as tingling, numbness, skin crawling or itching. Paresthesia can be temporary or chronic and stem from multitude of neurological diseases [38].

Hyperesthesia is a term describing increased sensitivity to stimulation including touch, thermal sensation and pain. It's a common symptom of neuropathic pain [39]. To describe hyperesthesia for a particular stimulus, Allodynia and hyperalgesia can be used. Allodynia refers to pain evoked by a stimulus that usually does not cause pain - for example light touch, while hyperalgesia is an increased response to a stimulus that normally causes pain - for example pricking [40].

2.2 Diagnostic measures for peripheral neuropathy

Although there is a variety of available measures of neuropathy, they differ in which phenotype of DSP they are best for. To address the problem of early detection of neuropathy, measures aimed at Small Fiber Neuropathy (SFN) are of the biggest value to us, since studies suggest that SFN is the earliest nerve damage. As stated in chapter 1 it is recommended to verify the presence of symptoms, signs and an abnormal NCS result before diagnosing DSP. A common issue with this recommendation is that an NCS tests the large fibers, meaning that it will often give a normal result in patients with early neuropathy. For this reason, measures of early neuropathy will generally leave out the NCS [33].

Below we introduce a subset of the common clinical measures that are used in assessing neuropathy. Most of the measures don't offer a cut-off point for diagnosis, leaving it up to the specialist to interpret the result.

2.2.1 Utah Early Neuropathy Scale (UENS)

UENS is a physical examination focusing on detecting and quantifying early signs of SFN in Diabetic Peripheral Neuropathy (DPN) [41][5]. An ACTTION systematic review of measures for DSP [5] found that 62% of the total UENS score is dedicated to the domain of SFN, with a minor focus on large fiber, reflexes and motor assessment. The examination requires a safety pin and a 128 Hz tuning fork, and consists of 5 examination areas:

1. Motor Examination (0-4 points)
2. Pin-prick Sensation (0-24 points)
3. Allodynia/Hyperesthesia (0-2 points)
4. Large Fiber Sensation (0-8 points)
5. Deep Tendon Reflexes (0-4 points)

The motor examination is performed on the great toes and 2 points are scored for each toe with notable weakness. A safety pin is used for measuring pin-prick sensation, which starts with testing a reference area to set a baseline. Then the sensation is measured in legs, starting with the top of the great toe and gradually moving up the foot and leg. The results are noted for 6 sections in each leg, scoring 2 points per section with absent sensation, and 1 point where it's diminished. If allodynia or any hyperesthesia is present in feet, 1 point per foot is added.

Large Fiber Sensation is examined in the great toes through vibration and joint position tests. The criteria for vibration are: 2 points per toe with absent sensation, and 1 point if the sensation diminishes in less than 10 seconds.Great toe joint position sensation can be determined as absent (2 points per toe) or diminished (1 point). Finally, deep tendon reflexes are tested in the ankles, each scoring 1 point if diminished, 2 if absent.

In each category both the left and right leg are examined, and the highest possible score is 42 points. Out of that, over half of the score is dedicated to the pin sensation. UENS across 20 comparisons is reported to have a 94% inter-rater reliability, as well as high specificity and sensitivity. No threshold recommended for diagnosis is provided [41].

2.2.2 Total Neuropathy Score (TNS)

TNS is a well established neuropathy measure testing both the legs and arms. It focuses on DPN and CIPN neuropathy types, offering a broad testing area not focused on a particular phenotype [5]. The components in this scale are:

1. Sensory Symptoms (0-4 points)
2. Motor Symptoms (0-4 points)
3. Pin-prick Sensation (0-4 points)
4. Vibration Sensation (0-4 points)
5. Motor (0-4 points)
6. Deep Tendon Reflexes (0-4 points)
7. Nerve Conduction Test (0-8 points)

Unlike UENS where individual scores are given for both legs, each component of TNS only records the result of leg or arm with the highest number of points.

Presence of sensory symptoms (tingling, numbness, neuropathic pain or cramps), decreased vibration and pin-prick sensations are scored based on most proximal point of

the limb where they occur: in fingers or toes (1 point), up to wrist or ankles (2 point), up to elbow or knee (3 points), beyond the elbow or knee (4 points).

Motor symptoms are assessed by the level of difficulty the patient has using their feet, legs, hands and arms in daily life: slightly difficult (1 point), moderately difficult (2 points), requires assistance (3 points), or the limbs have lost function (4 points).

The motor strength is evaluated with ankle dorsiflexion, finger spread and wrist extension and graded by the weakest of the three: mild weakness (1 point), overcomes gravity but not resistance (2 points), only moves when gravity is eliminated (3 points) or is paralyzed (4 points).

Reflexes are tested in the biceps, triceps, brachioradialis, knee and ankle. The scoring for this segment is: ankle reflex is reduced (1 point) or absent (2 points), ankle reflex is absent and other reflexes are reduced (3 points), all reflexes are absent (4 points).

Lastly, an NCS is performed, giving individual scores for motor and sensory study. The score is based on what percentage of normal value the result is: 75-95% (1 point), 51-75% (2 points), 26-50% (3 points) or below (4 points).

The highest score possible is 36 points, including PROs - the first two components. The test cannot be conducted remotely, due to the NCS, which requires a physical clinical setup. A common critique of TNS is that it weighs symptoms and signs in the arms equally with those found in the legs. As stated in chapter 1, neuropathy usually advances in the legs much earlier than in the arms [42].

TNSc is an alternative version of TNS that excludes the NCS, eliminating the need for specialized equipment. It adds autonomic symptoms to the measure instead. Up to 4 points are possible, 1 point is scored for each of the following symptoms the patient experiences: loss of bowel and bladder control, fainting, impotence and constipation. This makes the maximum score for TNSc 32 points [42]. A study on CIPN patients found TNSc to be a reliable measure, and a valid alternative to other TNS versions when an NCS is not feasible [43].

2.2.3 modified Toronto Clinical Neuropathy Score (mTCNS)

Another measure targeted at early DSP in the legs is the mTCNS. It consists of two parts, each with a set of sub-questions:

1. Symptom scores (0-18 points)
2. Sensory test scores (0-15 points)

Akin to TNS, each component score is decided only by the area that gives the highest score.

For the symptoms score the patient is asked how much their well-being and daily life is affected by the following symptoms in legs: foot pain, numbness, tingling, weakness, reduced muscle control, and any upper limb symptoms. Each symptom is scored if present: without interfering with well-being (1 point), interfering with well-being but not daily life (2 points), or interfering with both (3 points).

In the sensory test score it's determined how far up the legs the following sensations are reduced: Pin-prick, temperature, light touch, vibration and joint position sense. Each sensation is scored for reduced presence: in toes only (1 point), up to the ankles (2 points), reduced above the ankles or absent in toes (3 points).

The maximum score of mTCNS is 33, and the measure was found to have good inter-rater reliability [44].

2.2.4 Douleur Neuropathique 4 Questions (DN4)

In order assess the presence of neuropathic pain, the DN4 questionnaire was developed consisting of two parts:

1. Patient interview (0-7 points)
2. Patient examination (0-3 points)

In the interview the patient is first asked whether their pain can be described as feeling like burning, cold or electric shocks, scoring one point for each they agree with. Then, 1 point is scored for each of the following symptoms felt in the same area: tingling, pins and needles, numbness and itching.

In the examination, the patient is tested for allodynia and hyperalgesia, scoring 1 point for each that is present. Finally, if the pain can be provoked or worsened by stroking with the hand, 1 point is scored.

The maximum 10 points, and a score of 4 or above indicates neuropathic pain. The DN4 questionnaire showed good ability to discriminate neuropathic pain from non-neuropathic pain [45].

2.3 Existing tools for self-assessment of neuropathy

This section presents a short report on tools for unsupervised self-assessment of neuropathy and their application. The tools include both mobile applications and other forms of obtaining the PRO.

2.3.1 NeuroPad

NeuroPad [16] is a plaster you put on the sole of your foot for 10 minutes, after which it will change colors from blue to pink, if the foot's ability to produce sweat is normal [46]. The usefulness of Neuropad as a tool for diagnosing DSP has been the subject of many studies, and a study on its value as an early detector of neuropathy found it to have a sensitivity of 83.33% and specificity of 68.04% [17].

Unlike the self-assessment tool we will develop throughout this project, Neuropad is a physical product that must therefore be purchased and delivered to the user. The plaster is available for purchase online for residents of the UK and Ireland, but they are not available outside of that region.

2.3.2 Mobile application for self-management of neuropathy

A Korean mHealth application [47] was developed as a self-management program for CIPN patients. It uses various questionnaires to monitor the patient and assess CIPN's severity, negative impact on daily life and reduction in quality of life. Additionally, the application functions as a learning tool about CIPN and as a coach, suggesting exercises for relieving symptoms.

The reliability of the self-assessments were not studied since the focus of the paper was on the value provided by the application rather than its precision. We did not find evidence of the app being publicly available or that further research has been carried out.

2.3.3 NeuroDetect

NeuroDetect [48] is another mHealth application aimed at patients with CIPN which can be used for assessing the severity of the condition. The application makes an assessment

based on two types of functional tests, one testing the strength and dexterity of the hands, the other testing gait and balance.

A partial least-squares discriminant analysis of NeuroDetect results showed good ability to differentiate between CIPN patients and controls, but . The app is available on the Apple App Store [49], but it has not been subject to further studies.

2.3.4 PeriVib

PeriVib [50] is the combination of an mHealth application and a portable vibration device which enables assessment and diagnosis of CIPN. The vibration device is used for determining the vibration sensitivity threshold of the patient's feet with vibration tests, and the mobile application for performing gait and sway functional tests to assess presence of CIPN.

The vibration threshold test was correlated with a biothesiometer and tuning fork, with coefficients $R^2 = 0.68\%$ and $R^2 = 0.15\%$ respectively. The motor examination - balance and gait test correlated poorly with the threshold test and severity of CIPN. Neither the vibration device or the app are publicly available, and no further studies on the tool have been carried out.

2.3.5 DASH Outcome Measure

DASH Outcome Measure [51] is an extensively studied 30-item self-evaluation questionnaire for assessing muscular-skeletal disorders in the upper-body extremities. A study was carried out to validate the usefulness of DASH for assessing neuropathy in the elbow [52]. The study found that DASH scoring correlates well with clinical assessment of neuropathy in the elbow.

3 Research Methods

Creating a PRO grading tool that is suitable for assessing DSP is a major challenge, since the examination must be implemented in a way where even first-time users can conduct it accurately and confidently without guidance from a trained specialist. Therefore the tests, inspired by measures created for clinicians, should be easy to carry out and have unambiguous instructions.

For that reason, the design focus of this project has been on the set of tests used for grading, how they should be performed and how they are conveyed to the user. Medical doctors specialized in the field played a big role in formulating the questionnaire and providing feedback over the course of the project, and UX design methods have been employed in order to ensure that the self-assessment tool caters to the users. Figure 3.1 shows when during the design process we received feedback from doctors consulting on the project and when UX methods were used to validate the design.

The survey was developed in both Danish and English at the same time. The Danish version that was the one validated, while the English version followed the Danish phrasing.

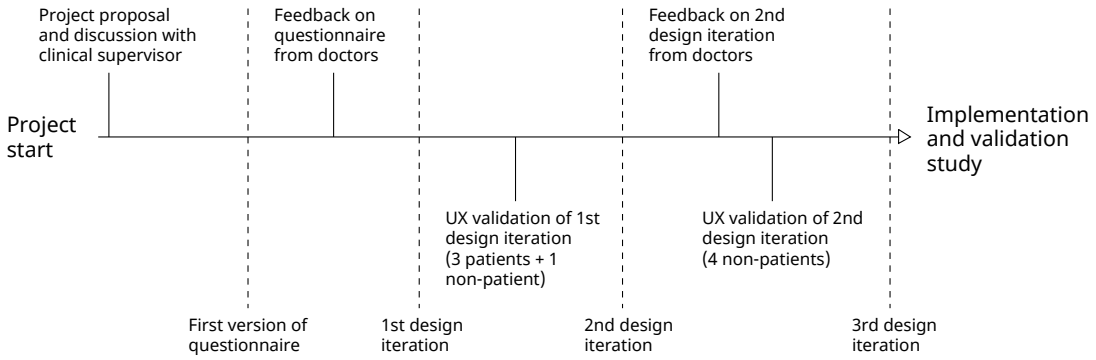


Figure 3.1: Overview of feedback and validation throughout the development of the questionnaire.

3.1 User Experience (UX) design methods

In essence, UX design is a process that is concerned with providing users a satisfactory experience with a product. For apps, many parameters contribute to the experience, including ease of use, enjoyment, efficiency and usefulness, but that is just a small part of what can constitute a good user experience. The main challenge in UX design is that it is very difficult to engineer an experience for other people [53].

Unless explicit measures are taken to validate and amend assumptions made about the users, there's a great risk for the resulting app to be unintuitive, ineffective or to simply be left unused [18]. For this reason UX design relies heavily on User-Centered Design (UCD) methods to ensure that the designs lead to a pleasant experience for users.

UCD is an approach to design where potential users are involved in the design process of an app in order to validate assumptions and get feedback and suggestions on the UX. They aren't given direct control over design choices, rather, their involvement is used by developers to learn about the target users, improve existing designs as well as guide future

ones. The participants in UCD should be representative of the target users of the app, and if they are included early in the design process, many design flaws can be corrected before they become too ingrained in the product. There exists a plethora of user-centered methods that are applicable in various phases of the design process, and each contribute with different insights [18][54][55]. Two of the most common user-centered methods will be explained further into this section after an introduction to iterative design, one of the cornerstones of UCD.

3.1.1 Iterative design

Iterative design is a methodology that acknowledges that designs rarely are, and in general cannot be expected to be, perfect at the first attempt. It instead focuses on testing, learning and then improving the designs iteratively. Testing is done by first creating a prototype that represents the current design or, in the late stages of design, using the functioning app itself. Then potential users are presented with the prototype and get to experience it in a specific way defined by the employed method. Learning happens during the tests either through observing, listening, asking questions or discussing, or a combination of those. Finally, the design is improved by using the insights gained to guide the development of the next iteration, after which a new cycle of testing, learning and improving begins. This pattern is repeated as long as time allows or until the results are satisfactory [55][56].

Design validation may lead to unexpected findings that can only be addressed by making drastic changes. Therefore it's inadvisable to spend a significant amount of time on implementation or grow too attached to the design without prior validation. Preferably the first iteration of a product should be a minimalistic implementation of the idea that allows for gradually building on top of it, otherwise known as a Minimum Viable Product (MVP). The premise of MVP is to distill the application you aim to create down to a simplified version of it. This allows the core concept to be validated before effort is put into less essential features like graphical design and login screens. The MVP doesn't need to be a functioning app implemented in code, but could instead be represented with a mock-up made with pen and paper or wireframes, as long as it accurately conveys the core idea [55].

A wireframe is a visualisation of all the possible states of an app, usually represented as a set of screens to the like of those you'd find in a finished product. Some common ways of creating wireframes are drawing them by hand or using a digital design tool. The screens should contain only the essential elements that the app would have, such as buttons, forms, images and relevant text [55].

3.1.2 Usability testing

Usability testing is a qualitative method used for the testing and learning phases of iterative design. Potential users interact either with a prototype of the app, or the app itself in the current iteration, while an observer is present to make note of how it is used. Testing is performed with one person at a time, and they are asked to carry out one or more tasks on their own while articulating their thoughts and actions; *thinking aloud*. If the app has a linear flow, a task could be to complete said flow. For a more complex case, a task could be to use a single feature to achieve a specific goal. The benefit of usability testing is that developers can discover how first-time-users perceive and interact with the app, and this method is likely to reveal the parts of the product that need UX improvements. It's common to use only around four people each iteration, since that is often enough to find the current major issues [54].

The main challenge of this method is that the observer must restrain themselves from

helping the user or answering questions, unless absolutely necessary. Instead a note should be made of the difficulties and questions that are asked, as well as how the user deals with problems they encounter. This part is crucial, as it provides insights into the main UX issues that could prevent users from successfully using the app on their own[55].

While usability testing on its own rarely provides immediate solutions to problems that it has revealed, discussing with the user after the test can be a good way to look for them. The discussion should be carried out with caution and attention to how questions are formulated. If the goal is to generate ideas it's important to avoid yes-no-questions as they prompt the user to validate solutions that have already been thought of, while open-ended questions encourage them to come up with their own. Examples of informative, open-ended questions that are applicable to most products are how they feel the test went, asking about points where they struggled, and what could improve the experience. There may also be other questions worth posing depending on the specific product. Additionally, asking the user to elaborate on their replies, and investigating their train of thought can provide invaluable information [55].

3.1.3 A-B testing

A-B testing is a quantitative method that can be utilized to decide between two or more design options. There are multiple ways to carry out an A-B test, and an example is to simply show the options to users and ask them which they prefer. This could be done by showing pictures of the options or by letting the user interact with prototypes of the options. With enough tests performed a preference for one of the options, or lack thereof, should become visible.

It's imperative that the options are presented in a non-biased way without signs of developers' preference for any of them. Randomizing the order of shown options greatly helps ensuring this is the case. Two common ways of conducting this test is either using a survey or doing it in person, both types having certain benefits. There exists many online tools for making and distributing surveys, making it possible to accumulate a large quantity of responses. It's generally more difficult to acquire participants for in-person A-B testing, however, a qualitative element can be added to each response by asking them to elaborate on their preferred choice. As with usability testing, this means asking open-ended questions in order to start a discussion, while taking care to avoid bias towards existing ideas [55].

3.2 Grading tool design

The initial foundation of the designed grading tool is the UENS metric, primarily owing to its emphasis on the SFN and early neuropathy. In its base form, the UENS scoring sheet is not easily readable for non-medical personnel and patients. It also requires specialist tools not commonly found in a household, like a tuning fork. An essential point of the design was developing step-by-step instructions to provide accessible and understandable tasks that patients could conduct themselves.

All of the discrepancies from the original measure were consulted with medical doctors specialized in the field, in order to ensure the most medically accurate results possible. The objective was to develop a grading tool that encompasses a comprehensive range of assessments congruent with the UENS in order to maximize its applicability and effectiveness.

It was decided that the questionnaire will contain questions and tasks in two categories - graded and non-graded. Responses in the non-graded part can still have individual scores, though they do not contribute towards the overall examination score. It was done

Table 3.1: Sections of the new grading tool and corresponding point ranges.

Section	Summary	Item count	Point Range
	Graded Sections		
General Symptoms	Symptoms in the legs and their severity	1	0-4
Pin-prick Sensation	Investigation into pinprick sensation in 6 sections in both legs	12	0-24
Allodynia/Hyperalgesia	Painful sensation to prick or touch	4	0-4
Large Fiber sensation	Feeling the vibration in three points in both legs	6	0-6
	Feeling changes in great toe joint's position	2	0-2
Motor Examination	Overcoming pressure on the great toes	2	0-4
	Non-Graded Sections		
Neuropathic Pain	Pain Level	1	1-100
	Painful Neuropathy (DN4) Questionnaire	4	0-10
Comments	Any additional information patient wishes to disclose	1	N/A

so for the tool to not only provide the score itself but also context around it - such as if the patient is also experiencing neuropathic pain or has any additional comments to share. An overview of the main sections of the questionnaire, along with the amount of points they can contribute to the score is presented in Table 3.1. Those sections are also referred to as *"parts"* or *"segments"* of the questionnaire.

The maximum points possible to obtain when using this score is 44, comparable to the 42 points in the UENS. The greater the number of points, the higher the likelihood of indicating neuropathy. The scoring in the sections is mostly consistent to that in UENS, even if the methods or questions vary, in aim to maintain a similar emphasis on early neuropathy detection. Thus still over half of the points is dedicated to investigating SFN in the form of pinprick sensation and sensitivity.

3.2.1 General Symptoms

While this questionnaire follows the UENS closely, one section in not present at all. During an examination, a clinician tests the patient's deep tendon reflexes. These reflexes are challenging to test on your own with no medical training. Additionally, their presence varies not only due to the disease, but due to other factors as well, such as age. To compensate for the lack of this part of the examination, the new questionnaire contains a part absent from UENS, but present in other scores like TNS - general symptoms. It consist of a single question, asking if the user is experiencing any pain, numbness, tingling or unexpected sensations in the legs. If so, the user is asked to specify up to which area the symptoms are present - toes, ankles, knees or above. The cut-off level of the symptoms determines the score for the segment, from 0 - no symptoms, to 4 - symptoms above the knees.

In order to avoid mixing symptoms and findings, this question has been placed first in the examination. This was due to a concern that if a user was asked such question after a different section, they might drive their answer by what they recorded there. For example, if they learned that they feel reduced pinprick sensation up to the knee, that could influence their choice in this section.

3.2.2 Pin-prick Sensation

In order for the user to be able to perform an investigation of pin-prick sensation in their legs, a fairly sharp, easy to use object is necessary. Ideally, the item would be also accessible and possibly present in the user's home. Three household items were discussed: a toothpick, a needle and a safety pin. After studying them and the variety of shapes and

(a) Pin-prick sections in UENS [41] (b) Pin-prick sections for the project.

Figure 3.2: Comparison between pin-prick sections.

sizes they come in, it was determined that a substantial portion of toothpicks may be too blunt to provide the required sensation. With that result, a design decision was made to use safety pins or needles as the tool for the examination.

UENS grades the presence of the pin-prick sensation in six areas, and uses an image presented in Figure 3.2a for reporting the findings. While informative for the clinician, the section layout may not clearly communicate which area of the leg the user should be investigating. The segments for this questionnaire were iterated on and validated in the design process, which is elaborated on in section 3.3 and section 3.4. The final iteration can be seen in Figure 3.2b. The chosen form captures not only the areas, but also the side of the leg that should be subjected to pricking.

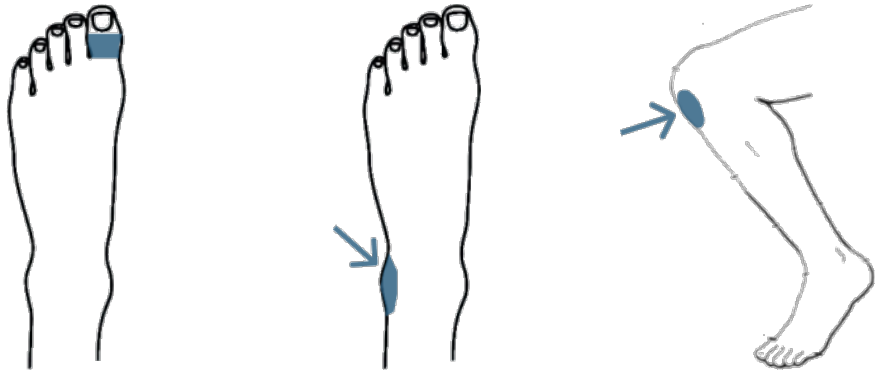
The score for each section depends on to what extent the pin-prick sensation is present. For that, a comparison to a healthy part of the body is necessary. For our tool, the recommended area it the clavicle. Since DSP affects extremities first and foremost, this area will be affected in a lesser extent. It is also accessible for the user, as clavicles are easy to reach. UENS grades each section between 0-2 points. Zero if the sensation is normal, one point if the sensation is reduced, and two points if sensation is absent. The carried out validation revealed the original phrasing: *"similar-reduced-absent"* was a source of confusion for the users. Thus, the phrasing for comparing sensation has been changed to *"same-less-none"* and seemed to have been met with more understanding.

In summary, in this section of the examination the user compares the feeling when pricking six areas in each leg to a reference area - the clavicle. This yields twelve steps to be completed, scoring between 0 to 24 points for this part of the examination.

3.2.3 Allodynia and Hyperalgesia

In the UENS 1 point is awarded if either allodynia or any hyperesthesia are present in the toes or foot of each leg. To extend the information of which types of hyperesthesia are present, we have decided to split them into two separate questions, one about allodynia and one about hyperalgesia. Thus, 4 points are possible in this segment of our grading tool, scoring 1 point per condition present in each leg.

The placement of the questions in the questionnaire was also a major point of deliberation. To avoid switching between legs too much during the examination, allodynia and hyperalgesia questions directly follow the pin-prick part for each leg. Since they refer to the just examined area, it was a proper place to ask about those hypersensitivities, while the user's memory about the sensations was still fresh.



(a) Great Toe (b) Ankle (c) Knee Cap

Figure 3.3: The points where large fiber sensation examination is carried out. Pictures present the areas for the left leg, in questions for the right leg they are mirrored vertically.

3.2.4 Large Fiber Sensation

As written in subsection 2.1.2, large fiber sensation is measured by applying vibration to the affected areas. After it is administered, the patient reports if they feel the vibration, and when it is no longer present. In UENS, large fiber sensation is only measured in the great toes. The scoring depends on how fast the vibration diminishes. As it is performed with a tuning fork, a tool that always vibrates at the same frequency, such information is a consistent measure. In mobile phones, the frequency varies depending on the manufacturer and model, varying between 130-180Hz [57]. Due to that, it was necessary to rework this section in order to make it applicable for a mobile questionnaire.

With guidance from the clinical supervisor, we have decided to measure vibration only based on it's presence, eliminating the "diminished" option from the question. To compensate for narrowing the scope of the test, we have extended the areas where the large fiber sensation test is performed. Users are to measure the sensation with their mobile devices on the great toe joint, ankle, and just below the knee cap, comparable to a test used in a different measure - TNS. an overview of how those areas were presented to the user can be seen in Figure 3.3. They remained unchanged since the first iteration (section 3.3).

The second part of this segment is examining the great toe's joint position. The users are asked to move their great toe with their hands, and answer if they are feeling the extension in the joint.

To keep a similar fraction of total points obtainable in this section analogous to the UENS, each question would score 0 or 1 points. In the vibration segment, one point would be assigned per leg area where the sensation is not present, totaling to 6 points possible (3 per leg). Not feeling the extension in the great toe would also score one point (per leg). Thus, the maximum point in the section is 8, which matches with UENS.

3.2.5 Motor Examination

Akin to UENS, motor examination is performed only for the great toes. UENS inspects dorsiflexion of the toe and determines if the strength of extension is normal or weak. The user is asked to apply pressure on top of their great toe and try to oppose it. Each question would score either 0 or 2 points, 2 being appointed if the pressure was difficult to overcome.

3.2.6 Neuropathic Pain

It was brought to our attention by the clinical supervisor, that it would be beneficial to include a section about neuropathic pain, as it would provide valuable information to the clinician reviewing the PRO output. Thus, we decided to incorporate questions from the DN4 questionnaire [58]. As the measure is clinically validated, questions in our tool follow it closely.

The only significant change is rephrasing the question if the pain is provoked or increased by brushing. Instead of simply asking that question in the questionnaire, we included instructions for the user. They are asked to gently stroke the area affected by the pain and answer if that provoked or increased the pain.

Additionally, we were asked by the clinical supervisor to include a general level of the pain, in the user's perception. To do that, the question in this segment provides a visual analogue scale [59] from 0 to 100, on which the user marks their pain level.

As mentioned before, the scores in this section are for context only, and they do not count towards the DSP grading score. That said, points possible to obtain in the DN4 part are exact to the tool - 0 to 10, and the pain level can be scored up to a 100.

3.2.7 Questionnaire design

Some of the design decisions have been mentioned in the previous sections. They mostly regarded phrasing, or the methods of the self-assessment. All discrepancies between the new tool and the inspiration - UENS were also presented. In this section, the point is to outline the overall design choices that concern the new grading tool specifically.

Firstly, the order of the questions and sections was a topic of discussion. It was important for it to be as smooth and accessible an experience for the user as possible, while still providing valuable and unbiased findings. The structure of the questionnaire is as follows:

1. Introduction to the self-examination
2. General Symptoms
3. Instructions for pin-prick
4. Pin-Prick - left leg (6 questions)
5. Hyperalgesia and Allodynia - left leg (2 questions)
6. Pin-Prick - right leg (6 questions)
7. Hyperalgesia and Allodynia - right leg (2 questions)
8. Pain presence question
9. Neuropathic pain (5 questions, omitted if there is no pain)
10. Instructions for Large Fiber Sensation
11. Large Fiber Sensation section - left, then right leg (8 questions)
12. Instructions for motor examination
13. Motor examination - left, then right leg (2 questions)
14. Additional Comments

The General Symptoms question (see subsection 3.2.1) was placed at the beginning of the self-assessment order. The questions were ordered to avoid bias, tackle the largest parts first, and produce a coherent flow. Thus, the Pin-Prick section (subsection 3.2.2) follows next, intertwined with Allodynia and Hyperalgesia parts (subsection 3.2.3) to avoid constantly switching leg sides. Since this examination can trigger or increase pain, the pain section (subsection 3.2.6) is placed afterwards. Then, the user moves onto Large Fiber Sensation (subsection 3.2.4) and Motor Examination (subsection 3.2.5), which are separated by instructions. Finally, the last screen of the grading tool is the Comments

section, providing an input opportunity for the user to elaborate on the experience or add any additional information.

Instructional and informational parts separate major sections, and introduce them to the user. For example, the Pin-Prick section instruction informs that a needle or safety-pin will be necessary for the next steps. In the Large Fiber Sensation section, users are directed which side of the phone to use when investigating, as well as that the test should be performed on bare skin. These instructions are necessary to provide context and relay important notes to the user, while putting them in separate segments avoids unnecessary repetition in the examination segments themselves. Despite that, some of the crucial instructions are repeated at every step of the assessment, as during the validation process that yielded the best results in correctness of performed tasks.

The Pain presence question asks about whether the user experiences pain in feet. While the DN4 questions (subsection 3.2.6) inquire about pain in general, if the user was asked about any pain, it could be answered *"yes"* for other non-neuropathic pain, for example headaches. After discussions with the supervising clinician, it was decided that asking about presence in feet makes it more likely that only users with neuropathic pain will go through the section.

Lastly, the Large Fiber Sensation segment is called and referred to as "Vibration Part" in the tool. It was decided that it is a more user-friendly language, as well as relates more to what is performed - investigating if the vibration is felt in the affected areas. Across the tool the wording and naming of sections and questions was altered or omitted entirely, if it was of no significance for the user. For example, the Hyperalgesia or Allodynia were not mentioned, as they are highly specialist terms. Instead, the grading tool contains questions about increased sensitivity to touch or prick, which encapsulates those disorders. Those professional terms are still relevant for the physician reviewing the results of the self-assessment, so we took steps to ensure that exported data properly describes the results.

3.3 First design iteration

UX and methods from UCD have been in the forefront of the design process during this project. The app has gone through three design iterations, starting with an MVP based on the first version of our questionnaire and feedback from medical specialists. At the end of the first iteration, user validation was carried out to discover issues with the design and gather general feedback. What we learned was then used to guide the second design iteration and improve the UX.

3.3.1 Minimum Viable Product

The first iteration was created as a wireframe mock-up with Figma [60], an online tool that can be used to design high-fidelity wireframes and turn them into interactive mock-ups. First, the screens for the individual app states are made with their buttons, text and any other content. The states can be connected with flow lines - interactions designed to happen between the elements, such as clicking button *X* will navigate to page *Y* or show picture *Z*. With these in place, Figma can go into Present Mode, where a single state is displayed and interacting with the components will change the displayed state, as defined by the flow lines. This allows emulating simple app flows, giving the impression of using a real app.

The MVP consisted of three components. Firstly, a *main page* from where a new examination can be started. Then the examination itself, consisting of the parts as listed in subsection 3.2.7 aside from a few differences. Originally, Additional Comments didn't

exist and General Symptoms came after the second set of Allodynia and Hyperalgesia. The third and final component was the *Examination Completed* page displaying a generic score. While designing the MVP, text on the screens was written in English and a parallel version was created with every element translated to Danish.

3.3.2 Validation goal

The goal of the first validation was to test the basic premise of the project, a self-assessment tool for neuropathy. The most important thing to uncover was whether the test participants would perform the examination as intended based on the instructions and could confidently answer the questions. This was investigated with the usability testing method by utilizing the mock-up made with Figma.

The *Examination Completed* page used in the MVP was a simple placeholder, so another goal was to get feedback on our ideas for possible implementations. For this reason, in-person A-B testing was performed on participants after the usability testing.

3.3.3 Validation participants and procedure

Four participants were recruited for the validation, of which three were patients from the outpatient clinic at the Neurological Department of Zealand University Hospital and one was not a patient. The participants ages were 45, 54, 65 and 69, and there was an equal number of men and women.

The usability testing was carried out in Danish with the the Danish MVP and using a predefined validation procedure. The procedure was as follows: First the test was explained to the participant, and permission was asked to record the test. Next, they were asked their age and were then provided with a smartphone displaying the interactive prototype as well as a safety pin. Finally they were asked to carry out the examination while thinking aloud, followed by questions and discussions as described in section 3.1. Participants were given permission to skip repetition across legs after having been observed performing a test on one leg, since the instructions were identical for both legs. The questions they were posed afterwards were asked in Danish, and listed below are their English equivalents:

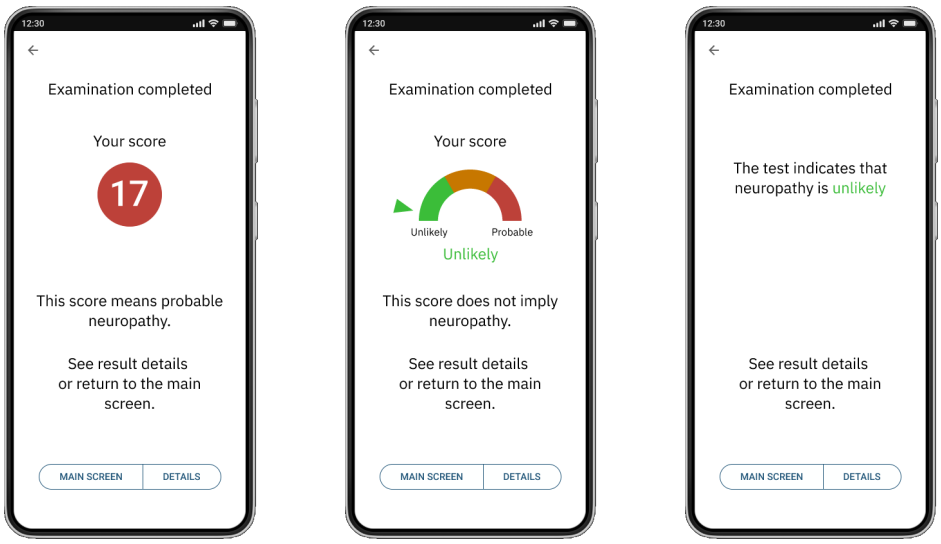
- How did it go?
- How easy were the instructions?
- Were there any difficulties?
- What do you think of the button size?
- What do you think of the text size?
- What would improve the experience?

After concluding the usability testing, we moved on to the A-B test. Here the participant was presented with three Figma screens of possible designs for the *Examination Completed* page, see Figure 3.4. The options were shown in a randomized order, and they were asked which they prefer and then elaborate on their choice.

3.3.4 Validation learnings

When the four participants had gone through the validation, the notes and recordings taken were analysed and collected to a set of learnings.

Carrying out an examination took around 10 minutes on average, but it should be reiterated that participants were allowed to skip repetition across legs. They were all positive about the potential of the self-assessment tool, but it was clear that perfection was not reached in the first iteration.



(a) Option A (b) Option B (c) Option C

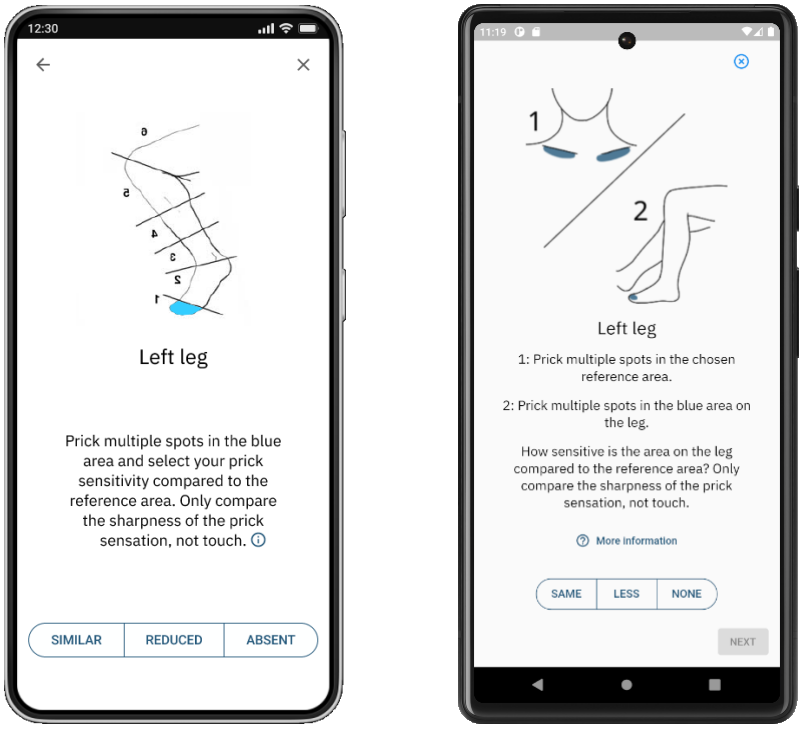
Figure 3.4: The three versions of the *Examination Completed* page that were subjected to A-B testing

The largest issue that emerged was rooted in instructions for the pin-prick test. Only one of the participants consistently used the reference point while the rest either forgot it for a large portion of the test or didn't notice the instruction at all. Also, a question mark icon that was intended to provide additional information was not used, even when the participants were in doubt. Occasionally, participants didn't notice which leg they were instructed to use. We also realized that some people may not be able to tell left from right without clear visual indication. Aside from these issues, pricking was generally performed as intended and in the correct locations.

There were some areas that were not clearly communicated in the prototype. Many participants were uncertain if they were supposed to take off shoes and socks, and those who wore tight pants were caught off-guard by the need to gradually roll their pant legs further and further up. Rolling up past the knee wasn't possible, forcing them to remove that particular piece of clothing after having already gone through the trouble of rolling up pant legs.

The names of the three answer choices in the pin-prick test also led to some confusion. Since it was the Danish version of the self-assessment tool being validated, the choices *"similar"*, *"reduced"* and *"absent"* had been translated to the Danish *"tilsvarende"*, *"reduceret"* and *"ingen"*. The Danish *"tilsvarende"* and *"reduceret"* are not as commonly used as their English counterparts. While the words were not unknown to the participants, some felt that they had to map the answer they wanted to give to one of the options available. One of the participants suggested using the Danish *"samme"* and *"mindre"* instead, which translate to *"same"* and *"less"*. The choice *"absent"* was problematic for a different reason, namely the fact that it isn't a natural response when asked to compare two things.

The A-B test found that one participant preferred option A while two preferred option B and one didn't give an answer. The reason for that was that they were very concerned that the presented *Examination Completed* pages looked a lot like a diagnosis. Despite not using any certainty levels higher than "probable", nor language implying a diagnosis, the *Examination Completed* pages still gave the impression of obtaining one. Later



(a) Iteration 1 (b) Iteration 2

Figure 3.5: First and second design iteration of pin-prick. The difference in phone types is due to (a) coming from Figma while (b) comes from an Android emulator.

consultation with the clinical supervisor revealed that the concern was well-founded.

3.3.5 Changes guided by validation

As a result of being the biggest cause of confusion, the pin-prick instructions underwent major changes between iterations 1 and 2. The instruction text was altered to directly state that the user must first prick the reference area and then prick the specified area on the leg. The instruction image was also reworked in order to illustrate both the reference area and address the issue of pricking the wrong leg. This was done by adding a torso where the clavicles are marked, and by showing the other leg behind the leg being tested. Finally, the answer choices were changed to the Danish and English version of *"same"* and *"less"*, and for the English version *"absent"* was changed to the more common *"none"*. A comparison is shown in Figure 3.5.

To better communicate that the user can get more information about tests, the help icon button was expanded to also include blue *"more information"* text. The button was moved to its own line, to not blend in with the instruction text.

Since there was a preference for option B (Figure 3.4b) in the A-B test, the gauge chart presentation was used as the main inspiration for the *Examination Completed* page. The concept was further built upon by displaying the number of points scored as well as the minimum and maximum number of points possible. A certainty level was no longer showed to the user and the gauge was changed from having *"green-yellow-red"* intervals to a gradient spanning the same colors.

3.4 Second design iteration

For the second iteration a functional app was developed based on the MVP and learnings from the validation. This also allowed vibration functionality to be added, making it possible to properly carry out the vibration test.

3.4.1 Validation goal

Since the self-assessment tool concept was met positively during the first validation (subsection 3.3.2), the focus of the second validation became to evaluate whether the problems from the first iteration had been mended and to discover new issues. This meant observing if the app provided a good user experience and whether the instructions were sufficient.

3.4.2 Validation participants and procedure

Another four participants were recruited, this time being solely non-patients. The ages were 19, 24, 72 and 75 with one being female and the rest being male. Usability testing was carried out in a similar manner to the first validation, after which the questions listed below were asked in Danish:

- How did it go?
- How easy were the instructions to understand?
- What do you think would improve the experience?
- Were there any difficulties?
- What do you think about the *Examination Completed* page?
- Is there any information you would have liked to see on the *Examination Completed* page?

Additionally, if any difficulties or concerns came up during the usability test, the observer would ask about those.

3.4.3 Validation learnings

The changes made to the pin-prick test turned out to be a step in the right direction, as all participants were using the reference area as intended. However, starting with the wrong leg was still a common occurrence and new issues were introduced as well. Some couldn't interpret what the torso image was, and due to the small size of the leg image some didn't notice the marked areas on the first two sections or the transition between them. There were also frequent requests for the option to answer *"more sensitive"* instead of having to use the *"same"* option.

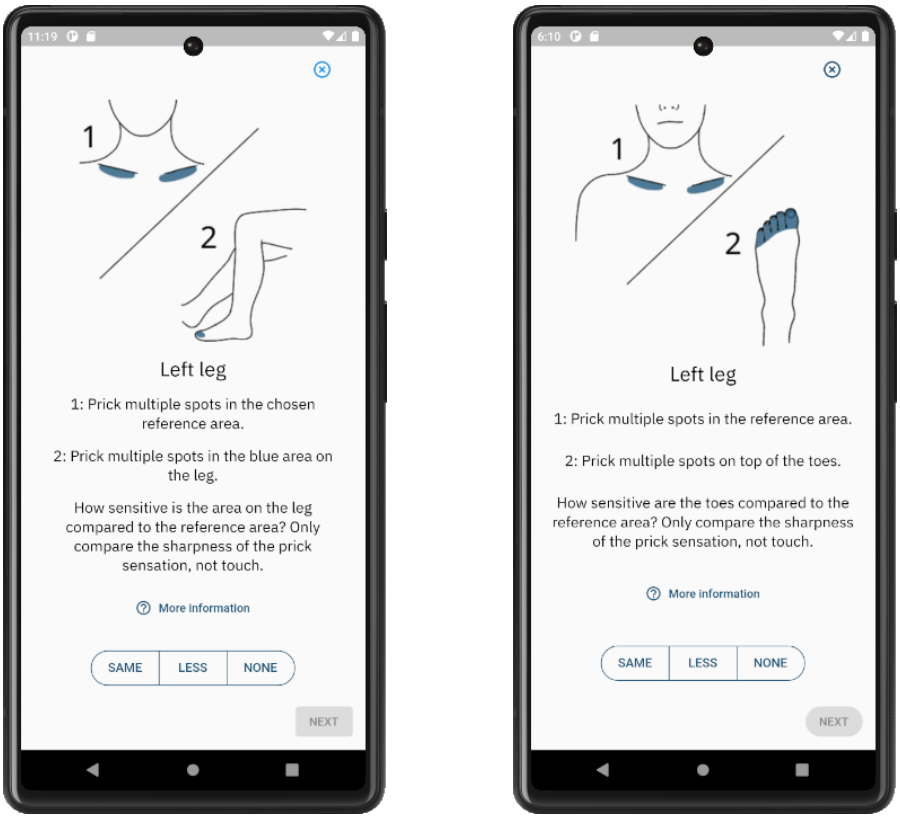
Two participants used the new *"more information"* text button, indicating that this feature had become more visible.

With the added vibration functionality the vibration test could be properly validated, and this revealed a significant limitation within the self-assessment tool. Since people had to hold the phone while pressing it against the leg, it proved challenging to discern vibration felt in the leg from vibration felt in the hands. One person attempted balancing the phone on the leg in order to not mix up the sensations, but the vibration still wasn't felt. Half of the participants only pressed the phone lightly against the leg, and there was one who kept socks and pants on for the test. These factors made it difficult for participants to confidently answer whether they could feel the vibration.

The change from color intervals to a gradient on the *Examination Complete* page did not prevent people from getting the impression of obtaining a diagnosis. They were shown the generic score of 14, and some interpreted it as *"being in the green"* due to the score lying between green and yellow on the gradient.

3.4.4 Changes guided by validation

The images for the pin-prick test were updated once again to add clarity. The torso image was extended to show more of the body, and a mouth and nose was added to the previously featureless head. For the first two sections, which are both located on the foot, the leg image was substituted for an image of just the foot. A comparison is shown in Figure 3.6.



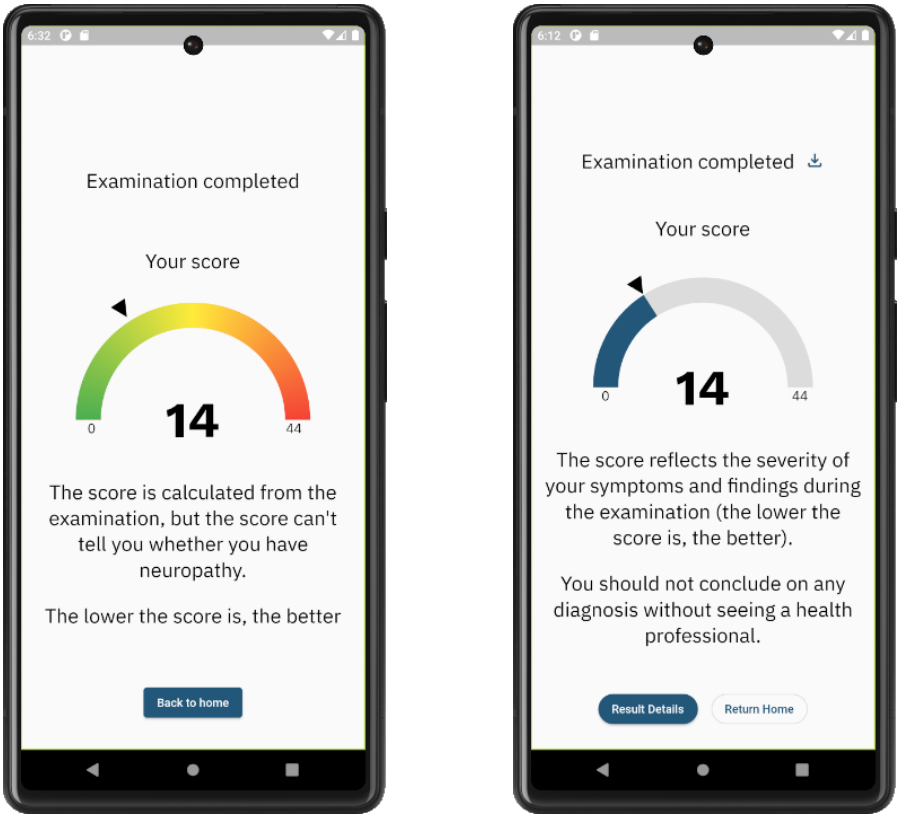
(a) Iteration 2 (b) Iteration 3

Figure 3.6: Second and third design iteration of the first pin-prick area

An additional requirement was introduced to the information page at the beginning of the examination. Since there had been inconveniences and confusion about performing tests on bare skin, a paragraph was added stating that the user must be able to access bare skin up to slightly past the knees. In this way people would know to expect it, and be able to decide before starting whether they can carry out the examination. This requirement is restated on the information pages right before the pin-prick and vibration tests as well.

The vibration instructions were extended by stating in bold text that the phone should be pressed firmly against the area being tested. This doesn't solve the problem of distinguishing vibration felt in the legs from that in the hands, but does increase the amount of vibration felt in the leg.

Finally, the gradient color on the *Examination Complete* page was substituted for a uniform blue color that fills up a portion of the gauge graph corresponding to the score, see Figure 3.7b. The text was changed as well to better communicate that the result shouldn't be interpreted as a diagnosis.



(a) Iteration 2 (b) Iteration 3

Figure 3.7: Second and third design iteration of the *Examination Completed* page.

4 Mobile Health Application for Self-Assessment of Neuropathy

This chapter presents the application developed for the project. It both describes the design and functionality, as well as implementation. If the reader is not familiar with programming or the technical aspect of the project is not of interest, section 4.2 can be skipped.

4.1 Design

The third design iteration was the last created during this project, implemented and used for the validation study (chapter 5). Many features of the app that aren't the grading tool flow itself were added at this point. The application has been titled *"Neuropathy Grading Tool"*. This section provides an overview of the final application design and presents the flow between the presented pages, explaining the behavior of the app.

4.1.1 Application flow

The navigation and interactions between pages in the application are presented in Figure 4.1. When the user opens the application, they see the splash screen. It will also be displayed if any changes to the saved results have been made - i.e. deleting the results, or finishing a new examination. All transitions between pages, except the splash screen, are slide transitions.

When the user taps the button to add examination, it begins the grading tool flow. It has 39 steps finishing with *examination completed page*. Navigation between those is identical, with the user tapping the *"next"* button after each step is completed. At any point of the examination it can be exited, with a warning dialog that the action will discard the result. When the examination is completed, the user can download the result, see detailed results or return to the *main page*. Download and detailed results can also be accessed from the *main page* itself. Downloading the examination result opens a file saver. It is not presented in the visualization of the flow (Figure 4.1), as it is not designed or implemented in this project, but native to the user's device.

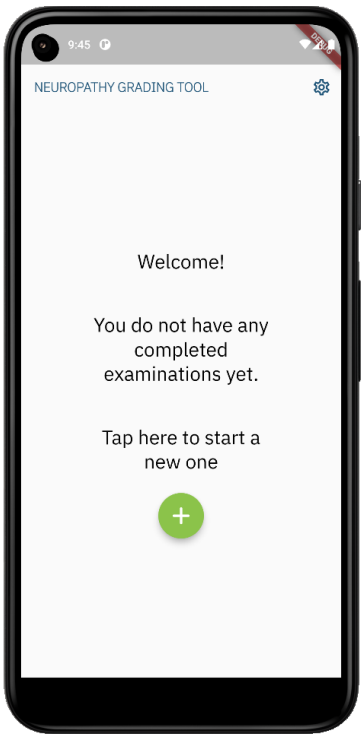
Both variants of the main page have access to the settings. When the user taps the desired setting, either a picker will be displayed on the bottom of the screen, or a dialog window will pop up. If the action will result in deleting existing results, a warning window will be shown warning the user and asking to confirm the operation.

4.1.2 Main Page

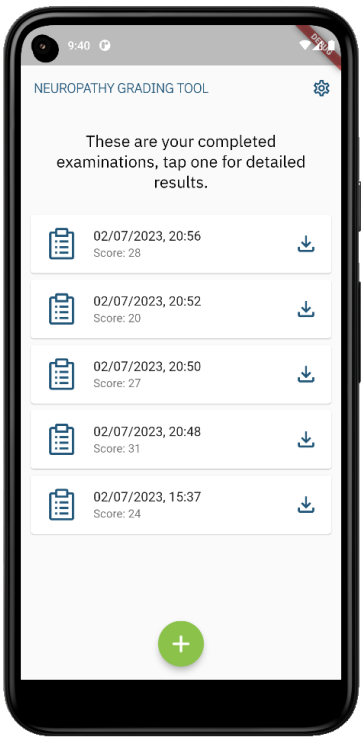
The *main page* is the entry point to the application, preceded with a simple splash screen. We designed two versions of the main screen to display - one welcoming a user that has not completed any examinations before, and one when returning to the app when there already are previous results. The design of those pages can be seen in Figure 4.2.

The text was designed to guide the user to the options they have on each screen. a floating button allows to add a new examination at any time, and if results are available, the user can see their details or download them. The detailed results displays the scored points in all parts of the examination, while downloading a result to csv is mostly meant for specialists or for data analysis in studies. A button for downloading results is present in many pages with the same icon for easy recognition.

Figure 4.1: Application flow between pages. Transitions happen when tapping on circled elements. Add and settings buttons on *Main Page with previous results* do not have arrows for better diagram readability. Their behavior is the same as on the other version of *Main Page*. The download icon has the same behavior across the application, and tapping it opens the device's file saver.



(a) Main page with no completed examinations



(b) Main page with previous examination results

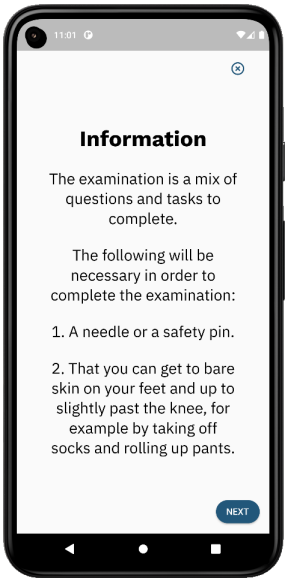
Figure 4.2: *Main pages* of the application. They are the entry point to the application, and all flows begin here. Their variant depends on whether there are previous examination results.

4.1.3 Grading Tool Questionnaire

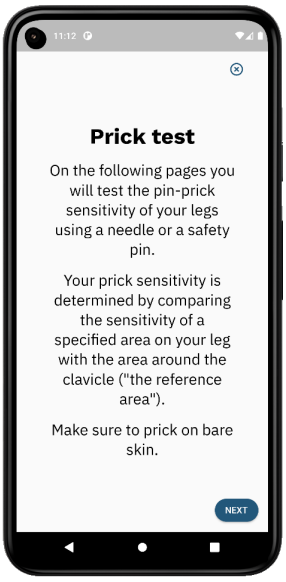
The grading tool has a total of 39 steps if the user is experiencing pain in their feet, or 34 if no pain is present. Each step is considered a separate page in the linear flow of the questionnaire. It follows the order described in subsection 3.2.7, finishing on the *Examination Completed* page. After clicking *"next"* the user is navigated to the next step. It is not possible to go back to previous questions. A lot of steps are repetitive tasks, where the instructions or questions are very similar, but target a different leg or leg area. This section describes the unique layouts of steps and what content changes in them if they are repeated.

All the non-question steps can be seen in Figure 4.3. The first step in the flow is the introduction to the examination (Figure 4.3a). It informs the user what will be necessary during the course of the examination, so they can prepare accordingly. *Prick Test* instructions (Figure 4.3b) precede the pin-prick section, and *Vibration Test* (Figure 4.3c) is displayed before the large fiber sensation segment. They both describe in general what the user can expect in the next steps, and how to properly do the tasks they receive. Note that both the introduction and those two instructional screens mention testing on bare skin. During the testing there was confusion around it, which is why it is repeated so often. The step introducing the motor examination (Figure 4.3d) does not include this, as it is not crucial. Additionally, since it's the last section of the examination, it's assumed the user has already taken off their shoes and socks for the previous steps. The final two steps of the questionnaire are the *Closing comments* (Figure 4.3e) and the *Examination Completed* page (Figure 4.3f). From the *Examination Completed* page the user can see their score,

download and access detailed results, and then navigate to the main page.

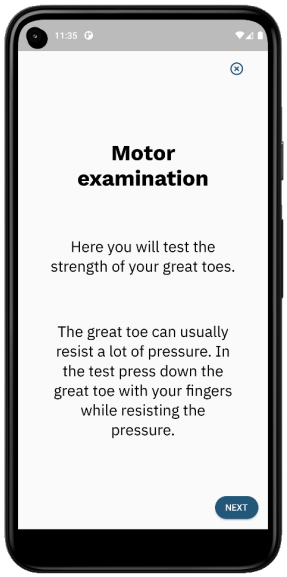


(a) Introductory step

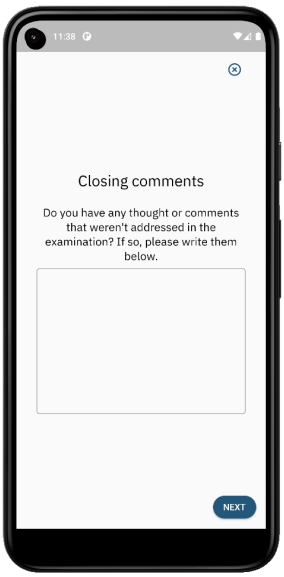


(b) Pin-prick section

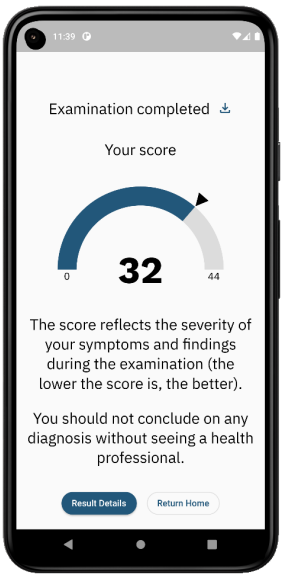
(c) Vibration section



(d) Motor Examination section



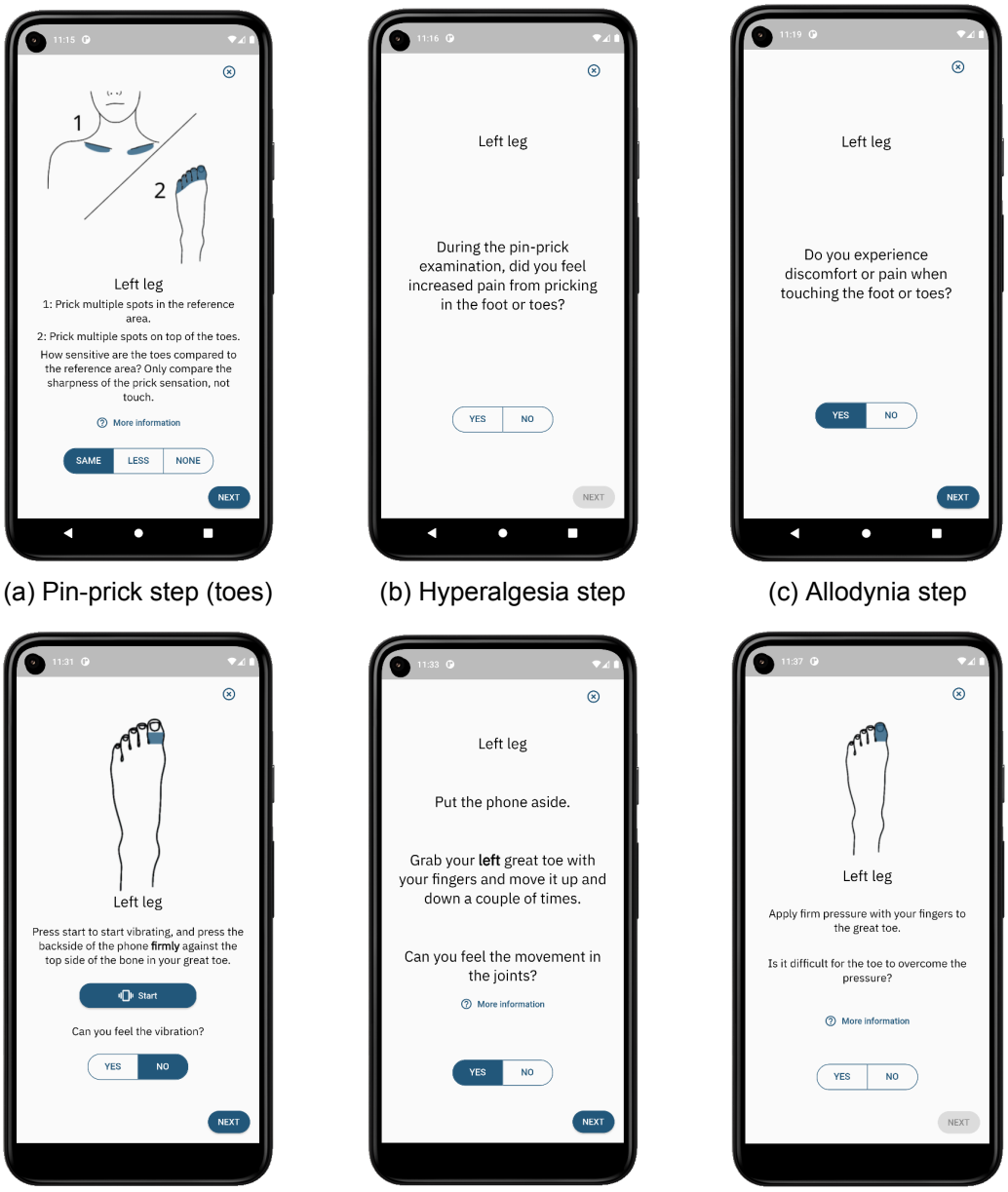
(e) Comments



(f) Examination Completed

Figure 4.3: Non-question steps in the assessment questionnaire.

The graded sections of the questionnaire that have multiple variants can be presented in six unique step layouts (see Figure 4.4). Each of the steps states the currently tested leg. All the steps that require an answer by default do not have an answer selected. The questions are single choice and mandatory, so if no answer is selected, the *next* button is grey and inactive. The pin-prick step (Figure 4.4a) provides two-step instructions for each section of the leg. In the variants of this layout, the image of the tested leg area changes - see all the areas in Figure 3.2b, and the images for left and right leg are vertically mirrored. When pressing *"More Information"* on this screen the user can see a detailed explanation what the *"same"*, *"less"* and *"none"* answers exactly mean. Hyperalgesia (Figure 4.4b) and allodynia (Figure 4.4c) steps come in two variants, changing only the tested leg.



(a) Pin-prick step (toes) (b) Hyperalgesia step (c) Allodynia step

(d) Vibration step (great toe) (e) Great joint position step (f) Motor examination step

Figure 4.4: The layout of repeated steps of the graded sections of the questionnaire with example answers or no answers selected. The steps vary in text and/or image content in the repetitions.

Vibration steps (Figure 4.4d) display different sections of the leg (see Figure 3.3), mirroring the images for left and right leg as well. Instruction for each of the points is different as well, directing precisely where to press the phone to - the user is asked to press firmly against the *"top side of the bone in your great toe"*, *"outside part of your ankle"* and *"bone just below your kneecap"*. When the user presses the start button, the phone vibrates, and the button changes its color to red, displaying *"stop"* instead, so the user can cancel the vibration. The great toe joint position question (Figure 4.4e) in the vibration test only changes the tested leg text. It directs the user to put the phone aside as it would be difficult to perform the task while holding the phone, and the *"More Information"* button is available on this screen if the user is unsure about what feeling movement in the joints

means.

For the motor examination step (Figure 4.4f), only the tested leg text changes, and the image is mirrored to represent the current leg, along with the text change. With *"More Information"* on this screen the user is informed that they can compare how well their fingers overcome applied pressure, and to answer *"yes"* if their great toe is weaker than that.

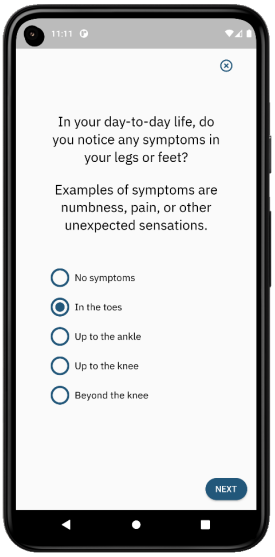


Figure 4.5: General Symptoms step

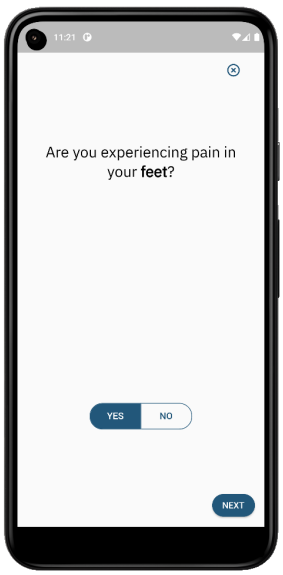
Finally, there are seven unique question steps in the questionnaire. Of these steps, only general symptoms step (Figure 4.5) is counted towards the assessment score. It's a single choice question asked immediately after the introduction to the examination.

The rest of the unique steps are connected to the pain section and are shown in Figure 4.6. As mentioned in subsection 3.2.7, the pain part of the questionnaire is omitted if there is no pain. If the user answers they experience no pain in the *Pain Presence* step (Figure 4.6a), clicking next will move them directly to the vibration section. If pain is present, the user will go through steps from Figure 4.6b to Figure 4.6f.

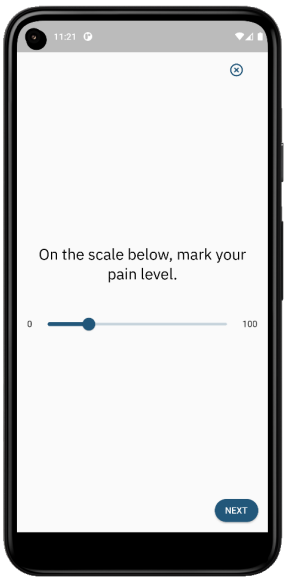
Pain characteristics (Figure 4.6c) and symptoms (Figure 4.6d) steps are the only multiple choice questions in the questionnaire. The user is asked to choose all that apply, and one of the options is *"None of the above"*. That option is not on the DN4 questionnaire, but on the paper form they can just refrain from selecting any of the options. Instead of making this step optional and allowing clicking *"next"* immediately, they need to confirm their choice by checking *"None of the above"*. This prevents accidentally skipping the question, since the user cannot go back to previous steps.

4.1.4 Detailed Results

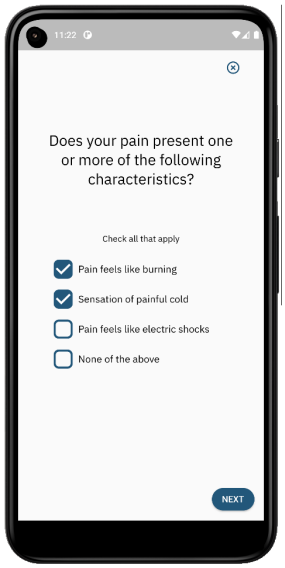
The *Detailed Results* page can be accessed from the *Examination Completed* screen or by tapping an examination listed on the main page. It displays the results sectioned in collapsible cards, which can be seen in Figure 4.7. If the pain section or the comments was skipped in the examination, they are not displayed. The initial state of the page is presented in Figure 4.7a. In the top of the page date of the examination and its final score is shown, while all the section cards are collapsed. The user can open and close multiple sections at a time, and the page becomes scrollable if the content is too large for the screen.



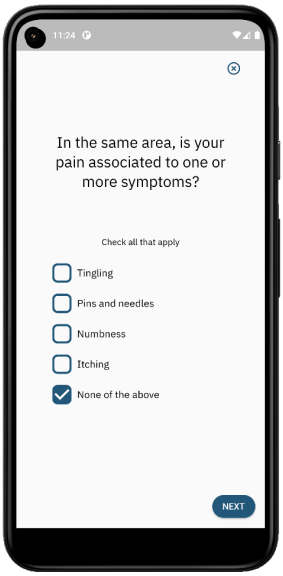
(a) Pain presence



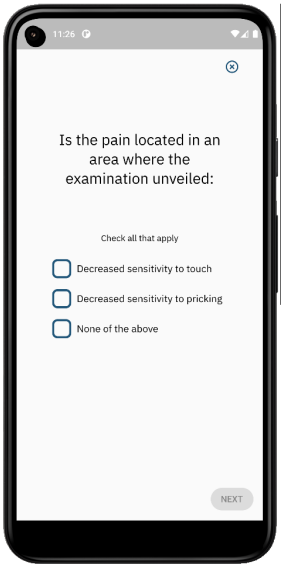
(b) Pain level



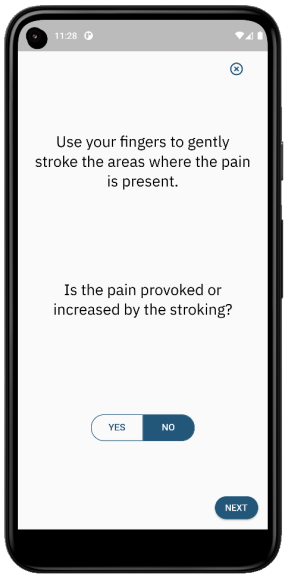
(c) DN4 Pain characteristics



(d) DN4 Pain symptoms



(e) Touch and Prick sensitivity



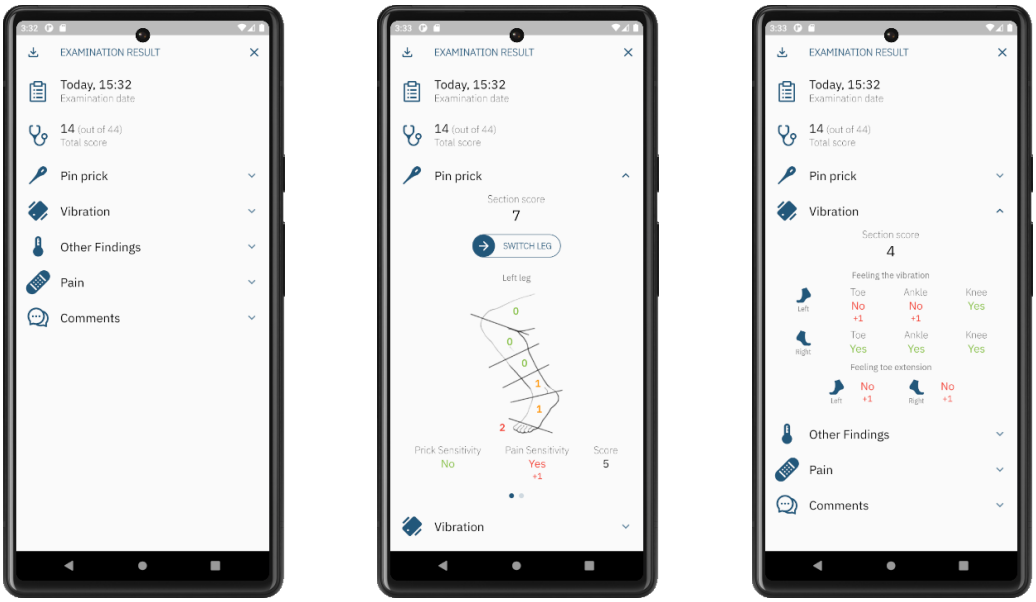
(f) DN4 Response to stroking

Figure 4.6: Pain Presence step and the Pain part of the questionnaire steps.

The pin prick section (Figure 4.7b) displays the total score for the section, and shows the number of points scored on each area of the leg, as well as and allodynia and hyperalgesia scores, captioned as touch and prick sensitivity. Tapping and swiping the slider or swiping the leg image changes which leg is shown. The vibration card (Figure 4.7c) encompasses the vibration and the toe extension tests which are both part of large fiber sensation. Figure 4.7d shows other findings, which consists of the motor examination and general symptoms.

The pain section (Figure 4.7e) displays the pain slider and the DN4 questionnaire responses. Any symptoms or pain characteristics found are listed along with an icon that symbolises it - for example, the *"pain feels like burning"* presented in the example has a fire icon next to it. The last card - Comments (Figure 4.7f) - displays the text the user

entered at the end of examination.



(a) All sections collapsed (b) Pin-prick section (c) Vibration section

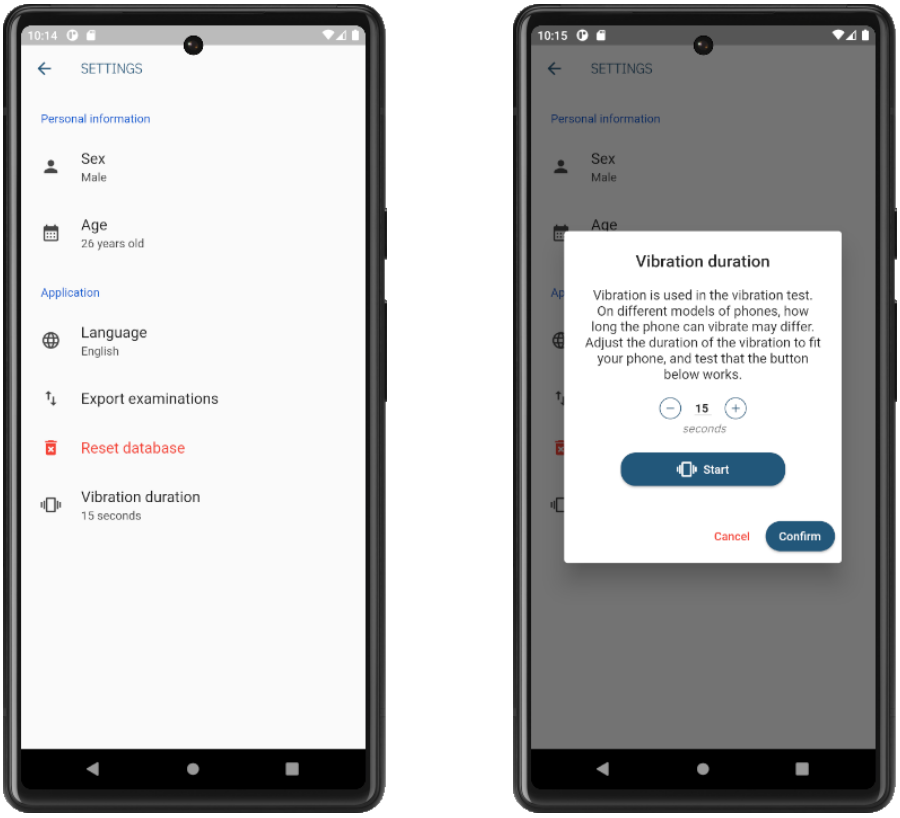
(d) Other findings section (e) Pain section (f) Comment section

Figure 4.7: *Detailed Results* page showing example examination results.

4.1.5 Settings

The settings page can be accessed from both variants of the *Main Page*, and it can be seen in Figure 4.8a. It is a list of tiles in two sections - Personal information and Application. In the first case the user can provide additional information: sex and date of birth that will be included in result exports. The date of birth is displayed as age on the tile. Examinations are associated with a single person, so after entering the personal information the first time it cannot be changed without deleting all previous examinations. For any action that would delete previous results, a warning is shown where the user must accept or cancel the operation.

In the Application section, there are 4 actions that can be performed. The user can choose between supported languages (currently Danish and English). The default language the first time the application is opened after installation is English. All examination results can be exported to a single file from this page as well, in the same format as single results are. Resetting the database will remove all existing results, so the previously mentioned confirmation dialog is displayed. Lastly, the user can define how long the phone will vibrate during the examination. This might be necessary as devices will not initiate vibration if the requested duration is longer than a set limit, which varies per device. A custom dialog where the user can test the new duration is displayed (see Figure 4.8b).



(a) Settings page overview (b) Vibration duration setting

Figure 4.8: Settings page with example entries

4.2 Implementation

Figure 4.9 shows the overall architecture of the mobile application. It consists of a set of User Interface (UI) screens (marked red) described in section 4.1, implementing the flow shown in Figure 4.1, and the non-UI components (marked green). The application is using the Research Package (RP) [61] framework to build the examination questionnaire, and relies on many flutter plugins. The main plugins have been listed in the Figure 4.9 (marked blue), but some have been omitted. The full list can be found in the `pubspec.yaml` file in the applications source code - see Appendix B.1.

Three aspects of the user's device are used, storage, vibration and file saver (marked yellow). A Not Only SQL (NoSQL) database is implemented using the Sembast package [62]. The implemented application does not connect to the internet, but uses local storage for saving data. A repository pattern is used, where our repositories for settings and results define simple commands for interacting with data, which the repositories translate to fit

the database language. The database is initialized on application start, registered as a singleton, and immediately used to fetch any existing settings and results. Preferred app language is also stored in the device, with the Shared Preferences plugin [63]. The Vibration package is used for accessing vibration functionality, which will be elaborated on later in this section. The device's File Saver is used by the to_csv package [64] which allows for downloading the results by the user.

This section goes into detail on implementing the grading tool using RP, as well as other app functionalities.

4.2.1 Research Package (RP)

Copenhagen Research Platform (CARP) [65] is a set of frameworks and components developed by Copenhagen Center for Health Technology (CACHET), mainly targeted at mHealth app development in Flutter. This project uses CARP RP which is a framework for creating surveys and collecting informed consent. The framework is built on the Business Logic Component (BLoC) architecture, meaning that a central component is handling states and logic that is common for the whole survey.

An RP survey is called a `Task` and consists of a series of `Steps`, most notably the `InstructionStep` that displays information and the `QuestionStep` that contains a question and a way for the user to provide an answer.

The survey (examination questionnaire) part of our self-assessment tool is implemented using the RP framework, but uses many modified components in order to add the needed functionality.

4.2.2 Modified components

The `InstructionStep` that comes with RP allows displaying a paragraph of text and, optionally, an image. Due the the importance of performing tests as intended, a lot of text is conveyed in the survey, and it is much easier to read if it's separated into paragraphs. Therefore a modified version of the `InstructionStep` was created that omits the image parameter which wasn't needed, but accepts a list of paragraphs that will be evenly spaced out on the screen. This version was used for all instructions in the survey.

The `QuestionStep` from RP only displays a paragraph of text, but as demonstrated in chapter 3, it is the images that are essential for conveying instructions in tests such as pin-prick. For that reason a new type of `QuestionStep` was created that accepts an image as well as multi-paragraph text. It also has the option of putting a *"More information"* button, that opens a `BottomSheet` with additional information when pressed.

A custom component that accounts for both single and multiple choice questions was created. It deselects existing choices that are incompatible with the latest choice (see algorithm in Figure 4.10). This algorithm is also applied to the toggle-button questions, like the pin-prick questions, however they skip the `Is SingleChoice` step and follows the left path since toggle-button questions can only be single choice. Thus, the *Pain level* question is the only question in the survey that doesn't rely on the algorithm.

The *Pain level* question appears in the pain section. A custom `PainSliderQuestionStep` that presents a slider for pain ranging from 0 to 100 was implemented. The default slider step provided by RP displays the selected score, which is not part of the visual analogue scale.

A feature that isn't part of RP is the ability to utilize the phone's vibration capabilities. Another type of `QuestionStep` was made that, in addition to taking an image and multi-paragraph text, also displays our custom vibration button.

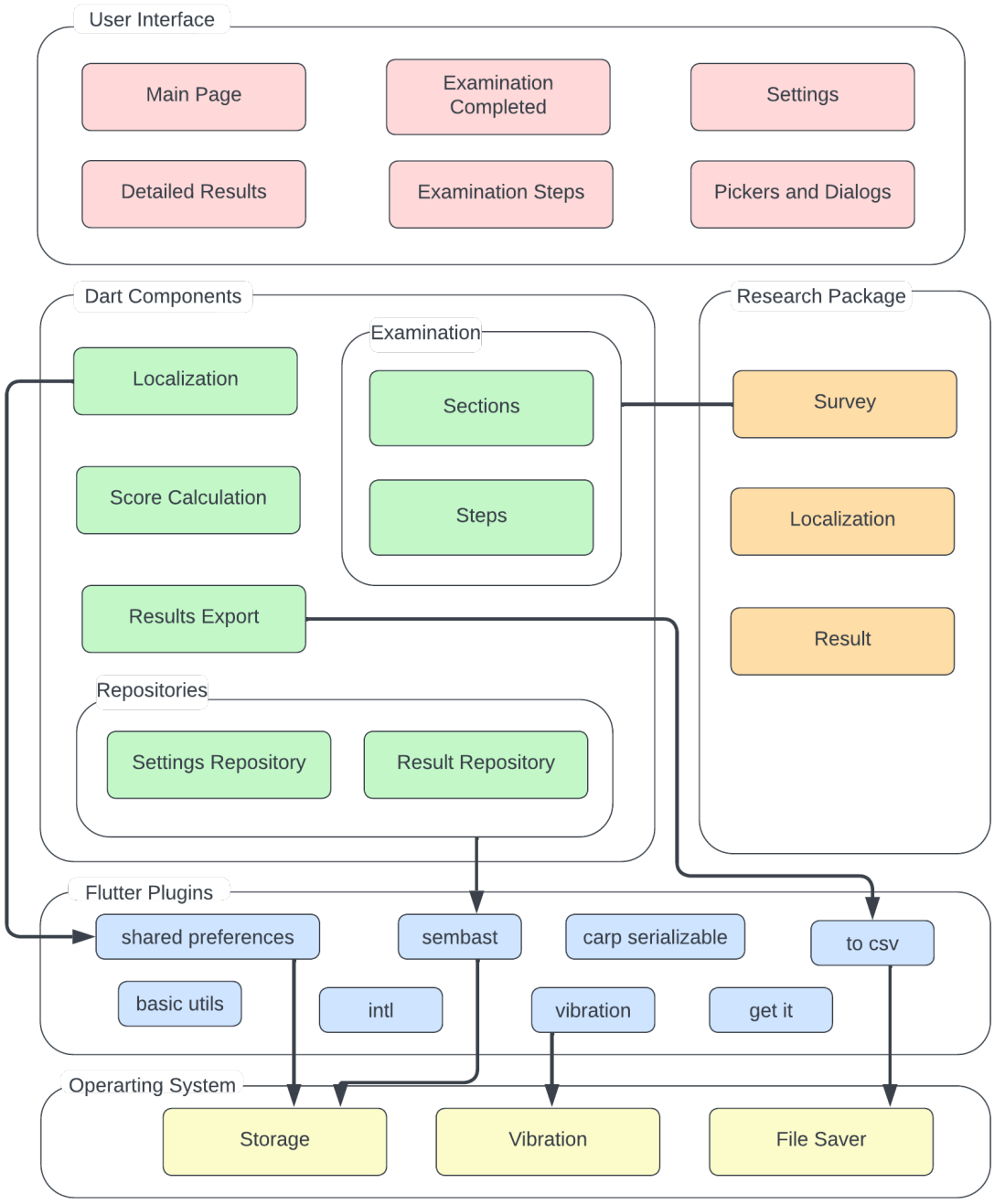


Figure 4.9: Architecture diagram of the ***Neuropathy Grading Tool*** app (red and green components), and its use of the Research Package (RP) (orange components), flutter plugins (blue components) and the device's system (yellow components). Only the most notable flutter plugins are listed.

4.2.3 Vibration

Phone vibration is used in large fiber sensation test - see Figure 4.4d. A vibration button that uses the `vibration` package [66] makes it possible to start and stop vibration on the device. Since the package doesn't allow for indefinite vibration, a default duration of 15 seconds is used for the button, after which the vibration stops. While the vibrating takes place, the button is turned into a red "stop" button, and when the vibration completes or is stopped by the user, the button returns to its original state and can be pressed again.

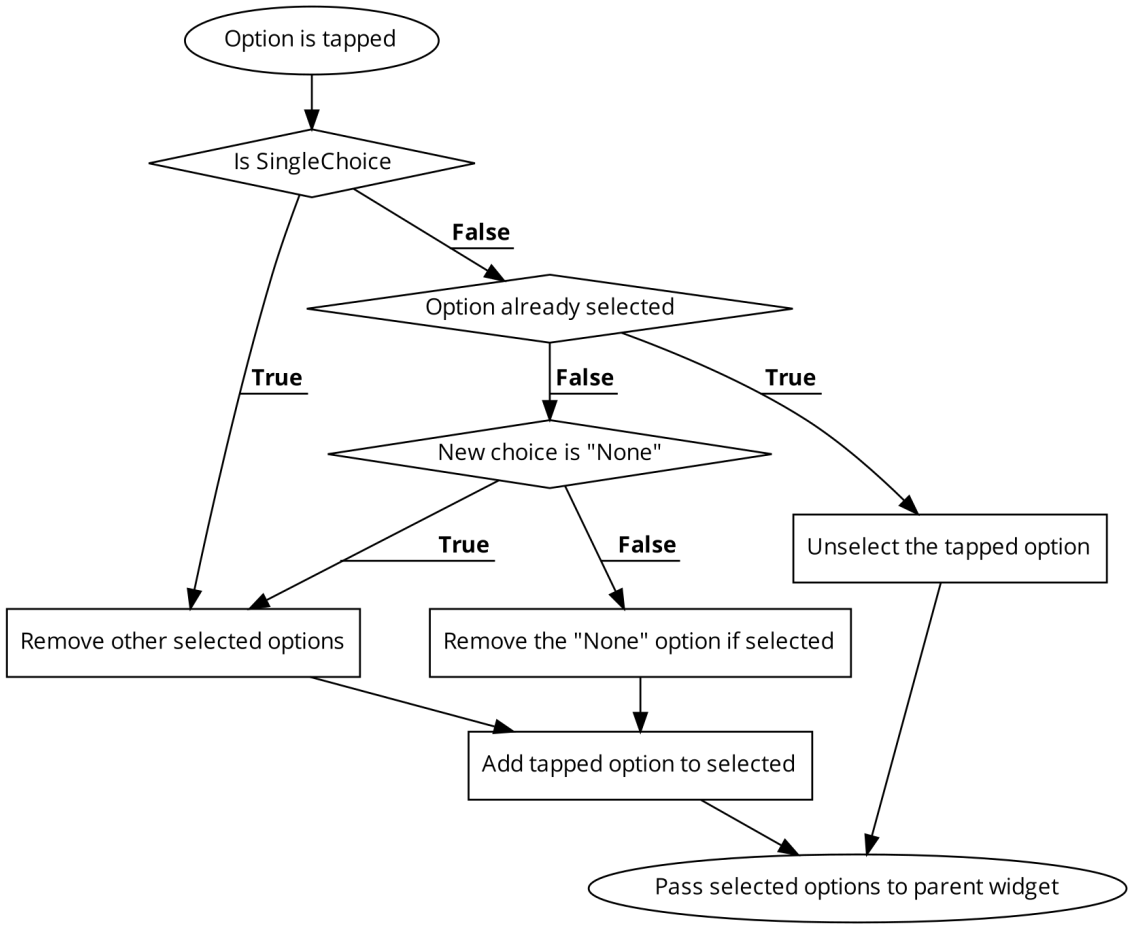


Figure 4.10: An algorithm flowchart for selecting an option in single and multiple choice questions.

The package doesn't provide any signals when the vibrating stops, meaning that the vibration button must set a timer to keep track of when it should update its state. The timer is implemented as a cancelable operation, that finishes either when the time runs out or when the user presses ”stop”. In the first case the vibrating runs out on its own, and in the second, the vibrating is stopped immediately. In both cases, the vibration button is returned to the initial state. It is necessary that the timer is cancelled and not just left running, otherwise, rapidly tapping the vibration button would start multiple timers, all returning the button to the initial state when they finish, regardless whether the phone stopped vibration or not.

Progressing to the next step also stops the vibration, but not the timer which will still attempt to update the state of the vibration button. A `Step` is disposed of when the user taps *”next”*, and attempting to update the state of a disposed `Step` results in an exception. For this reason the widget checks whether it is mounted whenever it tries to update its state.

4.2.4 Combining sections in the survey

When the app creates a `Task`, it provides a list containing all steps of the survey, except for the *Examination Completed* page which is not part of the `Task` object. This means that the `Task` receives 38 steps, and in order to keep them manageable, subsets of the full list are defined in separate files, one for each section of the survey.

Those files define each step of their respective section: Whether it's an `InstructionStep` or `QuestionStep`, the type of question as well as the image and text used. Additionally, an integer value is assigned to each answer option, which is used for score calculation. These values are in accordance with the point allocation from section 3.2. While some steps are unique and need to be defined individually, others are repeated many times after each other with small changes. In those cases, generator methods are used to reduce redundancy.

Listing 4.1 shows how the smaller lists for each section are combined into a single list and passed to the `Task`. It also shows the `NavigationRule` that is applied to the Skip pain step (Figure 4.6a). The value 0 is mapped to the first step of the vibration section, meaning that if the user gives the 0-answer, *"No"*, when asked if they experience pain in feet, the `Task` navigates to the vibration section skipping the rest of the pain section.

```
RPStepJumpRule noPain =
    RPStepJumpRule(answerMap: {0: vibrationInstructionStep.identifier});

RPNavigableOrderedTask examinationTask = RPNavigableOrderedTask(
    closeAfterFinished: false,
    identifier: 'ExaminationTaskID',
    steps: [
      introductionStep,
      symptomsStep,
      ...prickStepList,
      skipPainStep,
      ...painStepList,
      vibrationInstructionStep,
      ...vibrationStepList,
      ...motorStepList,
      freeTextStep,
    ])
  ..setNavigationRuleForTriggerStepIdentifier(noPain, skipPainStep.identifier);
```

Listing 4.1: `examination_survey.dart`. Collecting all parts and creating the `Task`.

4.2.5 Examination result score and persistence

The score of the examination is displayed in multiple places in the application and is used in result export as well. To calculate the score of an examination, an `RPTaskResult` is narrowed down to a list of step results that count towards the score. Then, the list is folded, retrieving the answer values from each of the steps and adding them together.

Examination results are stored using the result repository. After the examination is completed, the result object is serialized using the carp_serializable package [67] and inserted into the database. Each inserted result has a unique, automatically generated id. The id of the last result inserted is stored after it is generated, which enables retrieving the latest result. When a result is retrieved, it needs to be deserialized back to an `RPTaskResult` object. Deserialization is performed using RP mixed with custom elements to ensure proper type casting.

The score is not stored with the result, as it is not part of the result object. It is calculated separately in components that display it. This is also true for the *Examination Completed* page shown after questionnaire completion (see Figure 4.3f). The result is first saved to the storage, then the latest result is retrieved and made available on the *examination*

Completed page, making it possible to see the score, download the result or see detailed results. A loading screen is displayed for at least one second while the results are retrieved. If it was only shown for the duration of the operation, it would vanish so quickly that it would look like a glitch.

4.2.6 Exporting results

Individual or all results can be exported from the application to the user's mobile device as a csv file. In that format, each result is a row. The headers are the following:

1. Timestamp of the examination
2. User's sex (empty if not provided in settings)
3. User's date of birth (empty if not provided in settings)
4. Result - score of the examination
5. Graded tasks identifiers (27 headers)
6. Comments step identifier
7. Pain section identifiers (5 headers)

In the RP framework, each step has an identifier. All of the examination steps have been implemented with unique identifiers, so each of the questions will produce one header for the csv file. Each exported result is mapped to a list of strings that will become a row during export. For all headers in the above mentioned order, a cell value is retrieved by matching the header to a step result and extracting the answer value. Cell values for headers that are not part of the result object, are retrieved from the user's data from settings, or calculated. Then, the to_csv [64] plugin is used to create a csv file and open the device's file saver so the user can decide where to save the file.

4.2.7 Localization

The application is available in two languages - Danish and English. The language is not dependent on the user's device locale, and can be changed in the settings page. That choice is then persisted using the Shared_preferences plugin [63].

Displaying the text in the correct language is carried out with two localization delegates - one from the RP and a custom one. With built-in support for Danish and English, RP delegate translates parts of the survey - the *"next"* button, and the warning dialog when exiting the examination.

For the rest of the text, two files were created in the assets folder, one for each supported language. Each contains the same set of keys and translated strings in a nested structure, meaning that a key can indicate a section, which then contains key-value pairs for the translation. When the application is initialized or when the language is changed, a map of localized values is loaded. The nested translation file of the selected locale is flattened. An example of this would be:

```
"common": { "yes": "Yes"} => "common.yes": "Yes"
```

The flattened key contains both the section and the individual key. All components of the application that display text, render it by referring to the `translate` method of the localization class, providing the flattened key as parameter. An example of that would be:

```
Languages.of(context)!.translate('common.yes')
```

The `Languages` localization class is initialized during the application start in the widget on top of the widget tree, so it exists in the context of each child widget.

5 Validation Study

After arriving at our final design, we carried out a study on the scoring system in order to validate that the self-assessment tool could be useful in detecting DSP and assessing its severity.

We would determine this by comparing the scores from examinations using both our questionnaire and TNSc. Both would be carried out on each participant, thus making it possible to correlate one score with the other. Getting a high correlation would indicate that our self-assessment tool's ability to evaluate the severity of DSP is comparable to that of TNSc.

The study took place at the outpatient clinic at the Neurological Department of Zealand University Hospital with aid from the clinical supervisor. Even though the goal of the self-assessment tool is to be used at people's home without supervision, the clinical setting was chosen instead for practical reasons: We would have to be able to provide phones with the app on it for all participants, the third iteration had not been user validated which means that there still could be major problems, and the goal of this study was to validate the scoring system under the assumption that the tests are performed as intended.

5.1 Methods

The clinical supervisor provided an exam room for carrying out the study and time schedules for consultations at the outpatient clinic, as well as a reflex hammer and a tuning fork. The study had two phases: In phase I, the participant used the app to get a self-assessment, and in phase II TNSc was conducted on them. This ordering was chosen to reduce the amount of exposure participants had to the types of tests before the self-assessment, and so they would base the responses on their own experiences rather than our findings. Both phases were completed for each participant during the same appointment.

5.1.1 Preparations for TNSc

TNSc, as explained in subsection 2.2.2, is one of the diagnostic measures aimed at DSP. It is usually performed by a clinician, but we were trained by the clinical supervisor to be able to carry it out ourselves.

To ensure that all results from TNSc were collected in one place, a survey was created in Microsoft Teams containing all the examination questions. We filled out the survey alongside doing the tests, and upon completion, Teams automatically added the survey results to an Excel sheet. This also meant that the ordering of completed surveys was kept intact, ensuring that TNSc results could easily be paired with the app results. As a precaution, a question field was put in the Teams survey for entering the corresponding app score.

5.1.2 Recruiting participants

The time schedules provided by the clinical supervisor contained information on the kind of consultation each patient was coming in for, but not whether they were neuropathy patients. Alongside neuropathy appointments, patients came for more general neurological consultations such as Electromyography, making it difficult to know which patients were DSP patients. Recruited participants were not asked about diagnoses or other sensitive

information. Some were returning or control patients while others were new, meaning that their familiarity with the procedures in measuring neuropathy varied greatly.

Considering the goal of the study, it is a benefit if the participants are evenly distributed across the spectrum of neuropathy severity, and not clustered in either end. This allows the whole range of possible scores to be tested and help avoid artificial inflation of the correlation.

Since the majority of the general population will have none or very few symptoms and signs related to DSP, a cluster is more likely to appear at the low scores. To minimize the number of participants in this range, patients who had the most relevant consultation types were prioritized.

A total of 17 patients were recruited for the study, 5 being male and 12 female. Ages ranged from 40 to 75. The study was an organic process that shifted back and forth between recruiting and conducting the study on patients. This means that there wasn't and overarching phase I and II of the study, but that each participant was taken through both phases individually.

5.1.3 Phase I of study - Self-assessment tool

In the first phase, the participants performed the survey using the application. They were provided with a phone running the implemented app in Danish and a safety pin, and then asked to carry out the examination. Participants performing some of the survey steps can be seen in Figure 5.1.

(a) Pin-prick (b) Vibration (c) Motor

Figure 5.1: Study participants carrying out the self-assessment in phase I.

Since the goal was to validate the scoring system under the assumption that tests were conducted as intended, we didn't enforce strict restrictions on answering participants' questions.

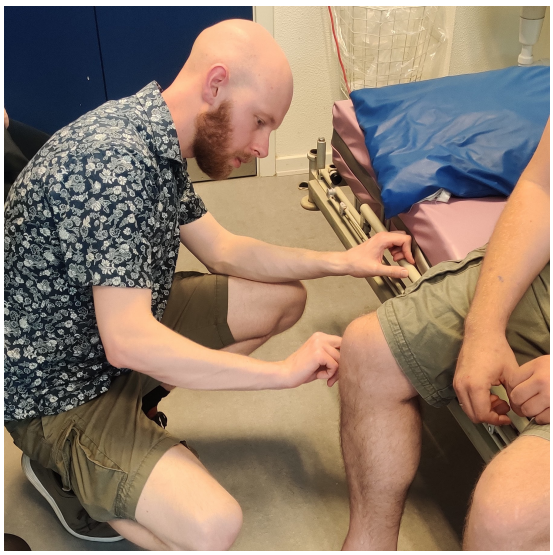
5.1.4 Phase II of study - Total Neuropathy Score clinical (TNSc)

In the second phase, TNSc was conducted on the participants, and Figure 5.2 shows some of the components of the measure being performed. The procedure was mostly as described in subsection 2.2.2, except that testing was mostly focused around the legs due to the previously mentioned fact that neuropathy develops there much earlier than in the arms.

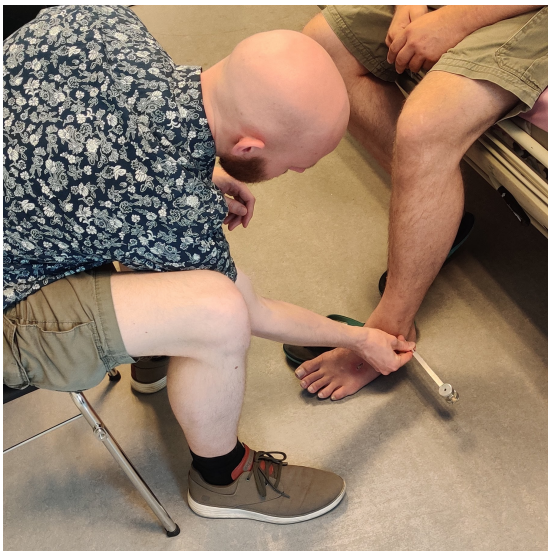
Participants were first asked about the four sensory symptoms. For each of the symptoms they were asked if they experience it in the feet and legs, and it was determined how far up. For motor symptoms, they were asked about the feet and legs, as well as the hands and arms, to uncover if they had any difficulty using them.

Pin-prick sensation was tested with the safety pin on the feet and legs, using the clavicle as a reference area, and it was determined how far up the leg the sensation was reduced. A 128 Hz tuning fork was used for testing vibration sensation in the leg. Four bones were used, the bone on top of the great toe, the outer ankle bone, the bone right below the knee and the hip bone. For each of those, the participant was asked how well they could feel the vibration.

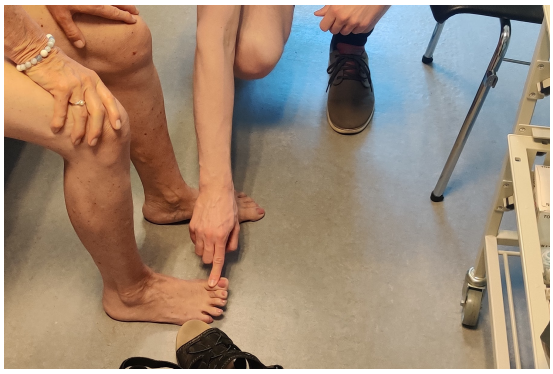
For motor strength, the ankles, fingers and wrists were all tested, but reflexes were only tested in the ankles with the Achilles tendon and in the knee with the patellar tendon. Finally patients were asked about the autonomic symptoms.



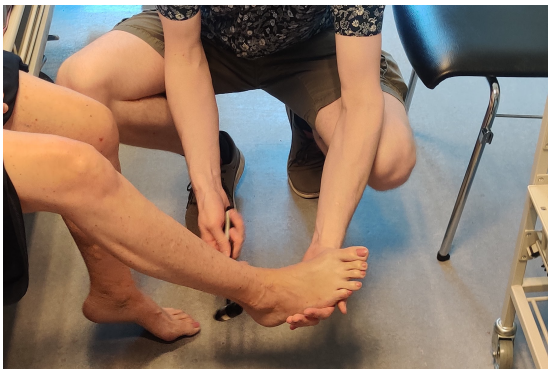
(a) Pin-prick



(b) Vibration



(c) Motor



(d) Reflex

Figure 5.2: TNSc carried out on study participants in phase II.

5.2 Results

Having conducted the study, the collected data was processed and prepared for analysis. Out of all the participants, the vast majority showed symptoms or signs relevant to DSP,

but they were not necessarily caused by it. An example was a person with pinched nerve in their hand and experiencing leg cramps, which was picked up by TNSc, but didn't correspond to any of the factors measured by our application. The total scores from our self-assessment tool were automatically computed and were in most cases ready to be used as-is.

There was one case where a study participant had given responses in the app that resulted in a score in the lower half of the spectrum, while the results from TNSc indicated a much more severe condition. Discussion with the patient revealed that there were major inaccuracies in their responses to the pin-prick and vibration sections, possibly due being inattentive to the answer options. After consulting with the clinical supervisor, it was agreed that since we assume correct usage of the tool and we knew what the score would have been, we would correct the app score.

Unlike the examination with the self-assessment tool, the TNSc examination was fully supervised, and no errors in the responses were discovered. The collected results were sub-scores that were computed into the total score after the whole study was complete.

5.2.1 Quantitative results

After arriving at the total scores from the study participants (see Appendix B.2), the Pearson correlation of the two measures was computed to 0.861, which is seen in Figure 5.3.

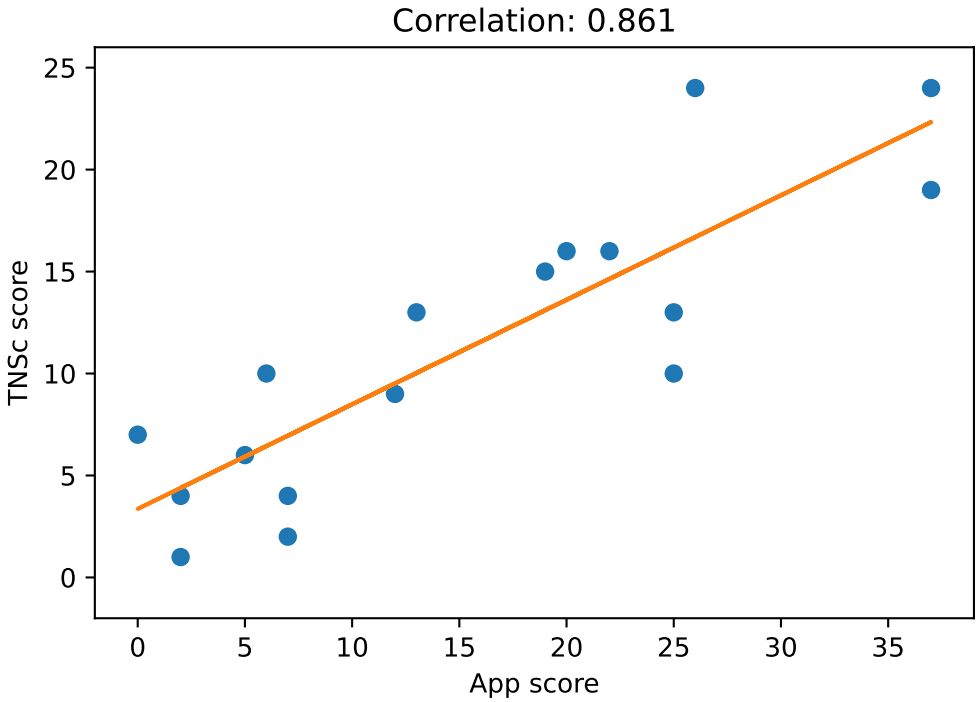
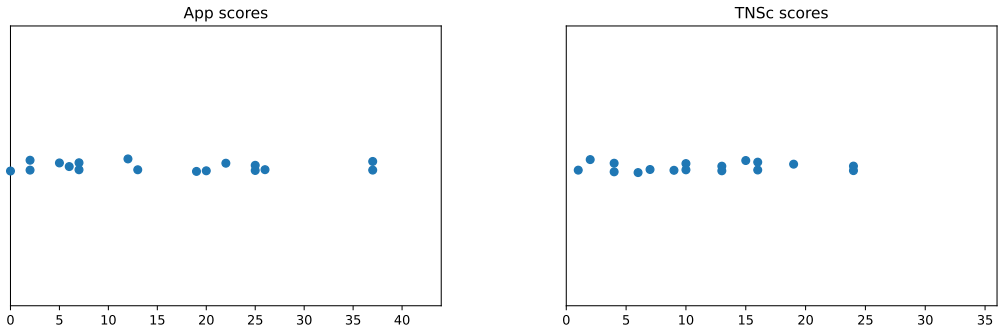


Figure 5.3: A scatter plot displaying scores from the application and TNSc, as well as a linear fit of TNSc score as a function of the app score.

It can also be inferred from the figure that the collected results for both measures are well-distributed, but Figure 5.4 shows it even more clearly. There are no major clusters in either measure, and only some gaps.



(a) Scores measured with the app.

(b) Scores measured with TNSc.

Figure 5.4: The scores collected during the study for both measures. The y-values are randomized in order to reveal duplicate scores. The x-axes span the ranges of possible scores.

5.2.2 Qualitative results

In phase I we found that the participants did not have much trouble conducting the self-assessment. Even those with the most severe neuropathic symptoms and signs were able to complete it on their own. Participants occasionally asked questions about the tests, but the main thing they asked for was getting confirmation whether they were doing it correctly, and there was no need for correction.

Participants were confident in their responses for most steps, though vibration was the one section that was troublesome. Similar to what was found during usability testing in chapter 3, participants had difficulty telling if the vibration is felt in the feet or hands. This meant that they often couldn't answer the vibration questions confidently.

During the study, it became very evident how important it is to make the survey questions easy to grasp. As previously stated, one participant happened to give responses to pin-prick and vibration opposite to what they felt, resulting in a large error in their score. This was a very clear case of an error happening in the self-assessment, and it draws to attention that there may be errors in other data points as well, though smaller and unnoticed.

Though TNSc was performed fully supervised, those measurements can contain some uncertainty as well. As mentioned in subsection 2.1.4, correct testing of reflexes can be difficult, and when we were doing it, it was often hard to tell if the reflex was absent or if the test was performed with incorrect technique.

6 Discussion

In this chapter it will first be discussed whether the research sub-questions, and by extension the overarching research question, have been answered over the course of the project. Then, more specific topics of the project, such as design, study and potential will be discussed.

6.1 Research question

State-of-art neuropathy diagnostic tools were researched in order to determine the feasibility of a PRO assessment tool that people can use without guidance from medical professionals. It was found that even though state-of-art measures are developed for use by clinicians, there is great potential for many of the individual measurements to be carried out by a person on their own. NCS is an example of a measure that is not possible, and reflex evaluation is unlikely due to the difficulty of reliable execution. On the other hand, symptom reporting, pin-prick testing and motor strength, which are among the common measurements of early neuropathy, were shown to be feasible without assistance if proper instructions are provided. We were not able to determine with high certainty if vibration, which is also common in early neuropathy assessment, is feasible as well.

A set of tasks and questions were developed in collaboration with medical specialists in the field and potential users of the self-assessment tool. Input from both sides was necessary in order to ensure that the steps are easy to perform, that users can be confident in their responses and that the collected results are useful for clinicians. This succeeded for most steps in the iteration used for the study, but some steps would benefit for further iterations, as will be discussed further in this chapter.

The scoring system of the self-assessment tool is based on the UENS, meaning that both the set of steps in our survey and the amount of points they add to the final score has much in common with that measure. While pin-prick and motor tests are identical in their scoring, testing reflexes was completely removed. Hyperesthesia was expanded to contain individual steps for allodynia and hyperalgesia. Large fiber sensation was drastically changed in order to simplify the test, only distinguishing between present and absent sense of vibration and joint position. For vibration, the ankle and knee were included as well. Despite those changes, the amount of points scored for this section remained the same as in UENS. General symptoms was added to determine how far up the leg users feel neuropathic symptoms. A pain section implementing the DN4 questionnaire and a visual analogue scale for pain were added as well. These scores are separate from the overall neuropathy assessment score given at the end of the survey.

A study was carried out, testing our final iteration of the self-assessment tool on patients alongside the clinical version of the widely used TNS measure. Each patient was scored with both tools to enable a comparison between them. There were 17 participants in the study and we arrived at a correlation between the two tools of 0.861, which is generally considered to be a high correlation [68]. The result will be discussed in more detail further in this chapter.

6.2 Design

When moving a measure of medical relevance from the clinic to an in-home setup, great care and effort must be put into the implementation. Though measures of neuropathy performed by medical professionals aren't infallible, the risk of errors is even greater without

their presence. Through UX design methods and consultations with medical specialists in the field, we have worked towards making the self-assessment as easy to perform and reliable as possible.

A major ambition of the project was to make the self-assessment tool accessible, both in terms of usage, and the required materials. The common measures of neuropathy are not publicly available, and many rely on specialized equipment, greatly limiting the accessibility of the measurements. Mobile applications have great reach due to the amount of people owning smartphones. As mHealth apps are increasingly seen as potential platforms for self-assessment, it makes sense to use them to reach people who would benefit from such self-assessment tools [18]. The vibration functionality of smartphones also has the the potential to act as a substitute for the tuning fork, which is used by many neuropathy measures.

From both usability testing and the study we have seen that most of the tests in the survey are easy to conduct and users can respond to questions with confidence. Motor examination is an example of a test that didn't change from the first design iteration to the last, due to it never causing difficulty for the users. On the other hand, the pin-prick test changed with every iteration, and it wasn't until the last one that the instructions were conveyed in a way that users fully understood. Validating and tailoring an mHealth application to a user with no medical background is not easy and doesn't always fully succeed. During a feasibility study of the NeuroDetect app [48], 58% of the study participants reported having issues with the study questionnaires. Further 21% of participants had issues with the functional tests of the product.

There were also some design problems that weren't solved in this project. Most notably, *"none"* is a confusing option when users are asked to compare pin-prick sensations, and that users had difficulty telling vibration felt in the feet apart from vibration in the hand. Possible solutions to this problem, that weren't investigated, could be using a reference point, asking someone else to hold the phone or putting padding between the phone and hand to reduce vibration felt in the hand.

From the study it was seen that patients, even those whose ability to move is impacted by their neuropathy, were able to carry out the steps of the survey. This indicates that the majority of the target users, people at risk of neuropathy or those with early neuropathy, would be able to use the self-assessment tool. It was rare for participants to ask questions due to not understanding tests. When they did ask questions they were mainly asked for confirmation if they performed tests correctly. We find that this is to be expected when partaking in a study under observation, and do not see it as alarming.

There was one participant of the study who had major errors in how they reported results in the survey. This was surprising, as no other instances were encountered of users consistently giving responses that differed from their experiences. While UX design aims to make a product as easy to use as possible, this is an example showing that it cannot guarantee that the product will suit everyone. It indicates that there may still be a need for either simplification of instructions or clearer communication.

In the last design iteration of the project, the *Examination Completed* page does not offer a diagnosis, and was iterated upon to dissuade users from inferring one. This is due to the restrictions that apply to products that can be considered medical devices, which includes apps that are intended for use in diagnosis of diseases [69]. Since this project is an investigation of the feasibility of such tool, implying a diagnosis at this stage of development could be dangerous and incorrect, as it has not undergone extensive validation.

Only the survey part of the application was validated, though there are other parts that would benefit from validation as well. The *Settings* page (Figure 4.8a), where users can enter basic personal information, choose language, download or delete all previous results and set vibration duration, needs validation. Especially the vibration duration setting, which is used for circumventing hardware limitations and therefore is more technical that the other settings. The *Detailed Results* page should ideally be easy to interpret for users interested in specifics about their condition, as well as medical professionals, who should be able to use the results to assist in diagnosis. That extends to the result exporting functionality where it should be easy for users to export results and deliver them to medical professionals, and for the medical professionals to import and get useful information from the results.

It was decided to make the app completely offline for this project, since adding it would impose regulations on data privacy and security, which would only distract from the goal of the project. This is why the two options for sharing data is by showing the app to someone else or downloading results and sending them through a different service. Data sharing, especially with medical professionals, could be made much more convenient by implementing an online functionality in the future.

6.3 Grading Tool Validation

The validation study was a small scale study that was meant to indicate if the tool can be considered for further, full evaluation of its clinical application and validity. The results from self assessment have shown high positive correlation to the TNSc measure.

The Pearson correlation of 0.861 that our tool shows with TNSc is comparable to how well different measures correlate to each other. When UENS was introduced, its correlation to two other state-of-art tools were 0.895 and 0.863 [41]. The correlation of our tool was investigated using patients on a wide spectrum of the grading scale (see Figure 5.4). This indicates that our tool agrees with TNSc on the whole scale evenly.

The study was performed in a clinical setting, and with the scoring correlating in a satisfactory degree, we believe it would perform well when used at home, as the patients required little to no guidance when using our application. This would require additional study to validate, for which we did not have available resources or time. Unsupervised examinations carry the risk of errors, which became apparent during our validation as well. A broader study would better measure the frequency of the issue.

It would be beneficial to expand the validation to compare the self-assessment tool to additional measures as well, with a detailed study of the correlation of similar segments of the measures - such as pin-prick or large fiber sensation. This would eliminate the risk of bias, and provide better insights to which areas of the examination may be a weak point of the tool.

6.4 Limitations

There are several limitations of the project that surfaced during design, implementation and evaluation. These points raise some concerns and require addressing:

- As previously mentioned, vibration aspect of both design and implementation may be unreliable in the current form. Due to hardware limitations - testing the application on two mobile phones owned by the authors, it is not certain how it will perform on different devices. Even with the two devices, there were issues with how long the phone was able to vibrate. Additionally, patients struggled to differentiate if they felt the vibration in the point of the leg or in the hand holding the phone.

- The study contained a medical examination - TNSc, but it was not performed by clinicians. This could lead to inaccurate measurements, especially in testing deep tendon reflexes. The technique is not easy to master, and even with training we obtained, we cannot be sure it was performed reliably by us.
- Unsupervised testing and usage of the application may lead to more errors than we were able to perceive during the study. One of the patients selected the opposite answers to most questions, and it was only noticed because there was someone observing the examination. This case happened for only one of the 17 participants, but it should be a concern and question in a larger-scale study.
- As the country of the project and the performed study was Denmark, only the Danish version of the grading tool was validated. The application is available in English, but we cannot speak on its correlation with TNSc.
- Parts of the final implemented design were not validated, and may require more development and adjustments. Segments that were not subject to an usability study were the settings, result export and detailed results page.
- Due to lack to access to iOS devices during implementation, it is not certain if the application performs as intended on that system.

7 Conclusion

Data analysis at the end of the project implied that it is possible for patients to use a mobile self-assessment tool for polyneuropathy with a high positive correlation to a diagnostic measure used by clinicians.

We developed a design of an mHealth grading tool for polyneuropathy, an innovation that has not previously been widely implemented, used and studied. The fundamental objective was shifting the user of a state-of-art measure from the clinician to the patient, providing a way for patients to conduct the examination themselves. Consulting with the clinical supervisor, we have modified and expanded the UENS, creating a questionnaire consisting of tests and question steps.

The new grading tool encapsulates a wide range of findings and symptoms of polyneuropathy, including small and large fiber sensation, hyperesthesia and motor strength tests that can be performed without the presence of medical professionals or specialized tools. Like the UENS, it still emphasises signs of early neuropathy, as over half of the points possible to obtain comes from the pin-prick sensation. In the same tool we included questions from the DN4 questionnaire, that do not contribute to the score, but provide context about the pain the patient may be experiencing. The total length of the questionnaire was 38 steps, with 0-44 point range, where more points indicate more findings and symptoms of polyneuropathy.

The user - a patient or person at risk of neuropathy, was the center of the design. The clinical measures, meant for medical professionals, rarely provide precise directions and instructions, and the phrasing used is complicated and not accessible to a user with no medical training. With an examination meant to be carried out without clinical supervision, it is imperative to be as certain as possible the patient will perform the tasks as intended. This is why we compiled a set of precise instructions, with pictures and additional information available. During usability testing we identified flaws and corrected them as much as possible. We noticed the users were not necessarily applying all instructions, especially in the pin-prick section, so the pictures were expanded to include all the steps and more precise depiction of sections meant to be tested.

The questionnaire has been implemented as an mHealth application, and the only things necessary to perform the examination are a safety pin and a mobile device with the application installed. The application is not available in any mobile stores, but can be compiled and installed directly from the code. We implemented the tool as a survey and added result screens, settings and a home page displaying all previous examinations. The application language can be English or Danish. Only the grading tool in Danish was validated during the study, as it was the focal part of the project.

A validation study was performed with 17 patients of the outpatient clinic at the Neurological Department of Zealand University Hospital. Even though the application is meant to be used outside of the clinical setting, we decided it was important at this stage to observe the use of the application, so the study was conducted at the hospital. The patients used our tool by themselves without assistance, but due to our presence in the room they sometimes asked questions, mostly to reassure that they were performing the instructions correctly. Despite the efforts to design precise instructions, some points of the examination still sparked confusion, especially sensing the vibration. The vibration was a challenge due to two reasons - phones can have different vibration frequencies,

and holding the phone in your hand while pressing it to your leg can make it difficult to ascertain if you're sensing it in the examined point. The instructions for this part could still be improved, for example by providing an reference area, such as is done with the pin-prick examination, provide padding between the hand and phone, or asking someone else to hold the phone if possible. One of the patients had marked opposite answers for pin-prick and vibration questions, seemingly not understanding the phrasing. This is a major concern with unsupervised examination, and it shows it cannot be assumed all users will receive accurate results.

For all patients that participated in the study, after they completed the self-assessment using our tool, we performed a TNSc examination. Since it was not performed by trained physicians, its results might not perfectly resemble those that medical professionals would obtain. The results from our tool and TNSc were analyzed, and showed a high positive correlation of 0.86. This is comparable to how well clinical measures correlate to each other. This data suggest that the new questionnaire might be a useful tool to assist in polyneuropathy detection or diagnosis.

The data we obtained is not enough to prove that the tool can be used to detect early neuropathy. It would require a more extensive study, with a much larger number of participants. The same is true for using the tool in a non-clinical setting. Though the patients did not take the phone or application home, they performed the self-assessment without our intervention or guidance, which leads us to believe that it can be used this way. The design includes data export, meaning that the application in its current state can be used for further studies.

In summary, the research question was answered, indicating that the proposed design and implementation shows potential to become a tool to assist clinicians in diagnosing polyneuropathy, designed to be performed by patients themselves. While showing good correlation with state-of-art neuropathy measures, further study including medical professionals is required to fully validate that, as well as ascertain the ability of the tool to detect early neuropathy.

Bibliography

[1] AH RAJPUT. “POLYNEUROPATHY”. eng. In: *Canadian Family Physician* 20.12 (1974), pp. 37–41. ISSN: 0008350x, 17155258.

[2] Michael Benatar. “Polyneuropathy”. eng. In: *Neuromuscular Disease* (2006), pp. 109–147. DOI: 10.1007/978-1-59745-106-2_8.

[3] Colin Quinn and Raymond S. Price. “Distal Symmetric Polyneuropathy”. eng. In: *Decision-making in Adult Neurology* (2020), 182, 183.e1, 183, 183.e2. DOI: 10.1016/B978-0-323-63583-7.00089-8.

[4] Brian C. Callaghan, Raymond S. Price, and Eva L. Feldman. “Distal symmetric polyneuropathy a review”. eng. In: *Jama - Journal of the American Medical Association* 314.20 (2015), pp. 2172–2181. ISSN: 15383598, 00987484. DOI: 10.1001/jama.2015.13611.

[5] Jennifer S. Gewandter et al. “Clinician-rated measures for distal symmetrical axonal polyneuropathy: ACTTION systematic review”. In: *Neurology* (2019). DOI: 10.1212/wnl.0000000000007974.

[6] Cleveland Clinic. *Cleveland Clinic Website*. URL: https://my.clevelandclinic.org/health/diseases/14737-peripheral-neuropathy (visited on 02/11/2023).

[7] Gabriel Hajaš. “Diabetic neuropathy in children and adolescents II - Importance of early diagnosis”. slo. In: *Lekarsky Obzor* 57.4 (2008), pp. 162–169. ISSN: 04574214.

[8] Rita McMorrow, Vanessa L. Nube, and Jo Anne Manski-Nankervis. “Preventing diabetes-related foot ulcers through early detection of peripheral neuropathy”. eng. In: *Australian Journal of General Practice* 51.11 (2022), pp. 833–838. ISSN: 2208794x, 22087958. DOI: 10.31128/AJGP-06-22-6456.

[9] Muneo Matsunaga. “Diabetic autonomic neuropathy: Importance of its early detection for improving morbidity and mortality”. eng. In: *Internal Medicine* 42.7 (2003), pp. 545–546. ISSN: 13497235, 09182918. DOI: 10.2169/internalmedicine.42.545.

[10] Solomon Tesfaye et al. “Diabetic Neuropathies: Update on Definitions, Diagnostic Criteria, Estimation of Severity, and Treatments”. In: *Diabetes Care* 33.10 (Oct. 2010), pp. 2285–2293. DOI: 10.2337/dc10-1303. URL: https://doi.org/10.2337/dc10-1303.

[11] John D. England et al. “Distal symmetric polyneuropathy: A definition for clinical research Report of the American Academy of Neurology, the American Association of Electrodiagnostic Medicine, and the American Academy of Physical Medicine and Rehabilitation”. In: *Neurology* (2005). DOI: 10.1212/01.wnl.0000149522.32823.ea.

[12] Fran Biggin et al. “Variation in waiting times by diagnostic category: an observational study of 1,951 referrals to a neurology outpatient clinic”. In: *BMJ neurology open* 3.1 (2021).

[13] Donald L Patrick et al. “Patient-reported outcomes to support medical product labeling claims: FDA perspective”. In: *Value in Health* 10 (2007), S125–S137.

[14] Prasanna R Deshpande et al. “Patient-reported outcomes: A new era in clinical research”. eng. In: *Perspectives in Clinical Research* 2.4 (2011), pp. 137–144. ISSN: 22295488, 22293485. DOI: 10.4103/2229-3485.86879.

[15] Nanja van Geel et al. “Development and validation of a patient-reported outcome measure in vitiligo: The Self Assessment Vitiligo Extent Score (SA-VES)”. In: *Journal of the American Academy of Dermatology* 76.3 (2017), pp. 464–471. ISSN: 0190-9622. DOI: https://doi.org/10.1016/j.jaad.2016.09.034. URL: https://www.sciencedirect.com/science/article/pii/S0190962216308763.

[16] Skyrocket Phytopharma. *NeuroPad Website*. URL: https://www.neuropad.co.uk/ (visited on 02/22/2023).
[17] N. Papanas et al. “A Prospective Study on the use of the Indicator Test Neuropad® for the Early Diagnosis of Peripheral Neuropathy in type 2 Diabetes”. In: *Experimental and Clinical Endocrinology & Diabetes* 119.02 (Aug. 2010), pp. 122–125. DOI: 10.1055/s-0030-1261934. URL: https://doi.org/10.1055/s-0030-1261934.
[18] Tara McCurdie et al. “mHealth Consumer Apps: The Case for User-Centered Design”. In: *Biomedical Instrumentation & Technology* 46.s2 (Sept. 2012), pp. 49–56. DOI: 10.2345/0899-8205-46.s2.49. URL: https://doi.org/10.2345/0899-8205-46.s2.49.
[19] Hanbin Zhang et al. “mHealth Technologies Towards Parkinson’s Disease Detection and Monitoring in Daily Life: A Comprehensive Review”. In: *IEEE Reviews in Biomedical Engineering* 14 (2021), pp. 71–81. DOI: 10.1109/RBME.2020.2991813.
[20] Michelline Joana Tenório Albuquerque Madruga Mesquita et al. “A mhealth application for automated detection and diagnosis of strabismus”. eng. In: *International Journal of Medical Informatics* 153 (2021), p. 104527. ISSN: 18728243, 13865056. DOI: 10.1016/j.ijmedinf.2021.104527.
[21] Miguel A. Camara et al. “MHealth tools for monitoring Obstructive Sleep Apnea patients at home: Proof-of-concept”. eng. In: *Proceedings of the Annual International Conference of the Ieee Engineering in Medicine and Biology Society, Embs* 2017 (2017), pp. 1555–1558. ISSN: 1557170x, 15584615, 26940604. DOI: 10.1109/EMBC.2017.8037133.
[22] Ivan Varga and Boris Mravec. “Chapter 8 - Nerve Fiber Types”. In: *Nerves and Nerve Injuries*. Ed. by R. Shane Tubbs et al. San Diego: Academic Press, 2015, pp. 107–113. ISBN: 978-0-12-410390-0. DOI: https://doi.org/10.1016/B978-0-12-410390-0.00008-1. URL: https://www.sciencedirect.com/science/article/pii/B9780124103900000081.
[23] Hyun-Jung Yu and Seong-Ho Koh. “Overview of symptoms, pathogenesis, diagnosis, treatment, and prognosis of various acquired polyneuropathies”. In: *Hanyang Medical Reviews* 37.1 (2017), pp. 34–39.
[24] Sasha Smith et al. “Pathogenesis of Distal Symmetrical Polyneuropathy in Diabetes”. In: *Life* 12.7 (July 2022), p. 1074. DOI: 10.3390/life12071074. URL: https://doi.org/10.3390/life12071074.
[25] Lisette R.M. Raasing et al. “Current View of Diagnosing Small Fiber Neuropathy”. In: *Journal of Neuromuscular Diseases* 8.2 (Mar. 2021), pp. 185–207. DOI: 10.3233/jnd-200490. URL: https://doi.org/10.3233/jnd-200490.
[26] Caitlin W. Hicks and Elizabeth Selvin. “Epidemiology of Peripheral Neuropathy and Lower Extremity Disease in Diabetes”. In: *Current Diabetes Reports* 19.10 (Aug. 2019). DOI: 10.1007/s11892-019-1212-8. URL: https://doi.org/10.1007/s11892-019-1212-8.
[27] David Kabagema Tumusiime et al. “Prevalence of peripheral neuropathy and its associated demographic and health status characteristics, among people on antiretroviral therapy in Rwanda”. In: *BMC Public Health* 14.1 (Dec. 2014). DOI: 10.1186/1471-2458-14-1306. URL: https://doi.org/10.1186/1471-2458-14-1306.
[28] P Kersten. “Chapter 3 - Principles of physiotherapy assessment and outcome measures”. In: *Physical Management in Neurological Rehabilitation (Second Edition)*. Ed. by Maria Stokes. Second Edition. Oxford: Mosby, 2004, pp. 29–46. ISBN: 978-0-7234-3285-2. DOI: https://doi.org/10.1016/B978-072343285-2.50007-3. URL: https://www.sciencedirect.com/science/article/pii/B9780723432852500073.
[29] Jonathan Cole. “Large-fiber Sensory Neuropathy: Effect On Proprioception”. In: *Encyclopedia of Neuroscience*. Ed. by Marc D. Binder, Nobutaka Hirokawa, and Uwe

Windhorst. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 2105–2107. ISBN: 978-3-540-29678-2. DOI: 10.1007/978-3-540-29678-2_2696. URL: https://doi.org/10.1007/978-3-540-29678-2_2696.

[30] Susan A. Darby and Robert J. Frysztak. “Chapter 9 - Neuroanatomy of the Spinal Cord”. In: *Clinical Anatomy of the Spine, Spinal Cord, and Ans (Third Edition)*. Ed. by Gregory D. Cramer and Susan A. Darby. Third Edition. Saint Louis: Mosby, 2014, pp. 341–412. ISBN: 978-0-323-07954-9. DOI: https://doi.org/10.1016/B978-0-323-07954-9.00009-8. URL: https://www.sciencedirect.com/science/article/pii/B9780323079549000098.

[31] H Kenneth Walker, W Dallas Hall, and J Willis Hurst. *Clinical methods: the history, physical, and laboratory examinations*. Butterworths, 1990.

[32] Steven McGee. “Chapter 62 - Examination of the Sensory System”. In: *Evidence-Based Physical Diagnosis (Fourth Edition)*. Ed. by Steven McGee. Fourth Edition. Philadelphia: Elsevier, 2018, 569–582.e3. ISBN: 978-0-323-39276-1. DOI: https://doi.org/10.1016/B978-0-323-39276-1.00062-7. URL: https://www.sciencedirect.com/science/article/pii/B9780323392761000627.

[33] R. A. Malik et al. “Small fibre neuropathy: role in the diagnosis of diabetic sensorimotor polyneuropathy”. In: *Diabetes/Metabolism Research and Reviews* 27.7 (2011), pp. 678–684. DOI: https://doi.org/10.1002/dmrr.1222. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/dmrr.1222. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/dmrr.1222.

[34] Jinny Tavee. “Chapter 14 - Nerve conduction studies: Basic concepts”. In: *Clinical Neurophysiology: Basis and Technical Aspects*. Ed. by Kerry H. Levin and Patrick Chauvel. Vol. 160. Handbook of Clinical Neurology. Elsevier, 2019, pp. 217–224. DOI: https://doi.org/10.1016/B978-0-444-64032-1.00014-X. URL: https://www.sciencedirect.com/science/article/pii/B978044464032100014X.

[35] H Kenneth Walker. “Deep tendon reflexes”. In: *Clinical Methods: The History, Physical, and Laboratory Examinations. 3rd edition* (1990).

[36] Kheng Seang Lim et al. “Wide range of normality in deep tendon reflexes in the normal population.” In: *Neurology Asia* 14.1 (2009).

[37] Henning Andersen. “Chapter 7 - Motor neuropathy”. In: *Diabetes and the Nervous System*. Ed. by Douglas W. Zochodne and Rayaz A. Malik. Vol. 126. Handbook of Clinical Neurology. Elsevier, 2014, pp. 81–95. DOI: https://doi.org/10.1016/B978-0-444-53480-4.00007-2. URL: https://www.sciencedirect.com/science/article/pii/B9780444534804000072.

[38] *Paresthesia*. URL: https://www.ninds.nih.gov/health-information/disorders/paresthesia#:~:text=Paresthesia%20refers%20to%20a%20burning,%2C%20skin%20crawling%2C%20or%20itching. (visited on 08/13/2023).

[39] Robert J. Maldonado and Orlando De Jesus. *Hyperesthesia*. StatPearls Publishing, Treasure Island (FL), 2022. URL: http://europepmc.org/books/NBK563125.

[40] Troels S Jensen and Nanna B Finnerup. “Allodynia and hyperalgesia in neuropathic pain: clinical manifestations and mechanisms”. In: *The Lancet Neurology* 13.9 (2014), pp. 924–935.

[41] J Robinson Singleton et al. “The Utah Early Neuropathy Scale: a sensitive clinical scale for early sensory predominant neuropathy”. In: *Journal of the Peripheral Nervous System* 13.3 (2008), pp. 218–227.

[42] Ellen M Lavoie Smith. “The total neuropathy score: a tool for measuring chemotherapy-induced peripheral neuropathy”. In: *Number 1/January 2008* 35.1 (1969), pp. 96–102.

[43] Guido Cavaletti et al. “Multi-center assessment of the Total Neuropathy Score for chemotherapy-induced peripheral neurotoxicity”. In: *Journal of the Peripheral Nervous System* 11.2 (June 2006), pp. 135–141. DOI: 10.1111/j.1085-9489.2006.00078.x. URL: https://doi.org/10.1111/j.1085-9489.2006.00078.x.

[44] V. Bril et al. “Reliability and validity of the modified Toronto Clinical Neuropathy Score in diabetic sensorimotor polyneuropathy”. In: *Diabetic Medicine* 26.3 (Mar. 2009), pp. 240–246. DOI: 10.1111/j.1464-5491.2009.02667.x. URL: https://doi.org/10.1111/j.1464-5491.2009.02667.x.

[45] Didier Bouhassira et al. “Comparison of pain syndromes associated with nervous or somatic lesions and development of a new neuropathic pain diagnostic questionnaire (DN4)”. In: *Pain* 114.1 (Mar. 2005), pp. 29–36. DOI: 10.1016/j.pain.2004.12.010. URL: https://doi.org/10.1016/j.pain.2004.12.010.

[46] Apostolos Tsapas et al. “A simple plaster for screening for diabetic neuropathy: a diagnostic test accuracy systematic review and meta-analysis.” In: *Metabolism-clinical and Experimental* (2014). DOI: 10.1016/j.metabol.2013.11.019.

[47] Pok Ja Oh and Jung Ran Lee. “Development and Effects of a Mobile Application-based Self-Management Program for Chemotherapy-induced Peripheral Neuropathy in Colorectal Cancer Patients”. In: *Korean Journal of Adult Nursing* (2022). DOI: 10.7475/kjan.2022.34.3.258.

[48] Ciao-Sin Chen et al. “Chemotherapy-Induced Peripheral Neuropathy Detection via a Smartphone App: Cross-sectional Pilot Study”. In: *Jmir mhealth and uhealth* (2021). DOI: 10.2196/27502.

[49] The University of Michigan. *NeuroDetect on the App Store*. URL: https://apps.apple.com/app/neurodetect/id1457083267 (visited on 08/12/2023).

[50] Peter G. Jacobs et al. “Design and evaluation of a portable smart-phone based peripheral neuropathy test platform”. In: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (2018). DOI: 10.1109/embc.2018.8513100.

[51] Pamela L. Hudak et al. “Development of an upper extremity outcome measure: The DASH (disabilities of the arm, shoulder, and head)”. In: *American Journal of Industrial Medicine* 29.6 (June 1996), pp. 602–608. DOI: 10.1002/(sici)1097-0274(199606)29:6<602::aid-ajim4>3.0.co;2-l. URL: https://doi.org/10.1002/(sici)1097-0274(199606)29:6%3C602::aid-ajim4%3E3.0.co;2-l.

[52] Neal B. Zimmerman et al. “Are standardized patient self-reporting instruments applicable to the evaluation of ulnar neuropathy at the elbow?” In: *Journal of Shoulder and Elbow Surgery* 18.3 (May 2009), pp. 463–468. DOI: 10.1016/j.jse.2009.02.010. URL: https://doi.org/10.1016/j.jse.2009.02.010.

[53] Interactive Design Foundation. *What is User Experience (UX) Design?* URL: https://www.interaction-design.org/literature/topics/ux-design/ (visited on 07/23/2023).

[54] Chadia Abras, Diane Maloney-Krichmar, Jenny Preece, et al. “User-centered design”. In: *Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications* 37.4 (2004), pp. 445–456.

[55] Laura Klein. *UX for lean startups: Faster, smarter user experience research and design*. ” O’Reilly Media, Inc.”, 2013.

[56] J. Nielsen. “Iterative user-interface design”. In: *Computer* 26.11 (Nov. 1993), pp. 32–41. DOI: 10.1109/2.241424. URL: https://doi.org/10.1109/2.241424.

[57] Jinho Yim, Rohae Myung, and Byongjun Lee. “The mobile phone’s optimal vibration frequency in mobile environments”. eng. In: *Lecture Notes in Computer Science (including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in*

Bioinformatics) 4559.1 (2007), pp. 646–652. ISSN: 16113349, 03029743. DOI: 10.1007/978-3-540-73287-7_75.

[58] Agency for Clinical Innovation. *DN4 Questionnaire*. URL: https://aci.health.nsw.gov.au/__data/assets/pdf_file/0014/212900/DN4_Assessment_Tool.pdf (visited on 07/30/2023).

[59] Nicola Crichton. "Visual analogue scale (VAS)". In: *J Clin Nurs* 10.5 (2001), pp. 706–6.

[60] Figma. *Figma: The Collaborative Interface Design Tool*. URL: https://www.figma.com/ (visited on 07/26/2023).

[61] Copenhagen Center for Health Technology. *Carp Research Package*. URL: https://carp.cachet.dk/research-package/ (visited on 05/11/2023).

[62] tekartik. *Sembast*. URL: https://pub.dev/packages/sembast (visited on 07/31/2023).

[63] flutter.dev. *Shared Preferences Plugin*. URL: https://pub.dev/packages/shared_preferences (visited on 07/31/2023).

[64] skillzupp.com. *to_csv*. URL: https://pub.dev/packages/to_csv (visited on 07/31/2023).

[65] Copenhagen Center for Health Technology. *Carp Website*. URL: https://carp.cachet.dk/ (visited on 02/12/2023).

[66] Benjamin Dean. *Vibration*. URL: https://pub.dev/packages/vibration (visited on 08/04/2023).

[67] cachet.dk. *carp_serializable*. URL: https://pub.dev/packages/carp_serializable (visited on 08/03/2023).

[68] M. M. Mukaka. "Statistics corner: A guide to appropriate use of correlation coefficient in medical research". In: *Malawi Med J* 24.3 (Sept. 2012), pp. 69–71.

[69] *How to determine if your product is a medical device*. URL: https://www.fda.gov/medical-devices/classify-your-medical-device/how-determine-if-your-product-medical-device (visited on 08/18/2023).

A Acronyms

BLoC Business Logic Component

CACHET Copenhagen Center for Health Technology

CARP Copenhagen Research Platform

CIPN Chemotherapy-Induced Peripheral Neuropathy

DN4 Douleur Neuropathique 4 Questions

DPN Diabetic Peripheral Neuropathy

DSP Distal Symmetrical Axonal Polyneuropathy

DTU Technical University of Denmark

mHealth Mobile Health

mTCNS modified Toronto Clinical Neuropathy Score

MVP Minimum Viable Product

NCS Nerve Conduction Study

NoSQL Not Only SQL

PRO Patient-Reported Outcome

RP Research Package

SFN Small Fiber Neuropathy

TNS Total Neuropathy Score

TNSc Total Neuropathy Score clinical

UCD User-Centered Design

UENS Utah Early Neuropathy Scale

UI User Interface

UX User Experience

B Results

B.1 Mobile Health (mHealth) Application

The source code for the implemented application can be found in a Github repository by following this link: https://github.com/jeyjey626/neuropathy-app.

B.2 Validation Study Data

Table B.1 presents the data points of the validation study. The date of appointment with each participant, their sex and age are listed. The table contains both scores obtained with TNSc and the project's self-assessment tool (Application score).

Table B.1: Data gathered from the validation study

Date (DD.MM.YYYY)	Sex	Age	TNSc Score	Application score
05.06.2023	Female	42	2	7
05.06.2023	Male	41	15	19
05.06.2023	Female	40	1	2
06.06.2023	Female	55	16	22
06.06.2023	Female	48	19	37
07.06.2023	Female	55	4	2
07.06.2023	Female	48	6	5
07.06.2023	Male	67	10	25
07.06.2023	Female	75	13	25
08.06.2023	Male	58	24	26
08.06.2023	Female	73	10	6
08.06.2023	Male	65	7	0
08.06.2023	Female	51	4	7
09.06.2023	Female	63	16	20
09.06.2023	Male	40	13	13
13.06.2023	Female	58	24	37
13.06.2023	Female	52	9	12

C Tables of contributions

C.1 Theoretical work

Chapter or section	Julia	Mads
1 Introduction	X	X
1.1 Research Question	X	
1.2 Research goals and methods	X	
1.3 Thesis overview		X
2 Background and Prior Work	X	
2.1 Medical background	X	
2.2 Diagnostic measures for peripheral neuropathy	X	X
2.3 Existing tools for self-assessment of neuropathy		X
3 Research Methods		X
3.1 User Experience (UX) design methods		X
3.2 Grading tool design	X	
3.3 First design iteration		X
3.4 Second design iteration		X
4 Mobile Health Application for Self-Assessment of Neuropathy	X	
4.1 Design	X	
4.2 Implementation	X	X
5 Validation Study		X
5.1 Methods		X
5.2 Results		X
6 Discussion		X
6.1 Research question		X
6.2 Design	X	X
6.3 Grading Tool Validation	X	
6.4 Limitations	X	
7 Conclusion	X	

Table C.1: Contributions to the thesis. The **chapter** contributions only encompass the text before the first section.

C.2 Empirical work

Area	Task	Julia	Mads
Software structure and collaboration	Creating GitHub repository and workflow	X	
	Creating GitHub pull requests	X	X
	Reviewing and merging GitHub pull requests	X	X
	Application structure	X	
	Code documentation	X	
Validation and study	Defining design validation goals	X	X
	Conducting design validation		X
	Extracting learnings from design validation		X
	Conducting study	X	X
	Extracting study results		X
Design and implementation of user interface	Main page	X	
	General symptoms section	X	
	Pin-prick sensation section		X
	Allodynia/Hyperalgesia section	X	
	Neuropathic pain section	X	
	Large fiber sensation section	X	
	Motor examination section		X
	Examination complete page		X
	Detailed results page	X	
	Settings page	X	
Other software components	Implementing result and settings database	X	
	Implementing result exporting	X	
	Score calculation	X	
	Making and translating localization files		X
	Implementing localization	X	

Table C.2: Contributions on the empirical work.

Technical
University of
Denmark

Richard Petersens Plads, Building 324
2800 Kgs. Lyngby
Tlf. 4525 1700

www.healthtech.dtu.dk