Accessibility in health mobile applications

University of Oulu
Information Processing Science
Master's Thesis
Maria Ollila
2023

Abstract

Nowadays, there is a vast number of mobile devices capable of storing an individual's entire life. There are applications for everything, from banking to ordering food and clothes, but also different health applications targeted towards different impairments and self-health care management. Self-health care management applications can have a significant impact on individuals with various diseases and impairments. However, it is essential that these applications are accessible to users with different impairments such as motor and vision impairments.

The purpose of this study was to examine accessibility concerns in mobile health applications for individuals with multiple sclerosis and evaluate how these concerns were addressed. Multiple sclerosis was chosen as the focus of this study because its symptoms encompass a range of impairments, including vision, motion, hearing, and cognitive limitations.

The study was conducted with benchmarking multiple sclerosis applications obtained in Google Play store. Benchmarking focused on accessibility, and measurements and metrics were gathered testing applications with Google Accessibility Scanner and TalkBack screen reader. Measurements were based on web content accessibility guidelines (WCAG) 2 and accessibility guidelines for mobile applications.

None of the tested applications followed accessibility guideline requirements based on benchmarking metrics. When examining the metrics from the perspective of impairments, it was found that applications had accessibility concerns related to motor and vision impairments. The applications addressed requirements for hearing impairments in applicable features, while testing cognitive impairment requirements proved challenging with the selected testing tools. In the future, it is recommended to conduct additional accessibility testing for cognitive impairments using methods such as manual accessibility testing and user testing.

Keywords
Accessibility, Mobile application accessibility, Mobile health applications, mHealth, Multiple sclerosis

Supervisor
MD PhD, Senior Researcher Guido Giunti

Foreword

This thesis took a long time. When I first started studying in the university of Oulu, I was certain that I was going to graduate in the target time, but life had other plans. I want to give thanks to my thesis supervisor Guido Giunti, who was patient with me and helped me through the whole process.

I also want to thank all my friends who have been there for me during my time in the university and during my thesis process.

Maria Ollila

Oulu 17.6.2023

Contents

1. Introduction

The number of mobile devices is huge now a days and a mobile device can hold one person's whole life inside it. In 2021, 84% of population in Finland used internet with their mobile phones (Official Statistics of Finland (OSF), 2021). Worldwide, approximately 4 billion people used the internet on their mobile phones and predictions suggest that the number of users will reach 6.129 billion by the year 2028 (Ceci, 2023). There are applications for banking, shopping, and ordering food. It is feasible for individuals to live their life inside their own home and never leave since tasks related to everyday life can be done with the mobile devices. Simultaneously, the availability of physical services has been decreasing or shifted towards time reservations, which makes need and use of different mobile applications even higher.

The significance and ubiquity of mobile devices in current society should prioritize ensuring effortless access to various applications for individuals with disabilities. This raises question about accessibility of mobile applications, that could these applications be used by everyone, and if not, then who are left out. Accessibility challenges can be physical or cognitive and solutions to these challenges can differ greatly. Users with physical disabilities have problems with interacting the devices, for example touch screens and users with cognitive disabilities might have difficulties to understand the content (Dias et al., 2012).

Mobile health or mHealth are mobile applications related to health and wellbeing used in mobile phones and wearable technologies (Cao et al., 2021). They can be used for stress relief, weight loss or to provide health information (Whittaker, 2012). The mHealth business globally was 40 billion U.S. dollars in 2018 and it was expected to grow to 332.7 billion U.S. dollars by year 2025 (Statista, 2018). Definition could be simplified that mHealth are health-related applications or services that can be accessed via mobile devices (Whittaker, 2012).

Multiple Sclerosis (MS) is a neurological condition that affects people in their most productive time of their life. MS may be one of most prevalent condition with accessibility concerns, because it impacts individual with fatigue, motor impairments, visual impairments, and cognitive impairments (Cao et al., 2021). Challenge with motor impairments for example paralysis, and tremors is usage of mobile devices by tapping (Ruzic & Sanford, 2017). MS patients with visual impairments, like partial blindness, have challenges with reading text, that can be too small or not enough contrast between text and background. Patients with cognitive impairments, meaning problems with memory, thinking and way of viewing things, could have challenges with complex interfaces and mobile menus. (Ruzic & Sanford, 2017).

Presented challenges above and other challenges which are related to different impairments can be met with different guidelines. One for website accessibility is Web Content Accessibility Guidelines 2 by W3C (W3C, 2022d). They also have own sections for mobile accessibility and cognitive accessibility (W3C, 2022a, 2022b). Also, mobile device platforms, Android and Apple has their own guidelines and guides on how to make accessible applications (Apple Inc, 2022; Google, 2022).

The present manuscript's research question is: how are accessibility concerns addressed in the state of the practice of mobile health applications for MS?

2. Prior research

Accessibility has been extensively studied over the years. The focus has been primarily on website accessibility, but mobile accessibility research has grown in recent years. This chapter describes various health impairments, accessibility requirements related to impairments and methods for assessing identified requirements.

2.1 Frameworks for assessing accessibility

Numerous frameworks have been developed for assessing accessibility, one being Web Content Accessibility Guidelines (WCAG). WCAG 1.0. was published 1999 to provide guidelines for accessibility in web, ensuring usability for users with different disabilities regardless of the platform being used. WCAG 1.0 had checkpoints, and different levels of accessibility from A to AAA, with AAA being a level were all the checkpoints in the site were crossed. (W3C, 1999)

WCAG 2.0 was published in 2008, and it aimed to fix accessibility issues that were not covered in 1.0. It provided detailed instructions, more guidelines for audio and video, and changed from checkpoints to success criteria levels (W3C, 2009). The most recent version, WCAG 2.1 was published in 2018. Currently the World Wide Web Consortium (W3C) is developing version 2.2. expected to be published in third quarter of 2023, along with WCAG 3.0, a completely new version of the guidelines (W3C, 2023). WCAG 3.0 is supposed to cover more user needs, for example covering more needs of people with cognitive impairments and be easier to understood than the 2.X versions. It is currently in draft process and will take few more years before it is ready. (W3C, 2022b)

WCAG has become the standard accessibility guidelines and the Finnish the Act on the Provision of Digital Services requirements includes WCAG 2.1 and the levels A and AA (Regional State Administrative Agency, 2022). The WCAG 2.0 and 2.1 can be applied to mobile devices, and 2.1. added own success criteria for mobile accessibility, for example app can be used with horizontal and vertical orientation without restrictions from content (W3C, 2020).

Another notable set of guidelines for mobile accessibility are BBC Mobile Accessibility Guidelines. These guidelines are part of BBC’s Accessibility for Products and are also based on WCAG (Richens et al., 2022b). These guidelines have three principles which are then covered in aspect of audio & video, design, and structure. They provide clear instructions how to follow and test the guideline, along with examples for different platforms, such as Android and iOS (Richens et al., 2022a). The guidelines are clear, but they have not been as widely researched as WCAG.

There have also been studies focused on creating own guidelines for mobile based on existing standards. Ballantyne et al., (2018) found 92 unique accessibility guidelines in their study, which was done as literature review. Those guidelines were then categorized into 11 different categories, by purpose and benefit. Benefit meaning for which disability the category would help. They also categorized the guidelines into three levels: design, system, and content, for helping to find, identify and fix accessibility problem.

2.2 The importance of accessibility in health

Accessibility in health is important because digitalization is affecting health care in positive and negatives ways. mHealth has potential to enhance user's quality of life, and patients believe that mobile health technologies will improve cost, quality and conveniency of healthcare (Lippincot et al., 2020). Developers and designers without disabilities are primarily responsible for the development of mHealth applications, leading to underrepresentation of people with disabilities in the field. This underrepresentation could lead to wider inequality between people with disabilities and general population (Lippincot et al., 2020). Incorporating accessibility into design and development is one way to prevent inequality and get more representation from people with disabilities and convert them to users of applications.

Accessibility is not one solution fits for all type of thing. Accessibility is determined by the type of impairment individuals have and what prevents them from using services or applications. Impairments can be classified into four main categories: visual, hearing, motor and intellectual/cognitive (Siebra et al., 2017). Each category has its own unique needs, requirements, and solutions, which will be discussed in detail in chapters 2.2.1-2.2.4.

2.2.1 Cognitive impairment

People with cognitive impairments often experience limitations in mental tasks such as planning, understanding numbers and symbols, conceptualizing, and remembering (Niman et al., 2015). This impairment can arise from various factors, including fatigue, migraines, concussions, amnesia, neurodivergent conditions, like ADHD and dyslexia or memory and mental health conditions, such as dementia or learning delays (BBC, 2022). It can be argued that cognitive impairment affects individuals in different ways and is therefore important to consider when developing accessible mobile services. Especially in health context, when one is ill and should be using application but cannot used it because it is not accessible to them.

Cognitive impairments can be classified into three levels of severity: deep, moderate, and mild, adapted from World Health Organization (WHO) (Siebra et al., 2017). However, this study did not identify any specific requirements for cognitive impairments, as it did for other types of impairments.

In a study conducted in 2007, when mobile phones were slowly gaining popularity, researchers explored the requirements of young adults with cognitive disabilities regarding mobile phones and gathered insights from their parents. The study tried to find out how young adults with cognitive disabilities used phones and what were requirements and hopes from their parents. Most of the research participants used phones only for calling and answering calls. All the requirements found are related to smart phones and mobile apps nowadays. For example, research participants expressed the need for a simple menu system with fewer options for easier navigation on phones. (Dawe, 2007)

Guidelines that specifically address cognitive impairments include avoiding links violating general flash threshold and content not flashing more than 3 times in any 1 second period (Ballantyne et al., 2018). Niman et al. (2015) developed their own European Telecommunication Standards Institute (ETSI) guidelines for design of mobile ICT devices and their related applications for people with cognitive disabilities. The guidelines give detailed instructions on how to design accessible mobile apps for cognitive impairments, for example the app should always be consistent and change only when user is indenting the change, and the app should support individualization (CML, 2016).

Individuals with multiple sclerosis can experience cognitive impairments that affect their information processing and memory. Approximately 34% - 65 % of adult patients with multiple sclerosis, have cognitive challenges with varying severity and occurrences (Kalb et al., 2018). Symptoms related to MS patients with cognitive impairments may be slower information processing, problems with verbal fluency and difficulties understanding complex issues (Kalb et al., 2018). WCAG 2.1. provides following guidelines to help with these kinds of issues: 2.2. Enough time, 2.4. Navigable and 3.2. Predictable (W3C, 2022a).

Enough time means, that application should provide enough time for users to read and use content, so the user with slower information processing can understand the situation of application. Navigable provides help for navigation, finding content and determining where user is in the application, while predictable applications operate and appear in ways that are obvious and natural for the users. Navigable and predictable should help users who have difficulties understanding complex issues. (W3C, 2022a)

Overall, cognitive impairments are among the most challenging to consider since they often involve non-physical problems related to an individual's perception, experiences, and memory. They are more related to how individual sees, experiences and remembers the world. One of the essential accessibility requirements for cognitive impairments is that app is consistent how it provides information with simple user interface.

2.2.2 Motor impairment

Motor impairment users have problems related to their hands, like tremors, unable to hold objects in their hands, limbs, too large or too small hands to use devices or movement restrictions (BBC, 2022). These impairments make it difficult to interact with smart phones, since there are no physical buttons, which may cause frequent mis-clicks. Siebra et al., (2017) contemplated motor impairments, to be the most important impairment to be considered with accessibility and usability of mobile applications.

Motor impairments can affect older individuals. Their range of movement can be limited, or they can have tremors. Nicolau & Jorge, (2012) studied elderly people with tremors and how could they type with virtual keyboard in mobile phones and tablets. They found out that elderly individuals could write better with tablets than with mobile phones. Study suggested as possible solutions, for wider virtual keyboard keys, options for personalisation and addressing poor aiming instead of accidental swipes caused by tremors (Nicolau & Jorge, 2012).

Motor impairments, including tremors, dystonia, and ataxia, are highly prevalent among multiple sclerosis patients (Ghosh et al., 2022). Patients with ataxia cannot control their muscles in arms and legs, or the control is poor, and patients with dystonia have unintended muscle contracts. Ataxia requires possibility to use touch screens with alternative ways, and dystonia have same kind of requirements as tremors.

Users who lack hand control need alternative interaction methods for devices and applications, such as voice commands or eyeball tracking. Additional solutions include resizable touch, automatic selection of fields and avoiding time limits for functions.

(Siebra et al., 2017). Ballantyne et al., (2018) present 12 requirements for mobile impairments that align with Siebra et al., (2017) requirements. These requirements include ensuring touch targets are large enough, alternative methods to interact with the screen, such as using a keyboard, and that text entries should be reduced with different inputs, like dropdowns and checkboxes.

Requirements that could enhance the user experience and accessibility of multiple sclerosis patients are making sure that touch events are triggered after touch has stopped, so that users with poor control do not fire unwanted events (Ballantyne et al., 2018). If the multiple sclerosis patient uses external devices, such as keyboard, the application should be usable when using it, meaning that elements can be focused and focus can be moved between elements by the keyboard as well as navigated with it (Ballantyne et al., 2018).

2.2.3 Vision impairment

Vision impairment encompasses complete blindness, partial blindness, and color blindness (Siebra et al., 2017). Vision impairment can be result of illness or injury and may require spectacles or cause blurriness in the vision (BBC, 2022). Users with vision impairments have higher accessibility needs as they cannot rely on visual cues or tangible buttons.

The requirements for vision impairment users include constant feedback from the applications, like indicating which buttons are pressed, available interactions and ensuring applications ease of use with screen readers (Siebra et al., 2017). Other guidelines emphasize that the user interface should be clear, be descriptive, have enough contrast and be resizable (Ballantyne et al., 2018). Vision impairment encompasses complete blindness, partial blindness, and color blindness.

The requirements differ between completely blind individuals and those with limited vision. Limited vision users benefit adjustable settings, such as brightness, contrast, or colour adjustments (Ballantyne et al., 2018). These requirements aim to prevent problems with mobile applications, such as text is hard to read, accidental clicks, complex menus, and inconsistent navigation (Ruzic & Sanford, 2017). Limited vision users encompass individuals with colour blindness, who can otherwise see normally but colours are mixed up or they cannot differentiate different colours.

Despite the existence of various guidelines and WCAG requirements, they do not always address the needs of blind users, particularly with websites and data visualization. Furthermore, app developers sometimes prioritize more design over accessibility, especially with data visualization, because using different patterns can make the graphs more complicated. One solution to this problem is to provided alternative way to read and access data, as highlighted in study by Fan et al., (2023). Additionally, the study found out that majority of the popular sites overlooked accessibility when making data visualisations.

Social media applications are currently filled with photos and videos. A study conducted in 2016 examined the impact of images on blind users' experience on Twitter. It was noted that blind users had problems with tweets with images, because images did not have alt-tags with meaningful information, and thus tweet lost its meaning (Morris et al., 2016). The study found out that videos performed better in it since they usually had voice that gave enough information of video’s status.

Vision impairments are closely linked to multiple sclerosis and at times vision impairments can be first symptoms of MS. Common visual impairments with multiple sclerosis, are problems with colour vision, temporary blindness, double vision, difficulties to direct gaze and unvoluntary movement of eyes (Hoff et al., 2019). These impairments all require own solutions for accessibility, from accessible colours and patterns for colour blindness to correct structure of applications for screen readers.

Vision impairment guidelines that could specially help users with multiple sclerosis, especially in the context of multiple sclerosis applications are information should be provided different ways besides colours, meaning data visualization should have alternative ways to be accessed (Ballantyne et al., 2018). Ensuring enough contrast between text and background due to the varying levels of vision impairments experienced by patients with multiple sclerosis (Ballantyne et al., 2018).

2.2.4 Hearing impairment

Individuals with hearing impairments may exhibit either total hearing loss, referred to as deafness, or partial hearing impairments, characterized by variations in auditory capacity such as unilateral deafness or different degrees of hearing loss (Siebra et al., 2017). Tinnitus is also classified as a hearing impairment. Additionally, individuals who utilize hearing aids, despite their improved hearing abilities with the aids, are still considered to have a hearing impairment (BBC, 2022).

Although individuals with hearing impairments may have better accessibility for mobile applications due to possibility to see and use their hands, their inability to hear presents significant obstacles. Many of them rely on sign language as their primary mode of communication which can impact their reading skills (Yeratziotis et al., 2022). Accessibility requirements for hearing impairments have taken this into account. Siebra et al., (2017) list seven requirements for hearing impairment users, like providing captions or transcriptions in sign language, or using images beside text content and vibrations for applications' notifications. It could be simplified that videos and audio recordings, live or pre-recorded, should have captions or transcriptions (Ballantyne et al., 2018).

The social media has become one of the main ways to communicate and deaf people should be able to use their preferred language to discuss with others and for example interact with mobile application with sign language. (Yeratziotis et al., 2022). Videos have emerged as a popular means of communication in social media platforms, with TikTok leading the way with their short videos. Youtube has been around since 2005 and in 2015 Shiver & Wolfe conducted a study about web-based multimedia and deaf accessibility. Study found out that deaf participants preferred captions in videos, and that automatic captions had potential, but there were too many errors. Youtube users can add automatic captions to their videos, but these captions might not be accurate due machine learning algorithms (Google, 2023c). Instagram also offer a similar feature to add automatic captions based on video audio (Instagram, 2023).

In the context of multiple sclerosis (MS), some patients may experience hearing impairments related to the condition. Sudden sensorineural hearing loss (SSHL), rapid hearing loss, is one symptom that has been studied with MS patients. However it is a rare symptom of MS, and indivual can be recover from it, meaning that hearing loss is only temporary (Atula et al., 2016). Patients with temporary hearing loss also need accessibility, like subtitles on videos. These can be achieved by following guidelines of providing captions for prerecord videos and audios, as well as not using background music in applications (Ballantyne et al., 2018).

2.3 Previous uses of accessibility of mHealth

There have been studies of mHealth applications and how they affect managing diseases, how useful and effective they are but there have not been many studies about accessibility of these applications. (Zhou et al., 2020). One notable example where accessibility has been considered is mobile health system called iMHere. It was developed for improving user self-management skills, especially when user has dexterity problems (D. Yu et al., 2019).

The iMHere consist of several apps or modules, and there has been several studies how to make iMHere more accessible with good results (D. Yu et al., 2019; D. X. Yu et al., 2014; Zhou et al., 2020). The addition of accessibility features to the apps resulted in increased user satisfaction due to their ease of use and learnability (D. X. Yu et al., 2014).

While there have been studies on the accessibility and usability of mobile applications, most of them have focused on a wide range of application types, such as testing government applications (Antonio Mateus et al., 2021). One study examined the usability of applications specifically designed for Multiple Sclerosis patients. However, it was noted that these MS apps were not tested with the targeted population, and it remains unclear whether users could effectively use the apps prior to the study (Ruzic & Sanford, 2017).

3. Research Method

The research method employed in this study was benchmarking, which aimed to assess how accessibility concerns are addressed in the state of the practice of mobile health applications for individuals with MS. To achieve this, a selection of mobile health applications was tested using the Android Accessibility Scanner and Android TalkBack. The Android platform was chosen for its widespread usage and availability of these testing tools, making it a suitable platform for evaluating accessibility. This chapter provides detailed overview of the selected research method, including testing practices and measurements employed in the study.

3.1 Benchmarking

Benchmarking, traditionally used in software engineering for testing system performance, can be extended to cover other properties such as security, reliability, and in this case, accessibility. (Kounev et al., 2020). It can be defined as an evaluation tool for comparing systems and components properties such as dependability, security, and performance (Kistowski et al., 2015).

Benchmarks possess five key characteristics that serve as quality criteria: relevance, reproducibility, fairness, verifiability, and usability. Relevance ensures that the benchmark aligns closely with the user's interests and behaviors. Reproducibility verifies that consistent results are obtained with identical configurations. Fairness allows test configurations to compete based on their own achievements. Verifiability ensures accurate results, while usability minimizes roadblocks in running benchmarks in users' testing environments. (Kistowski et al., 2015)

Benchmarking encompasses three fundamental aspects: metrics, workload, and measurement methodology. Metrics are the values derived from measurements and serve as the basis for benchmark results. The workload comprises scenarios and conditions, such as the execution of specific programs. Measurement methodology encompasses the measurements themselves, the benchmarking process, and the production of results. (Kounev et al., 2020)

In addition to these core aspects, benchmarking involves other essential components. Motivating comparison provides a need or purpose behind the benchmarking study, which, in this case, is evaluating the accessibility of mobile health applications for individuals with MS. The task sample represents real-world usage scenarios, ensuring that the benchmark reflects actual practice. Performance measures, whether quantitative or qualitative, are used to evaluate the accessibility of applications, either through automated tools like the Accessibility Scanner or through human observation, in this case TalkBack reading. (Hasselbring, 2021)

Benchmarking was selected for this study due to its inherent nature of comparing and evaluating. Its clear structure offers an easy way to plan and execute research when assessing accessibility. Additionally, benchmarking provides a reliable means to measure accessibility and opens possibilities for future research.

By utilizing the benchmarking approach, this study aimed to provide insights into how accessibility concerns are addressed in mobile health applications for individuals with MS. The Android platform and the selected testing tools offered a representative task sample and user-friendly environment. However, it is important to acknowledge the limitations of this study, such as the small number of tested applications and the focus on a single platform. Future research could consider expanding the number of tested applications, including those on iOS, and incorporating manual testing or user testing to further explore the accessibility requirements and adaptability of health applications to the evolving needs of individuals with MS.

3.2 Choosing the test method

There are couple ways to test the accessibility of mobile applications. One is users testing, where user uses the applications and review makes notes on all the barriers that come up during the testing session. Other way is to test the application with different testing tools that check problems in screen or in the source code.

In this research the chosen method was testing tool approach. User testing has the potential for future use as testing tools may not identify all problems, such as those related to content or user experience, and can sometimes produce inaccurate results. The testing tools typically cover range of accessibility guidelines including WCAG and BBC Mobile guidelines.

Android has it is own testing tool, Accessibility Scanner which scans application's screen, and provides suggestions based on content labels, touch target size, clickable items and text and image contrast (Google, 2023a). It has default options for text and image contrast 4.5:1, which is same as in WCAG 2.1, and for touch target size, which default is 48 device-independent pixels (dp). 48 dp corresponds to 9 mm on device, and thus following recommendation in WCAG 2.1. and accessibility guidelines for mobile applications (Ballantyne et al., 2018). Suggestions are issues found on application. Accessibility Scanner identifies issues with elements and provides suggestions on how to improve or resolve the identified issues.

Initial research brought up different ways to test accessibility of mobile applications. Beside testing every screen manually, there are also options for automatic testing. Android has Espresso testing suite, which checks application's views during development (Google, 2023b). Two automated testing tools were found, which check same attributes as Android Accessibility Scanner and Espresso. First one was Mobile Accessibility Testing (MATE), which tests applications with random exploration and filling inputs with random data (Eler et al., 2018). Second one was xBot, which checks application's screen for accessibility issues (Chen et al., 2022).

This research was originally thought to be done with MATE, since there had been other research using MATE, and it was also surveyed together with couple of other testing tools (Silva et al., 2018). Eventually, due to time and device constraints, the decision was done against MATE, and Android Accessibility Scanner was chosen as main testing tool. It was also noticed that the result could be analysed better when researcher went through application manually, because of knowledge and experience which accessibility problem were vital, and potential threats to usability.

TalkBack was selected to complement Accessibility Scanner and verify its suggestions. It is a screen reader provided by Google and included in Android devices. It gives the user possibility to use the device eyes-free. It can be turned on from the device's accessibility settings. It reads out loud applications content and adds special gestures for easier application usage, (Tomlinson et al., 2016). TalkBack gestures and other accessibility features can be integrated to applications for better user experience, but it works generally in every app (Jain et al., 2021).

TalkBack works the most effectively when content has clear structure, and all buttons have labels on them. If buttons do not have labels, TalkBack cannot read them properly and user does not know what buttons do. (Jain et al., 2021). Buttons with labels is one of the WCAG requirement.

3.3 Testing plan

The testing process employed two key tools: Android Accessibility Scanner and TalkBack, both specifically designed for the Android platform. The decision to use these tools was based on their proven effectiveness in evaluating the accessibility concerns of mobile applications.

Android Accessibility Scanner was chosen for its ability to provide detailed insights into accessibility issues. Selected applications were obtained from Google Play Store with following criteria:

- Application must be available in Finland.
- Application's language must be English.
- Application must be targeted towards Multiple sclerosis.
- Application's purpose is either to educate or help managing user's health.
- Application was open to everyone to use without needing to register with phone number.

The criteria were chosen to ensure that the research could be easily maintained and reproduced without requiring the inclusion of personal details for applications and without any costs associated with accessing relevant features.

To conduct the evaluation, each selected application underwent a systematic testing procedure. First, the applications were subjected to the scrutiny of the Android Accessibility Scanner. During this phase, the application's screens were recorded while the tool was active, capturing the user's typical interactions and navigation within different screens. A sequential approach was followed, systematically pressing buttons and links from top to bottom and left to right. The recording results was then saved for further analysis.

Following the Accessibility Scanner test, the applications were subjected to evaluation using TalkBack, a screen reader integrated into the Android platform. The screens were assessed in the same order as during the Accessibility Scanner test. TalkBack's integrated gestures were utilized, with a primary focus on examining the screen's elements' reading order and verifying the accuracy of label suggestions provided by the Accessibility Scanner.

It is important to note that the chosen testing tools, Android Accessibility Scanner and TalkBack, have distinct advantages and limitations. Accessibility Scanner offers comprehensive insights into potential accessibility issues, providing suggestions for improvement. TalkBack, on the other hand, allows for a more user-oriented assessment by simulating the experience of individuals with visual impairments. However, it is important to consider that these tools may not address all aspects of accessibility, such as cognitive impairment requirements, which may require additional manual testing or user feedback.

By adopting this testing methodology, incorporating the strengths of Android Accessibility Scanner and TalkBack, and considering the limitations, a comprehensive evaluation of the selected applications' accessibility was conducted, providing valuable insights into the state of accessibility concerns in mobile health apps for Multiple Sclerosis.

3.4 Measurements

This study employed a range of measurements derived from the data collected during the testing process. The Android Accessibility Scanner generated reports containing suggestions for improving the accessibility of the tested applications. These suggestions were categorized into specific measurements, aligned with established accessibility guidelines. The measurements derived from the Accessibility Scanner reports were as follows:

- Item label: This measurement focused on assessing the presence of labels associated with buttons, links, and other interactive elements within the applications.
- Touch target: The touch target measurement evaluated the size and spacing of interactive elements to ensure they met the minimum requirement for ease of use, particularly for individuals with motor impairments.
- Text contrast: This measurement examined the contrast ratio between the text and its background to ensure it met the minimum standards for legibility, catering to individuals with vision impairments.
- Image contrast: This measurement specifically assessed the contrast ratio of images within the applications, considering their visibility and legibility for users with vision impairments.
- Item description: This measurement focused on providing information of elements with same descriptions, ensuring that individuals with vision impairments could distinguish different elements when using screen readers.

In addition to the Accessibility Scanner, TalkBack observations during the testing sessions provided further measurements to assess the applications' accessibility. These observations were categorized into the following measurements, based on the established accessibility guidelines for mobile applications (Ballantyne et al., 2018):

- Nameless buttons and inputs: This measurement identified instances where buttons and input fields lacked descriptive labels, making it challenging for individuals using TalkBack to understand their purpose.
- Reading order and state of application: The reading order and status measurement assessed whether the elements within the application were presented in a logical sequence and provided accurate status updates, ensuring a seamless user experience with TalkBack.
- Reading images and icons: This measurement examined TalkBack's ability to accurately read and describe images and icons used within the applications, ensuring that individuals with visual impairments could comprehend the visual content.

- Data visualization: The data visualization measurement specifically focused on assessing TalkBack's ability to convey information from data visualizations, ensuring that individuals with visual impairments had access to alternative formats or descriptions.

The measurements were carefully chosen to ensure that the requirements and guidelines identified in prior research, particularly those relevant to multiple sclerosis patients, would be thoroughly tested across the selected applications. Accessibility Scanner measurements provided by the application itself, but they were aligned with the accessibility guidelines and requirements.

To facilitate analysis, the metrics obtained from these measurements were compiled into an Excel file, with each application having a dedicated sheet. This organization allowed for easy comparison and comprehensive analysis of the data. Summary sheets were also created for each testing tool, enabling a holistic assessment of the overall accessibility landscape.

By utilizing these measurements and consolidating the data in a structured manner, a thorough analysis of the accessibility concerns in the tested applications was conducted, providing valuable insights into their adherence to established accessibility guidelines.

4. Findings

10 applications were selected for accessibility testing. All applications were available on Google Play and targeted MS patients. The testing was conducted using Accessibility Scanner and TalkBack. The results of testing sessions are presented in this chapter.

Eight of applications were disease management applications, designed to help users. Users interacted with these applications by inputting various data and engaging in chat conversations with other users. Two were disease and treatment information applications, without any user interactions involving data input.

4.1 Accessibility Scanner

Accessibility Scanners provided suggestions for five measurements: item label, touch target, text contrast, image contrast and item description. The number of screens, suggestions, and division by measurement can be found in Table 1. Table 2 presents percentages per category per application as the number of suggestions and screens differed, making direct comparisons challenging and in Figure 1 the same percentages can be seen in bar chart. In total, 560 screens were scanned, resulting in 6670 suggestions provided by Accessibility Scanner.

Table 1 Accessibility Scanner results

Name	**Screens**	**Suggestions**	**Item Label**	**Touch Target**	**Text contrast**	**Image contrast**	**Item description**	**Average suggestions per screen**
SelfInsight	44	495	86	181	169	38	21	11,25
BelongMS	69	1916	514	359	433	594	16	27,77
MS tim	80	998	325	511	114	4	44	12,48
icompanion	134	412	134	250	13	15	0	3,07
My MS-UK	47	677	141	265	162	103	6	14,40
Ms Notes Journal	18	199	14	32	123	16	14	11,06
The Msing Link	45	689	0	371	318	0	0	15,31
Symptom & Mood Tracker	70	1014	201	379	388	14	32	14,49
MyTherapy	36	175	3	14	138	6	14	4,86
Multiple Sclerosis Symptoms	17	95	2	72	21	0	0	5,59
Min	17	95	0	14	13	0	0	3,07
Max	134	1916	514	511	433	594	44	27,77
Average	56	667	142	243,4	187,9	79	14,7	12,03
Median	46	586	110	257,5	150	14,5	14	11,86

Table 2 Accessibility Scanner percent results

Name	Item Label %	Touch Target %	Text contrast %	Image contrast %	Item description %
SelfInsight	17,37 %	36,57 %	34,14 %	7,68 %	4,24 %
BelongMS	26,83 %	18,74 %	22,60 %	31,00 %	0,84 %
MS tim	32,57 %	51,20 %	11,42 %	0,40 %	4,41 %
icompanion	32,52 %	60,68 %	3,16 %	3,64 %	0,00 %
My MS-UK	20,83 %	39,14 %	23,93 %	15,21 %	0,89 %
Ms Notes Journal	7,04 %	16,08 %	61,81 %	8,04 %	7,04 %
The Msing Link	0,00 %	53,85 %	46,15 %	0,00 %	0,00 %
Symptom & Mood Tracker	19,82 %	37,38 %	38,26 %	1,38 %	3,16 %
MyTherapy	1,71 %	8,00 %	78,86 %	3,43 %	8,00 %
Multiple Sclerosis Symptoms	2,11 %	75,79 %	22,11 %	0,00 %	0,00 %
Min	0,00 %	8,00 %	3,16 %	0,00 %	0,00 %
Max	32,57 %	75,79 %	78,86 %	31,00 %	8,00 %
Average	16,08 %	39,74 %	34,24 %	7,08 %	2,86 %
Median	18,60 %	38,26 %	29,04 %	3,53 %	2,02 %

Table 3 Accessibility Scanner refined results

Name	Screens	Suggestions	Item Label	Touch Target	Text contrast	Image contrast	Item description
SelfInsight	28	70	12	16	32	8	2
BelongMS	37	136	37	41	25	26	7
MS tim	48	50	14	20	7	1	8
icompanion	64	15	1	8	2	4	0
My MS-UK	39	26	2	9	11	0	4
Ms Notes Journal	12	17	2	1	11	2	1
The Msing Link	22	31	0	20	11	0	0
Symptom & Mood Tracker	49	129	27	43	51	3	5
MyTherapy	26	20	2	3	10	3	2
Multiple Sclerosis Symptoms	11	11	0	10	1	0	0

The Accessibility Scanner results were further refined to provide a clearer understanding of the situation for each individual element. Duplicate screens were removed. Only suggestions pertaining to individual elements, such as navigation items, were considered, and suggestions with similar elements were grouped together. Furthermore, for text and image contrast, only one suggestion with the same colours and contrast ratio was kept.

This gave a good view of elements, what cause troubles with accessibility. The refined results can be found in Table 3. Figure 2 is a bar chart containing refined results in percentages.

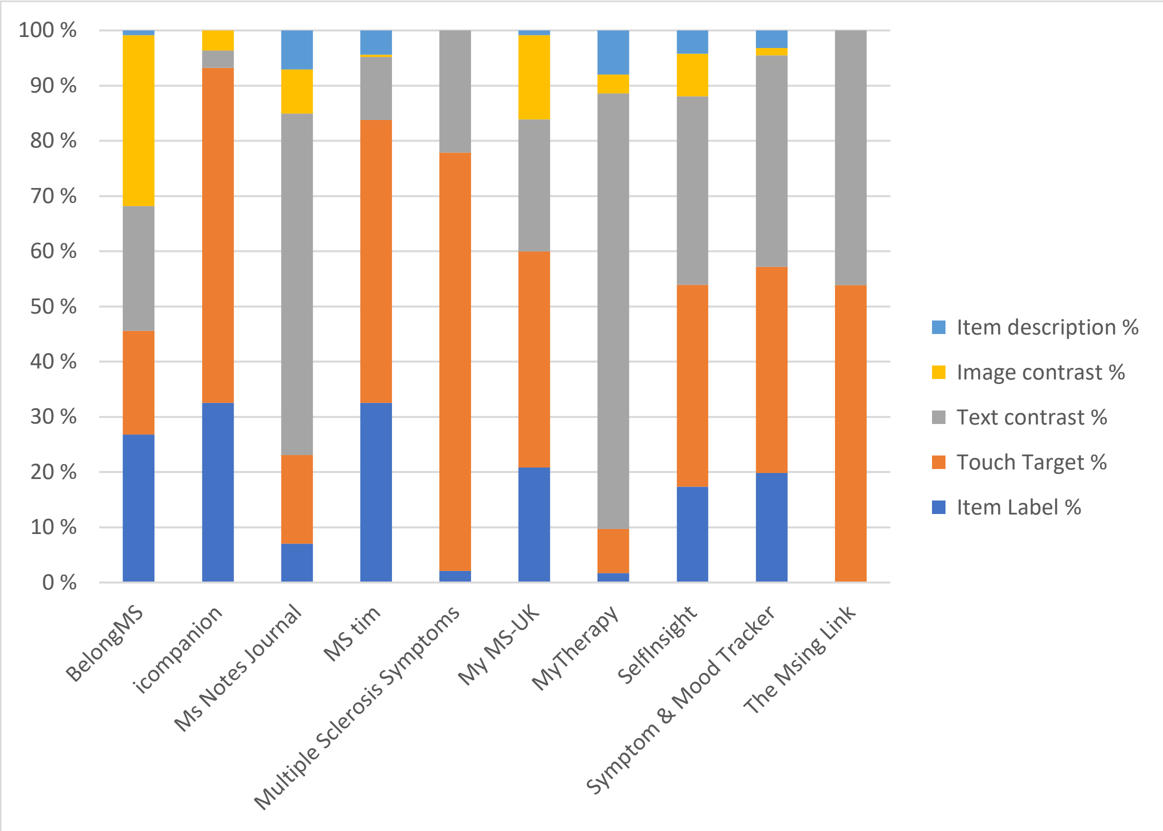

Figure 1 Bar chart of original percentages

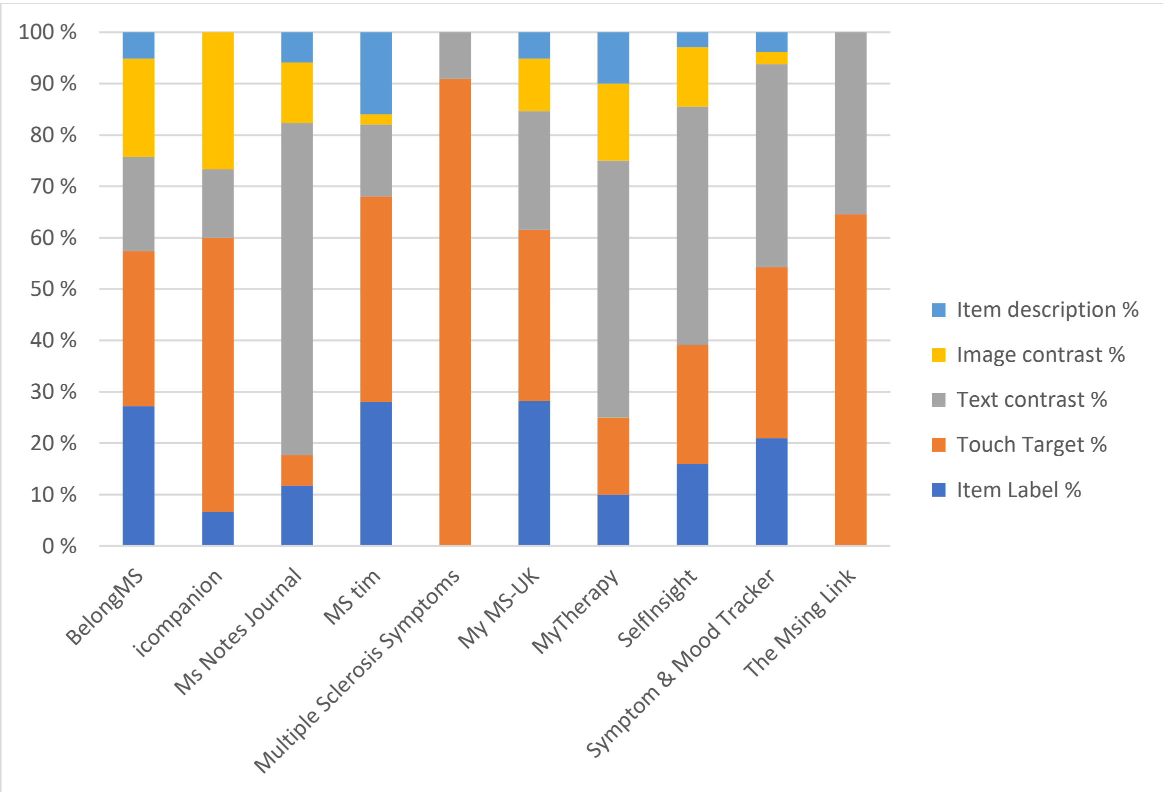

Figure 2 Bar char of refined results as percentages

4.1.1 Item labels

Accessibility Scanner provided 1420 suggestions for item labels, indicating that certain items lacked labels, potentially affecting their accessibility for screen readers. Many of the item label findings were also identified during the TalkBack testing.

The suggestions highlighted various types of items, with buttons being the most common type lacking labels, followed by inputs and finally images. Buttons without labels were typically navigation, menu, and calendar buttons. Inputs without labels were primarily scale inputs used for logging symptoms. In Figure 4, all the highlighted pictures and buttons do not have item labels.

In Figure 3 items without labels are back and close buttons, and inputs for adding symptoms. Additionally, the symptoms inputs, which are also devoid of labels, lack explanatory text that would normally indicate the purpose of these input fields. Consequently, this absence of contextual information renders the symptoms inputs visible to all users without providing a clear indication of what each input represents.

In the original data (Table 1), only one application did not receive suggestions for item labels. However, in the refined data (Table 3), two applications had no suggestion for item labels. In Table 2, can be seen that item label suggestions were average 16,08 % of all suggestions made by Accessibility Scanner with a maximum percentage of 32,57%.

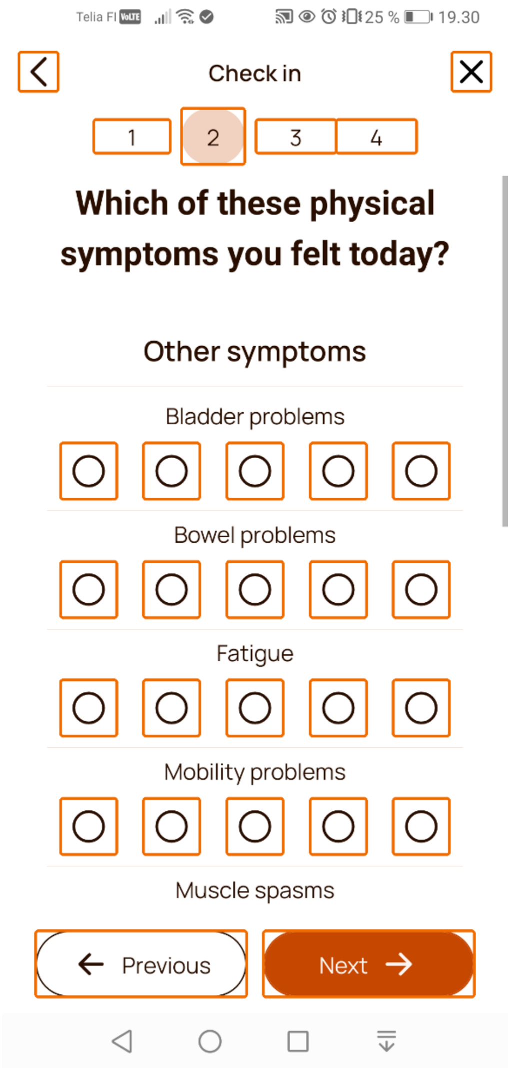

Figure 3 MS time screen caption

4.1.2 Touch Target

Touch target suggestions were the most common suggestions provided by Accessibility Scanner. It was suggested 2434 times, with average percent being 39,74 %. These suggestions align with the recommendation that items should have a minimum width and height of 9mm or 48dp, accounting for padding around the items. Items were highlighted if their width or height fell below 48dp.

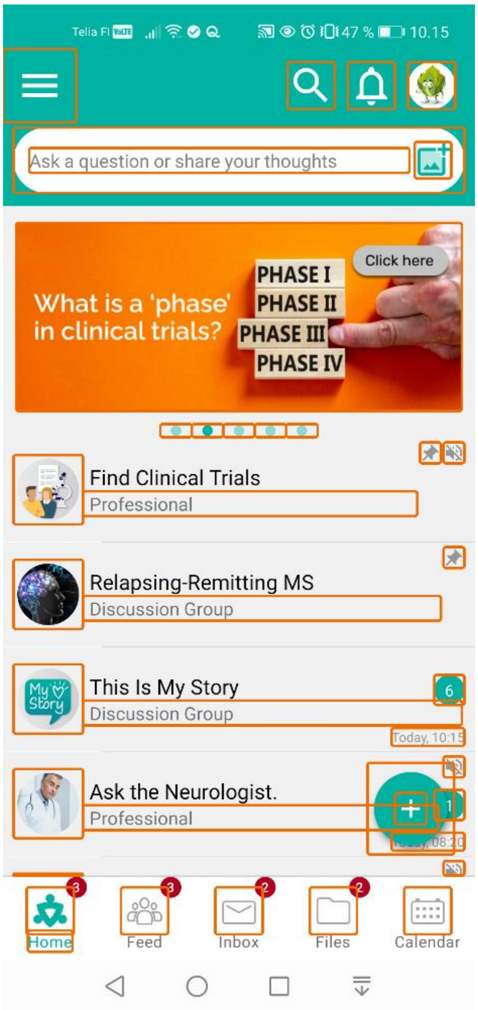

Figure 4 BelongMS screen caption

The size of touch target suggestions was also evaluated. The evaluations were based on refined data. Suggestions with a value above 38dp were considered tolerable, while those between 37dp and 25dp were classified as poor, and values under 25dp were deemed severe. The evaluation results can be found in Table 4 and bar chart of results can be found in Figure 5.

Table 4 Touch target results

Touch target	Tolerable	Poor	Severe
SelfInsight	0	4	5
BelongTail	9	20	12
MS tim	6	9	5
icompanion	3	2	3
My MS-UK	3	7	3
Ms Notes Journal	1	0	0
The Msing Link	5	10	5
Symptom & Mood Tracker	11	21	11
MyTherapy	1	2	0
Multiple Sclerosis Symptoms	0	0	0

For suggestions with both height and width, the evaluation was based on the lower value. A tolerable evaluation indicated that suggestion could be easily fixed, poor evaluations meant that fixing suggestions would require more effort, and severe evaluations indicated that suggestion would require significant work, with the fix being noticeable to user.

The touch target suggestions were applicable to buttons, inputs, and links. Inputs typically met the suggested width requirements but fell short in height requirement. Button sizes significantly, often with both height and width below the suggested values. In Figure 4 the highlighted buttons in upper right corner have suggestion for touch target due to their size being only 36dp x 36dp.

In Figure 3 all the highlighted items were under 48dp x 48dp. However, a slight deviation from this requirement is observed in the buttons located at the bottom of the screen, as their height measures 47dp, coming close to the specified dimension. On the other hand, the inputs designated for adding symptoms have a uniform size of 40dp x 40dp, falling short of the prescribed 48dp x 48dp guideline.

All tested applications received touch target suggestions, and this aspect was particularly emphasized in situations where one or more categories had no suggestions. The highest percentage of touch target suggestions among all suggestions for an application was 75.79%, while the lowest percentage was 8%. Within the refined data Table 3, one application did not receive any touch target suggestions.

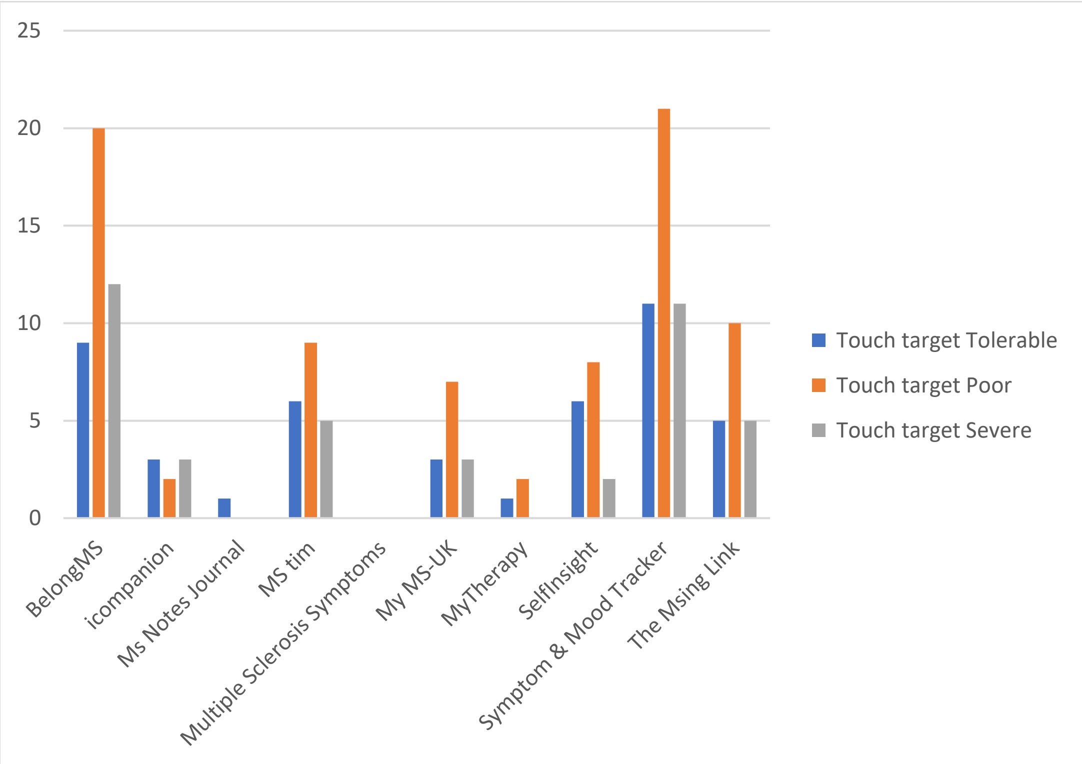

Figure 5 Touch target bar chart

4.1.3 Text contrast

Text contrast measurement provided suggestions, when the contrast between texts and background fell below 4.5:1. While a contrast ratio of 3:1 may be acceptable for larger text, the Accessibility Scanner specifically tested contrast using a threshold of 4.5:1. To facilitate evaluation, the refined data values were categorised: tolerable, poor, and severe.

Ratio for tolerable was over 3.8, poor 3.7-2.5 and severe under 2.5. Table 5 presents the corresponding results. Figure 6 contains the same information as Table 5 but in bar chart format for alternative view of the data.

The text contrast measurement applied to various elements, including plain text, buttons with icons and text, and inputs. Issues were identified with grey text on white backgrounds, disabled buttons, and color schemes that lacked adequate contrast with white color.

All the applications received suggestions regarding text contrast, making it the second most frequently encountered suggestion in terms of both frequency and average percentages. In Figure 7 can be seen text contrast suggestions with disabled number buttons. Their contrast ratios were 1.28:1 which falls below the default value, even with larger fonts. Within same figure, under the 'Sleep Interruption Level' heading, the number 1 is highlighted also for text contrast. Its contrast ratio is 3,86:1, which could be deemed acceptable given its larger font size.

Table 5 Text contrast results

Text contrast	Tolerable	Poor	Severe
SelfInsight	6	16	10
BelongTail	5	14	6
MS tim	0	3	4
icompanion	0	1	1
My MS-UK	3	5	1
Ms Notes Journal	1	4	6
The Msing Link	1	3	7
Symptom & Mood Tracker	1	11	39
MyTherapy	3	5	2
Multiple Sclerosis Symptoms	0	0	0

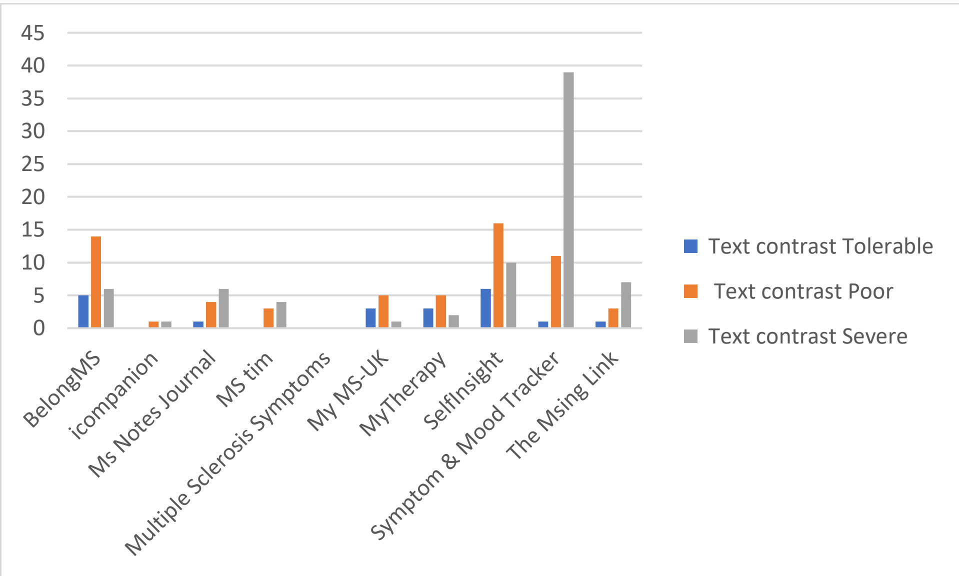

Figure 6 Bar chart of text contrast numbers of suggestions

Figure 7 SelfInsight Accessibility Scanner screen caption

4.1.4 Image contrast

Image contrast measurement assessed the contrast between images and the background, with minimum required contrast ratio of 3:1 to ensure adequate visibility of the images. This measurement was specifically applied to images. Two applications did not receive any suggestions regarding image contrast. The remaining applications had scale from 0.4 % to 31 %, and average was 7.8 %. Image contrasts values were evaluated following the same approach as text contrast utilizing categories tolerable, poor, and severe, based on the refined data. Contrast evaluations followed the criteria of tolerable (ratio above 3.8), poor (ratio between 3.7 and 2.5), and severe (ratio below 2.5). The results can be accessed in Table 6 and in Figure 8.

Table 6 Image contrast results

Image contrast	Tolerable	Poor	Severe
SelfInsight	0	4	5
BelongTail	0	14	12
MS tim	0	0	1
icompanion	0	0	4
My MS-UK	0	2	2
Ms Notes Journal	0	1	1
The Msing Link	0	0	0
Symptom & Mood Tracker	0	1	2
MyTherapy	0	0	3
Multiple Sclerosis Symptoms	0	0	0

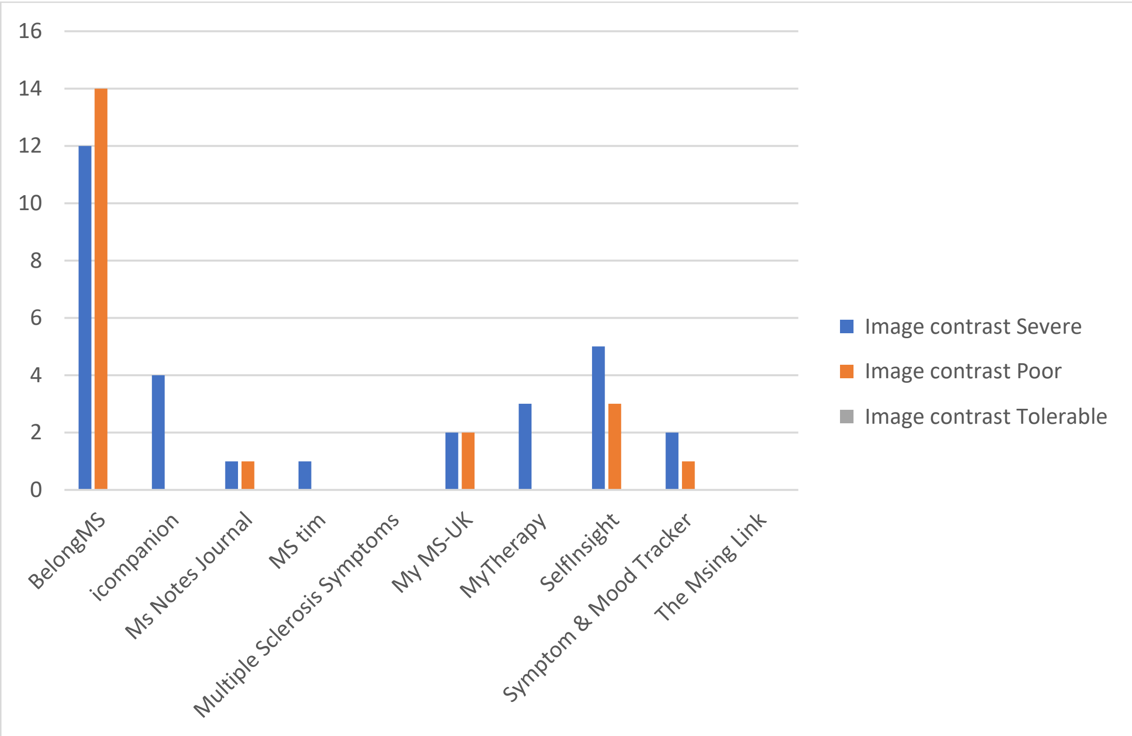

Figure 8 Bar chart of image contrast number of suggestions

Figure 4 showcased the icons highlighted at the bottom of the screen, exhibiting a contrast ratio of 2.52:1. In Figure 10 food images had contrast ratios varying from 1.78:1 to 2.74:1, illustrating examples of images that fail to meet accessibility requirements. It is recommended to manually assess highlighted images, if necessary, as suggested by the Accessibility Scanner.

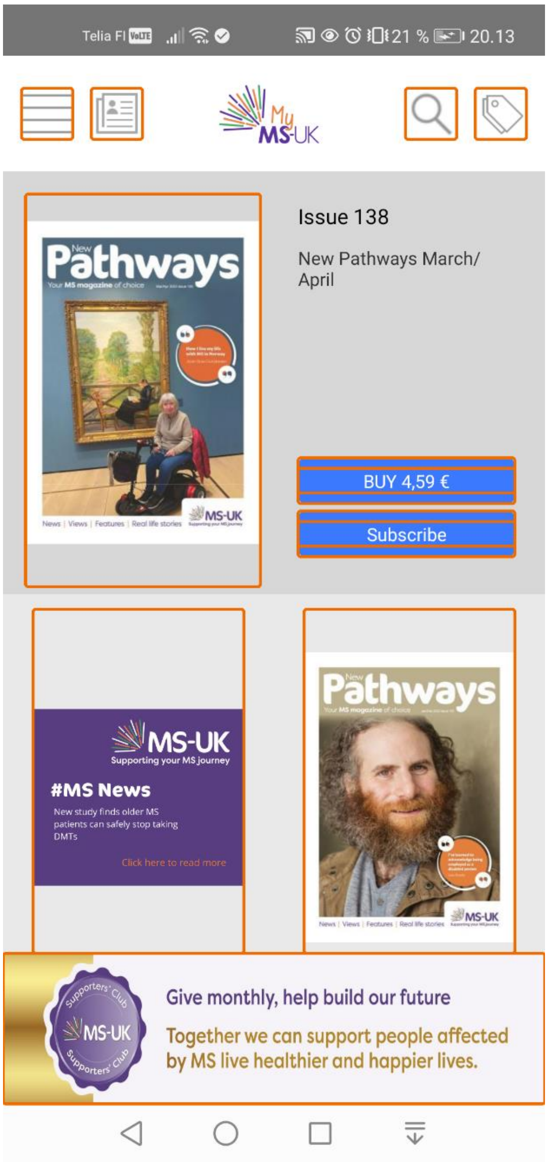

Figure 9 My MS-UK screen caption

In Figure 9 the icons in the top bar are highlighted with poor image contrast ratio. The contrast ratio values of these icons range from 2.85:1 to 2.68:1, which fall below the recommended threshold. Considering the relatively large size of the buttons, it is essential that the contrast ratio exceeds 3 or higher to ensure optimal visibility and legibility for users.

4.1.5 Item description

The final measurement focused on item descriptions, which received the fewest number of suggestions. Three applications did not receive any suggestions regarding item descriptions. Item description could be applied to various elements, such as text, buttons, images, and inputs.

The issue with item description arises when multiple items share the same description. For instance, in Figure 7, the suggestion pertains to the first number 2 button, as there are two additional number 2 buttons on the screen that share the same description. Descriptions should differ for each individual element for better user experience.

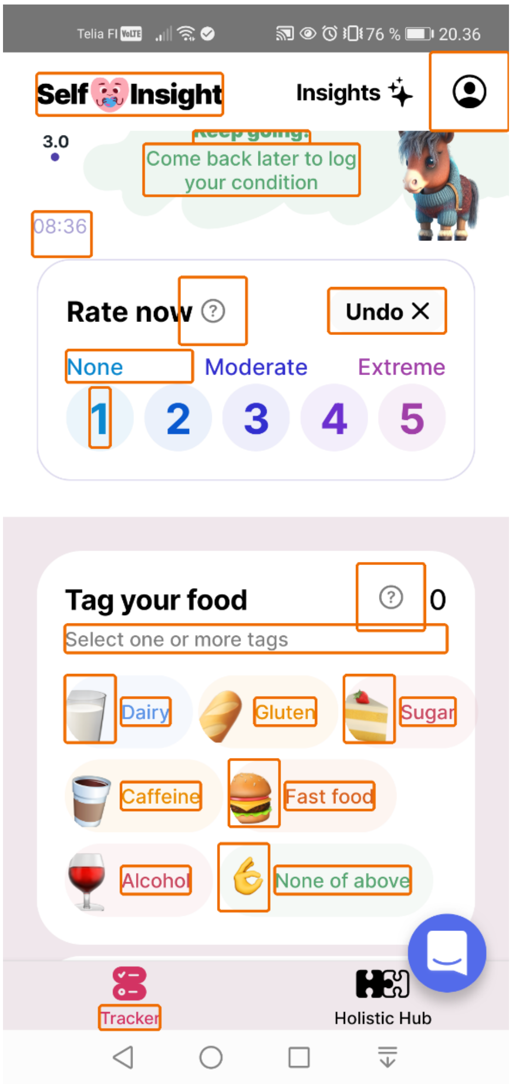

Figure 10 SelfInsight Accessibility Scanner capture

4.1.6 Summary of Accessibility Scanner findings

Accessibility Scanner identified numerous suggestions from the applications. Every application received suggestions in at least three measurements. Some of the suggestions were corroborated using TalkBack. Table 1 illustrates that the average number of suggestions per screen was 12.03. Touch target measurement yielded the highest numer of suggestions and a comparatively higher average percentage. Item description measurement received the fewest suggestions and had the lowest average percentage.

Applications characterized by a simple structure and absence of photos exhibited better performance compared to those with discussion forums and a high number of inputs. The measurements conducted by the Accessibility Scanner were clear and effectively emphasized crucial aspects of the applications. The suggestions served as metrics for the analysis.

4.2 TalkBack

The TalkBack findings were categorized into four measurements to identify the encountered problems within the application. The findings were consolidated into one table (Table 7) for subsequent analysis. The four selected measurements were nameless buttons and inputs, reading order and status of application, reading images and icons, data visualization.

Table 7 TalkBack results

Name	Nameless buttons and inputs	Reading order and state of application	Reading images and icons	Data visualization
SelfInsight	Missed items	Neutral	No	Alternative way to present data
BelongMS	Nameless items	Bad	No	-
MS tim	Nameless items	Neutral	No	Labels only
icompanion	Nameless items	Neutral	No	No
My MS-UK	Missed items	Neutral	No	-
Ms Notes Journal	No	Good	No	-
The Msing Link	No	Good	Yes	-
Symptom & Mood Tracker	Missed items	Neutral	No	Premium only for charts
MyTherapy	No	Good	No	Alternative way to present data
Multiple Sclerosis Symptoms	No	Good	No	-

4.2.1 Nameless buttons and inputs

The measurement of nameless buttons and links was established by observing how TalkBack read various buttons and links. Table 7 presents three metrics under the Nameless buttons and links column: missed items, nameless items, and no. Nameless items refer to buttons or links in the tested application that lack names yet are still recognized as items by the screen reader. Missed items were those that the screen reader failed to recognize altogether. No indicated that all buttons and links had names and their purpose was evident. Out of the ten tested applications, seven had one or more buttons or inputs without names, yet TalkBack was able to recognize them as buttons and inputs.

Nameless buttons encompassed navigation buttons, calendar navigation buttons, and settings or more information buttons. These buttons posed the most frequent issue identified by TalkBack, aligning with the Item label suggestions from the Accessibility Scanner. Inputs without names or labels were less prevalent compared to nameless buttons. Inputs without name were scale inputs employed to record varying levels of symptoms, such as fatigue and brain fog.

Among the tested applications, three applications had buttons or inputs that were entirely unreadable by TalkBack. The missed buttons consisted of navigation buttons and calendar buttons, whereas the missed input buttons were scale inputs related to various symptoms. Pressing these buttons did not trigger TalkBack to tell what kind of element was pressed. In total, only three applications had successfully named all buttons and provided labels for inputs. Example of inputs that were not read at all can be found in on the top of Figure 11.

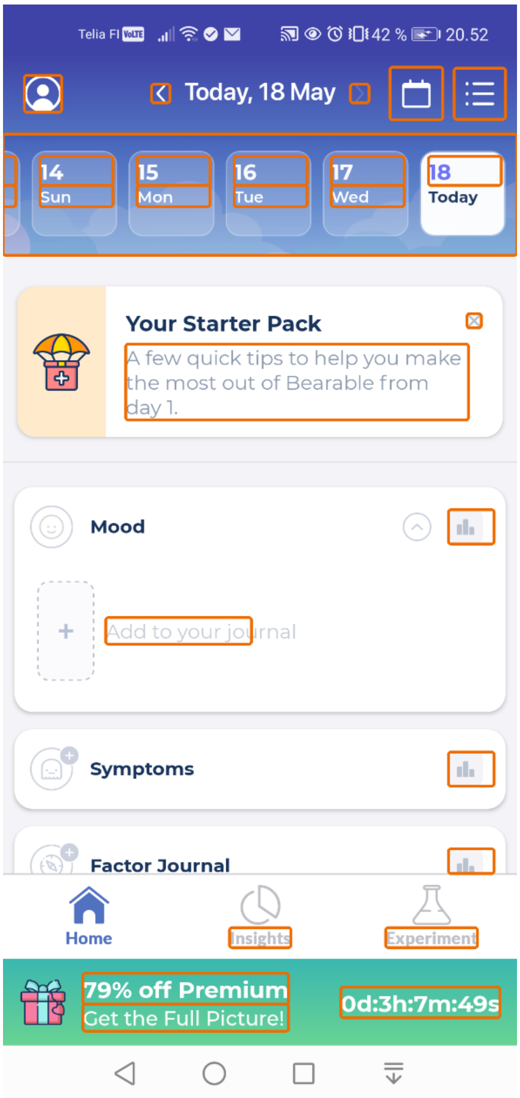

Figure 11 Mood and symptoms screen capture

4.2.2 Reading order and state of application

The reading order and state of application measurement focused on how TalkBack navigated through the application's screen and the sequence in which different elements were read. Table 7, under the Reading order and situation column, presents three metric options: good, neutral, and bad. An application was deemed good, if the reading order of the elements was understandable, and it was clearly indicated if tab or input was selected or active. A neutral assessment was assigned when the reading order of the elements was understandable or when there was clear indication of chosen or active tab or input. Applications assessed as bad indicated the presence of unreadable order of elements and a lack of information from TalkBack regarding active tabs or inputs.

Out of the tested applications, only three received a good assessment. Six applications received a neutral assessment, meaning that there were some issues with either reading order or status of buttons, inputs, and tabs. One application received bad assessment, meaning that there were problems with both reading order and status of buttons, inputs, and tabs.

Figure 12 shows an example of an application and a screen, where the state of the tabs was unclear when reading the screen with TalkBack. Even though 'Mood'-tab is highlighted and active tab, it was obvious to the screen reader and therefore not read at all.

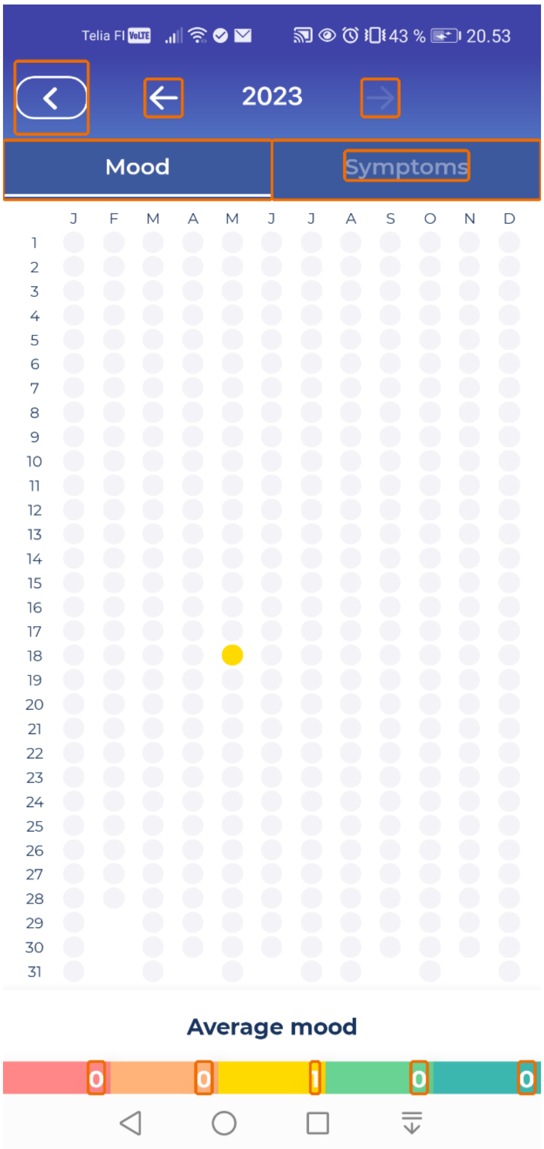

Figure 12 Symptoms and mood tracker screen capture

Two out of three applications with assessment good, were simple applications without pictures, and the one bad application was complex application with discussion forums. The most frequently encountered issue was TalkBack reading elements in the wrong order. The second most prevalent issue was TalkBack's inability to identify if an input or tab was active or selected.

4.2.3 Reading images and icons

The reading images and icons measurement was determined based on number of images described by TalkBack. If an application contained image, that TalkBack could not read it, it was assessed as “No”. If TalkBack could read the images, it was assessed as “Yes”. These assessments are listed in the Table 7. The reading of images and icons is based on the requirement that images should have alternative descriptions for screen readers.

Out of all the applications, only one had alternative descriptions for images, which TalkBack could read. In the remaining applications TalkBack skipped images entirely without indicating their presence. The reading images and icons measurement exhibited a correlation with the Item label suggestions from the Accessibility Scanner.

Figure 13 is an example of an application where the images could be read with TalkBack and Figure 14 shows an example of screen where the image could not be read. Image in Figure 14 was used as reward image for completing daily task, so description of the image should be there for screen reader.

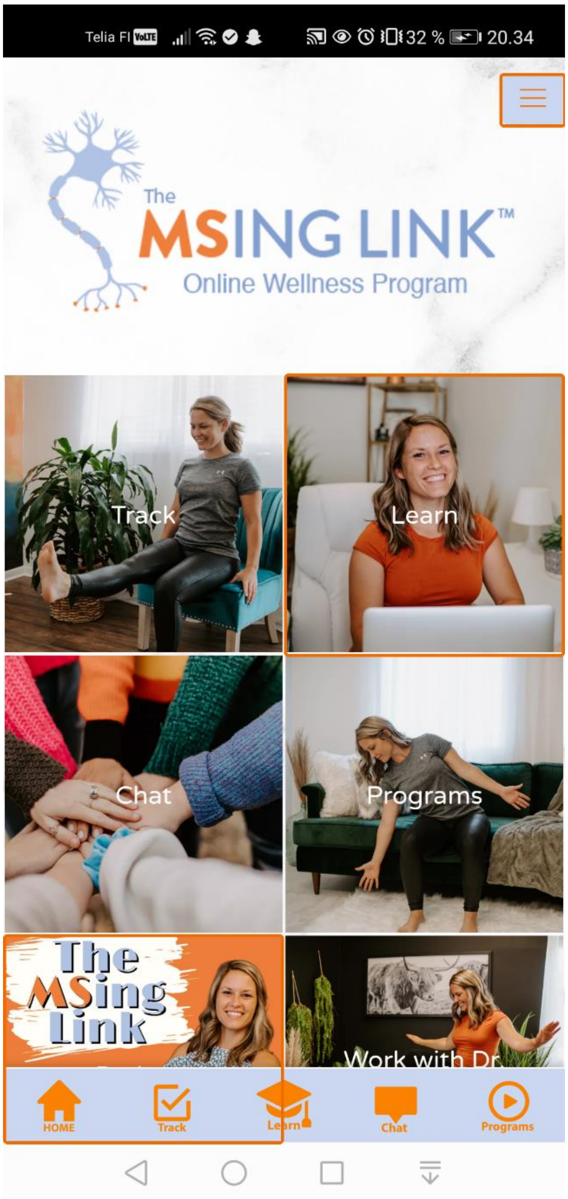

Figure 13 The MSing Link screen capture

Figure 14 My therapy screen capture

4.2.4 Data visualization

The data visualization measurement was not applicable to all applications, but it was chosen as a measurement because some of the applications collected users' inputs and presented them with data visualizations. It could have been combined with the Reading images and icons measurement, but due to the importance of data visualization in conveying information to the user, a separate measurement was created.

Out of the five applications with data visualization elements, only four were assessed, as one application's data visualization was a premium feature. TalkBack could not read the data visualizations in any of the four assessed applications. One application read the headings of the data visualizations, one also read the labels, while the others skipped the data visualizations entirely. The results were consistent with the findings from the reading images and icons category, as well as the suggestions provided by the Accessibility Scanner regarding item labels.

Out of the four assessed applications, two provided alternatives for data visualization. The alternative was in both cases list of data and data points. The alternatives consisted of a list of data and data points, which were readable with TalkBack and easily understood. Other two applications did not provide alternative view for data visualization.

In Figure 15 is an example of a data visualization which was not read at all. The screen reader just skipped it and did not read labels. The application did not provide alternative way to read or view inputted data.

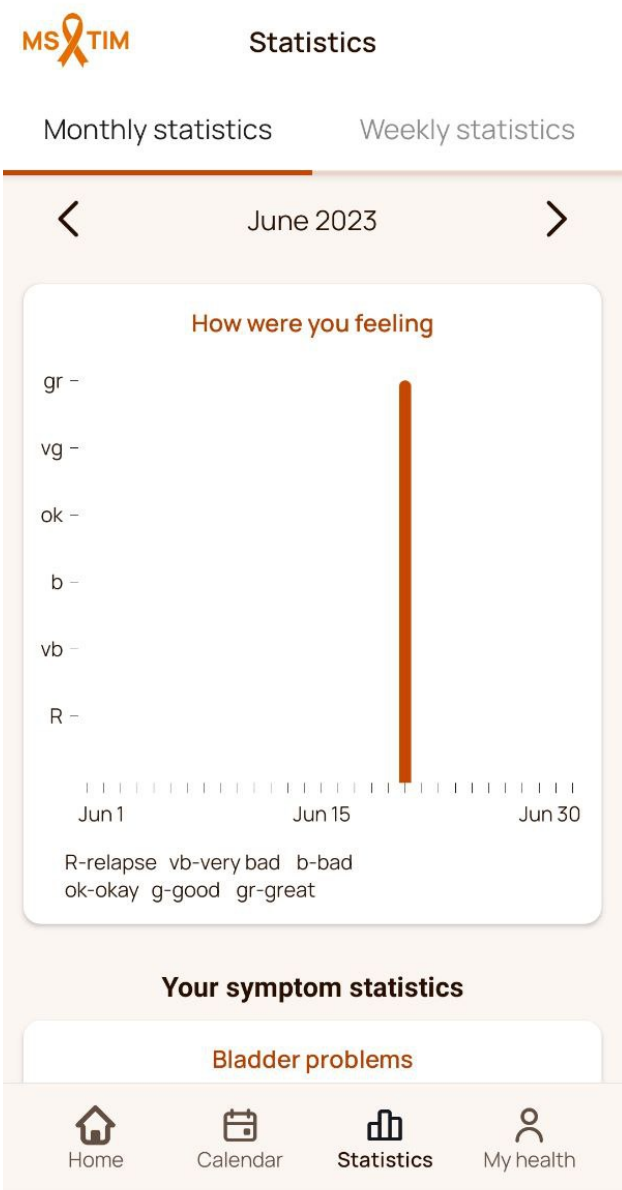

Figure 15 MS tim screen capture

4.2.5 Other findings with TalkBack

During the testing sessions, several observations were made outside of the designated measurements. One notable observation with TalkBack was that the application's functionality and TalkBack's gestures did not always align seamlessly. In some cases, using TalkBack's gestures resulted in the application navigating to incorrect locations due to misinterpretation of the gestures.

Another issue encountered was related to pop-up windows within the applications. Firstly, some pop-up windows were not accessible to TalkBack, meaning that their contents could not be read aloud. Secondly, in one instance, a pop-up window appeared last in the reading order when it should have been presented first. Additionally, there was an incident where a pop-up window could not be closed because TalkBack gestures and clicks only interacted with elements that were beneath the pop-up.

Figure 16 shows an example of problems with pop-up which could also be recorded with Accessibility Scanner, which gave suggestions below pop-up. The figure is from the application where pop-ups were read last.

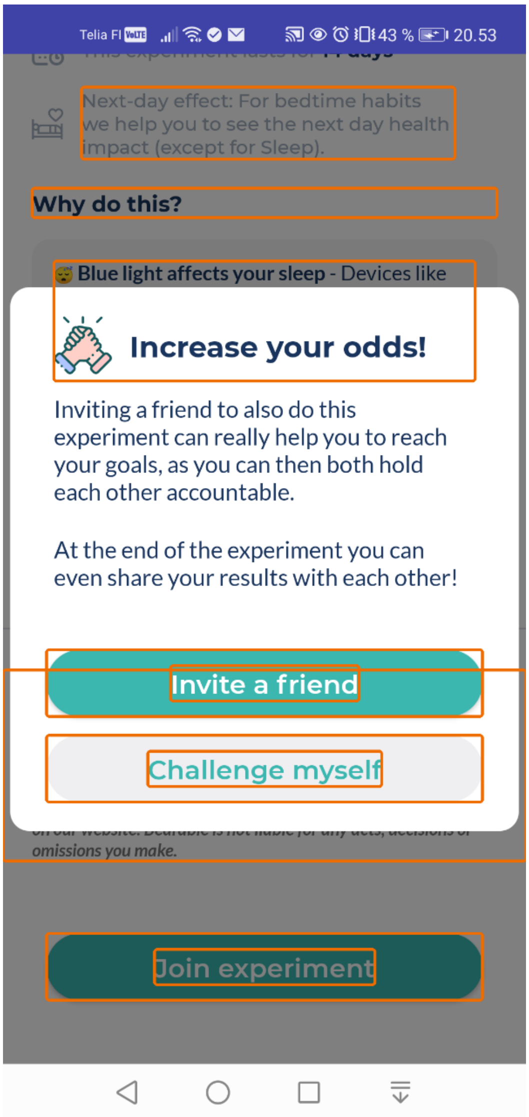

Figure 16 Symptoms and mood tracker screen capture

These observations highlight potential challenges and areas for improvement in the interaction between TalkBack and applications, particularly concerning gesture interpretation and the accessibility of pop-up windows.

4.2.6 Summary of TalkBack findings

TalkBack testing proved to be valuable in identifying issues that could not be assessed solely with Accessibility Scanner. It not only helped validate the item label suggestions provided by Accessibility Scanner but also revealed additional accessibility problems. Among the most common issues were the absence of alternative descriptions for images and the presence of nameless buttons and inputs.

It is important to note that none of the assessed applications achieved a perfect score in terms of accessibility. Every application had at least one problem in at least one measurement category. Interestingly, the type of application, whether focused on disease management or disease and treatment information, did not appear to have a significant impact on the overall results.

These findings emphasize the importance of conducting comprehensive accessibility testing using multiple approaches, such as combining automated scanning tools like Accessibility Scanner with manual testing using tools like TalkBack. This holistic approach helps uncover a wider range of accessibility issues and ensures a more inclusive user experience for all individuals.

5. Discussion

The research question for this study focused on examining how accessibility concerns are addressed in the current state of mobile health apps for Multiple Sclerosis (MS). The findings presented in this chapter are then compared to the requirements and aspects identified in prior research.

By evaluating the accessibility of the tested mobile health apps for MS using Accessibility Scanner and TalkBack, the study aimed to shed light on the existing practices and identify any gaps or areas for improvement in addressing accessibility concerns. The findings from the measurements and observations were then analysed and compared to established requirements and previous research in the field.

This comparative analysis allows for a comprehensive understanding of the current state of accessibility in mobile health apps for MS and provides insights into the extent to which these apps align with established accessibility guidelines and recommendations. It helps identify areas where improvements can be made to enhance the accessibility and usability of these apps for individuals with MS, thereby promoting inclusive healthcare experiences for all users.

5.1 RQ: How are accessibility concerns addressed in the state of the practice of mobile health apps for MS?

The simple answer to the research question is that accessibility concerns are not addressed well enough in the state of practice of mobile health apps for MS. This conclusion is based on the findings of the study, which revealed that all the assessed applications had suggestions from the Accessibility Scanner or encountered problems when being read aloud by TalkBack.

While one application could have passed the Accessibility Scanner assessment as all the suggestions were for advertisements within the application, it is important to note that different impairments have different requirements for accessibility. Therefore, the results should be evaluated individually in relation to the specific accessibility requirements of different impairments and according to the chosen measurements used in the study.

5.1.1 Cognitive impairments

Cognitive impairments do not have countable requirements as other impairments. Accessibility mobile for mobile applications by Ballantyne et al., (2018) had two guidelines specially targeted to cognitive impairments. Both of those were related to content flashing. Neither testing tools could test this aspect, but none of the applications had any flashing content.

TalkBack measurements related to cognitive impairments revealed an unclear state of tabs and menus. In some applications, although the current page or tab was visually highlighted, this highlighting was not conveyed by the screen reader. One of the WCAG success criteria is providing location information to users, indicating their position within the application or webpage (W3C, 2018). This criterion is particularly relevant to cognitive impairments, as simplified menu structures are considered beneficial. Since the screen reader did not provide an indication of location, the requirement for fulfilling the location succession criterion was not met in some of the tested applications.

5.1.2 Motor impairments

Accessibility Scanner results are in line with motor impairments requirements, like touch target size and that touch target are surrounded by inactive space (Ballantyne et al., 2018). Accessibility Scanner provides touch target suggestions that include a minimum touch target size and consider the padding around the element.

There were 165 individual suggestions for 9 applications. Suggestions were categorized as tolerable if the touch target size exceeded 38dp, indicating that the touch target size could be easily adjusted. Touch target sizes were classified as poor if they ranged between 37dp and 25dp, and severe if they were below 25dp. If a size was assessed as severe, it indicated that it would be challenging to fix. If it the size was assessed as severe, it meant that it was not easy to fix. These results can be found in Table 4.

Many of the suggestions were categorized as poor, while the number of tolerable and severe suggestions was equal. Suggestions below 37dp were generally more common. There were differences between applications, but most of them had at least one poor suggestion. Out of the ten tested applications, eight did not meet the requirements for touch target size, and these requirements could not be fulfilled with minor adjustments.

Multiple sclerosis patients who would use applications for managing and keeping track their own health would not be able to use them, if they had motor impairment symptoms, like ataxia or dystonia. Inputs too small and too close to each, especially in case of scale inputs for symptom tracking, would make the usage of application hard. Almost all the applications with symptoms logging had the suggestions with inputs. However, most of the time these inputs were assessed as tolerable, so they would be easy to fix.

Navigation buttons, like back button, next button, and settings button, were also too small in almost every application, but most of the time they were assessed as tolerable. The most severe problem was with more information buttons, which were always assessed as severe. They usually provided necessary information to the element below, that would have helped with usage of the application, but their size made them inaccessible.

Links were assessed as severe if they were within text, which is understandable, but in those case, the links could be changed to buttons for better accessibility. In all applications which had text inputs, those inputs could be easily fixed by adding them more height. They were all full width when compared to width of the screen.

All the elements with suggestions for touch target were vital for using the application. Some of the elements could be easily fixed as mentioned, but others require more work or completely different way to show them, like links within text. Accessibility of multiple sclerosis applications requires work for motor impairment users.

It should be also noted that touch target is important requirement because it is applicable to motor impairments as well as vision impairments. All the mentioned problems with accessibility are also applicable to vision impairment users. Also, other impairments and people without impairments can benefit from it because touch target is part of applications usability.

5.1.3 Vision impairments

Vision impairments has the greatest number of requirements that can be tested with both TalkBack and Accessibility Scanner. Discussion about requirements for touch targets was covered in Motor impairments. This chapter covers additional requirements such as color contrast, item labeling, content order, and alternative descriptions of images, as discussed by Ballantyne et al. (2018).

Accessibility Scanner checks for text and image contrast. The minimum is 4.5:1 contrast ratio, but most of the tested applications did not meet this requirement. There were fewer suggestions related to image contrast compared to text contrast, but all suggestions were classified as poor or severe. Eight applications had at least one suggestion for severe or poor categories regarding their images. However, images are sometimes used as decorations in applications and thus image contrast is not one of the most important accessibility requirements to fill.

Images with poor image contrast, were often icons in navigation button or inputs. If the input or button had a label, the low image contrast would not be a huge problem in accessibility point of view. But if the icon was only thing which differentiated the usage of button or input, then the low image contrast would be a problem for accessibility. Most of the applications had labels with inputs, but navigational buttons did not have visible labels, making those applications inaccessible.

Text contrast ratio on other hand is important because text always provides information. There were 158 individual suggestions for text contrast ratio. Severe category, meaning that contrast ratio between text and background was under 2.5, was the biggest category. In poor category the ratio was between 3.8 and 2.5. The suggestions that fell into the tolerable category, which were easily fixable, were in the minority. Issues with the contrast ratio arose from the use of multiple color combinations or if the application's brand color did not meet the 4.5:1 requirement.

The Accessibility Scanner only assessed a single contrast ratio, although it should be noted that the contrast ratio can be lower for larger font sizes. In the tested applications the lower contrast ratio would not have helped since, the same colours were used consistently in applications. Colour contrast is important aspect because there are lots of different vision impairments, like colour vision and blindness. Colour contrast requirements are also good regarding usability of applications.

The text contrast ratio was mostly related to buttons, and in most cases the problem was that the applications theme colour was not accessible in the beginning. Two applications had several colours in their application and most of the time different colour combinations were inaccessible. Types of elements with inaccessible colour contrast were usually inputs for logging symptoms, data visualization of those symptoms and buttons. Data visualization and logging symptoms colours should be accessible, because of their purpose and usage.

TalkBack measurements for nameless buttons and links align with the results from the Accessibility Scanner regarding item labels. Items without labels could not be read with TalkBack, and there were 105 individual suggestions for item labels. For vision impairments it is important that the application can be used with screen reader, and adding labels to items is important part of screen reader usability and mobile accessibility. Most of the applications encountered issues with item labels. Three applications received a severe assessment as TalkBack failed to recognize certain items entirely.

Completely nameless items were navigation items and more information-buttons, which could be interpreted as that the whole application was inaccessible at that point for vision impairment users, since they provided important information related to application usage. Sometimes navigation items were nameless, which is not better for screen reader, but at least the user can know if the pressed item was button or input.

TalkBack could not read images, or data visualizations in most of the applications. Data visualization part could be only applied to those applications which had it. It is understandable that sometimes data visualization can be hard to read and understand with screen reader, thus skipping data visualization is acceptable if alternative for data visualization is provided. Out of the tested applications, only two implemented the requirement of providing alternative solutions for data visualizations.

Regarding images, only one application had provided alternative descriptions for images and those could be read with screen reader. Other applications did not provide descriptions. Descriptions may not be necessary for images used purely for decorative purposes. In many applications, images were used in relation to input information. In such cases, providing descriptions for inputs containing images, such as smiley faces, would have been beneficial.

The Accessibility Scanner also identified items with identical descriptions. This issue was also observed in the TalkBack results, as elements with the same descriptions could not be differentiated when using TalkBack. This was problem when logging moods and other symptoms. Related to these was also reading order and status of the applications. Most of the applications had a well-defined and coherent reading order that could be followed using TalkBack. Out of the tested applications, one received a bad assessment, four were assessed as good, and the remaining applications were categorized as neutral. Neutral assessment was that there were some problems but those could be improved easily, when bad assessment had a lot more to do regarding accessibility.

Identical descriptions happened when there was navigation bar at the bottom and on the top of the screen, were the screen’s name. Another case was usually when there were several symptom logging inputs in the same screen. Most of the time inputs were from 1 to 5 and all the inputs had the same description, only change was title of the symptom. The same description in several inputs was a problem with screen reader, since all the inputs were same and depending on the reader of the application, those could easily be misremembered and confused to another input.

Vision impairments have the greatest number of requirements and most of those requirements were not met within tested applications. The severity of problems depended on what kind of messages and usage applications have. Generally, the absence of image descriptions and low color contrast are considered detrimental to accessibility. The assessed elements provided valuable information of how severe concerns related to accessibility were within vision impairments.

5.1.4 Hearing impairments

Hearing impairment requirements are not covered by TalkBack or Accessibility Scanner. Hearing impairment requirements are related to videos and their captions. Only two of the tested applications had videos, and those videos were hosted by Youtube. One application had readymade captions in their videos, and in other application it was possible to enabled automatic captions in videos.

The requirements for hearing impairments include captions for videos and no background music (Ballantyne et al., 2018). Although these requirements were not specifically tested, it was evident that all the applications adhered to the requirements. There were no background sounds and videos had possibilities to captions.

5.1.5 Summary of discussion

The research question was: How are accessibility concerns addressed in the state of the practice of mobile health apps for MS? Regarding different impairments most of the applications address accessibility concerns poorly. There were lots of problems with colours, item labels and touch targets. Some of the accessibility problems could be easy to fix but others would require a lot of work.

During the analysis of elements, it was observed that crucial components such as navigation buttons or symptom logging inputs did not adhere to the established guidelines or requirements. This lack of compliance rendered the entire application inaccessible and unusable for individuals with multiple sclerosis. As a result, the intended purpose of these elements and applications, specifically catering to the needs of different multiple sclerosis patients, was not fulfilled due to their inherent inaccessibility.

Two of the most important requirements that need to be addressed are touch targets and item labeling, as they encompass a significant portion of the requirements for motor and vision impairments. In addition, motor and vision impairments are one of the common symptoms in multiple sclerosis, and requirements related to these impairments are easy to test. Hearing and cognitive impairment requirements could not be tested as easily, but some concerns came up during testing sessions, and those concerns were assessed.

6. Future research

The research for this study involved benchmarking, as well as the use of Accessibility Scanner and the TalkBack screen reader. The benchmarking process revealed a significant need for improvements in almost all of the applications, particularly in the context of motor and vision impairments. However, it is important to note that cognitive impairments were not included in the testing due to certain limitations imposed by the selected testing applications and the specific nature of cognitive impairment requirements. Consequently, there is a clear need for further research that specifically focuses on addressing cognitive impairments in application accessibility.

One potential approach for future research is manual testing, where each application is thoroughly examined by human evaluators who assess different cognitive impairment requirements. This method would involve a comprehensive review using the human eye to identify areas where improvements are needed to cater to cognitive impairments. Another viable testing possibility is to engage in user testing, where applications are tested by individuals with multiple sclerosis (MS) who currently experience cognitive symptoms such as fatigue or brain fog. By collecting observations and feedback from these users, researchers can gain valuable insights into the cognitive impairment requirements and identify strategies for enhancing accessibility.

It is worth noting that the research was conducted solely on Android applications. Expanding the research to include iOS applications would provide a more comprehensive understanding of how accessibility concerns are addressed in iOS applications designed for individuals with multiple sclerosis. This expansion would enable researchers to gain insights into the approaches and considerations employed in iOS development, further enhancing the overall accessibility landscape for individuals with MS.

An innovative approach for future research of mobile health accessibility research involves selecting health applications that cater to different specific target groups. One such group is femtech, which encompasses devices, applications, and services designed for female health (Stewart, 2023). Femtech is a rapidly growing industry, with a global market size of 51 billion dollars in 2021, projected to reach 103 billion dollars by 2030 (Stewart, 2022). Despite its significance, femtech remains relatively underexplored in terms of accessibility research.

Ensuring the accessibility of femtech applications is crucial due to the diverse health experiences of women. For instance, fatigue is a common symptom during pregnancy, and menstrual periods can significantly impact the daily lives of women, causing symptoms such as cramps and fatigue. By investigating the accessibility considerations within femtech applications, the specific needs of women and enhance their overall user experience could be addressed.

Conducting thorough research on femtech accessibility could contribute to improving the usability and inclusivity of these applications. It could also foster a better understanding of how technology can support and empower women in managing their health. By considering the unique challenges and symptoms faced by women in different life stages, such research could pave the way for the development of more accessible and user-centric femtech solutions.

Overall, there are several opportunities for future research related to accessibility in mobile health applications as presented above and more. The research opportunities could have great impact on inclusivity, usability, and accessibility of mobile health applications, which would benefit wide range of users.

7. Conclusions

This thesis aimed to address the question of how accessibility concerns are addressed in the state of practice of mobile health apps for MS. The research approach involved testing the accessibility of various applications based on prior research findings. It was observed that all the applications had accessibility problems, except for one that passed the assessment when excluding advertisements from the data.

The research method was benchmarking which provided structure for testing plan and measurements. The suggestions generated by Accessibility Scanner served as metrics and measurements. Additionally, TalkBack measurements were derived from observations made during testing, serving as additional metrics in the final analysis.

Across different impairments and their corresponding requirements, most of the tested applications did not adhere to the necessary accessibility standards. Only the requirements for hearing impairments were adequately addressed. Numerous challenges were identified in meeting the requirements for motor and vision impairments, indicating a poor level of accessibility in the tested MS applications.

It is important to acknowledge the limitations of this study, which include a small sample size of tested applications and the restriction to a single platform. Not all the found application could be tested because of area restrictions. Moreover, the chosen testing tools did not encompass all aspects of accessibility requirements, particularly those related to cognitive impairments, except for one or two specific requirements. Nevertheless, the selected tools and applications offered a representative sample for testing and benchmarking, as they are widely available and user-friendly on Android phones.

For future research, it would be valuable to focus on accessibility requirements that cannot be adequately tested using automated tools. This may involve manual testing or involving users in the evaluation process. User testing can provide deeper insights into the essential requirements and the inclusiveness of these requirements. Considering the variability of multiple sclerosis symptoms and the evolving accessibility needs of individuals over time, user testing could shed light on how health applications can adapt and meet the changing requirements.

Furthermore, expanding the scope of the study to include iOS applications would provide a more comprehensive understanding of the state of accessibility in mobile health apps for MS across different platforms.

Another possibility for future research of mobile health accessibility lies in researching mobile health applications targeted towards women and how accessibility concerns are addressed within them. This line of inquiry is particularly significant due to the emergence of femtech as a burgeoning industry, which caters to the diverse health needs of women across different life stages, including pregnancy, menopause, and various symptoms experienced throughout their lives. By investigating the accessibility considerations specific to femtech applications, we could ensure inclusivity and usability for a large user base, including women with disabilities or impairments. This kind of research would contribute to enhancing user experiences, improving health outcomes, and fostering inclusivity within the femtech industry.

Overall, the research highlights the critical importance of addressing accessibility concerns in mobile health apps for individuals with MS. Inaccessibility can hinder their ability to manage their condition, access information, and engage in self-care. App developers and stakeholders in the field of mobile health should prioritize accessibility in their development processes to improve health outcomes and enhance the user experiences of individuals with MS.

References

Antonio Mateus, D. A., Aparecido Da Silva, F. A. C., & Pimenta Freire, A. P. (2021). Pandemic Crisis Brings More Digital Governmental Services to Mobile Devices-But Are They Accessible to People with Disabilities? *Proceedings of the 39th ACM International Conference on the Design of Communication: Building Coalitions. Worldwide, SIGDOC 2021*, 197–204. https://doi.org/10.1145/3472714.3473905

Apple Inc. (2022). *Accessibility on iOS - Apple Developer*. https://developer.apple.com/accessibility/ios/

Atula, S., Sinkkonen, S. T., Saat, R., Sairanen, T., & Atula, T. (2016). Association of multiple sclerosis and sudden sensorineural hearing loss. *Http://Dx.Doi.Org/10.1177/2055217316652155*, *2*. https://doi.org/10.1177/2055217316652155

Ballantyne, M., Jha, A., Jacobsen, A., Scott Hawker, J., & El-Glaly, Y. N. (2018). Study of accessibility guidelines of mobile applications. *ACM International Conference Proceeding Series*, 305–315. https://doi.org/10.1145/3282894.3282921

BBC. (2022). *Glossary - Accessibility for Products - BBC*. https://www.bbc.co.uk/accessibility/forproducts/guides/mobile/glossary/

Cao, J., Lim, Y., Sengoku, S., Guo, X., & Kodama, K. (2021). Exploring the Shift in International Trends in Mobile Health Research From 2000 to 2020: Bibliometric Analysis. *JMIR Mhealth Uhealth 2021;9(9):E31097 Https://Mhealth.Jmir.Org/2021/9/E31097*, *9*(9), e31097. https://doi.org/10.2196/31097

Ceci, L. (2023). *Global: mobile internet users 2019-2028 | Statista*. https://www.statista.com/forecasts/1146312/mobile-internet-users-in-the-world

Chen, S., Chen, C., Fan, L., Fan, M., Zhan, X., & Liu, Y. (2022). Accessible or Not? An Empirical Investigation of Android App Accessibility. *IEEE Transactions on Software Engineering*, *48*(10), 3954–3968. https://doi.org/10.1109/TSE.2021.3108162

CML. (2016). *Human Factors (HF); Guidelines for the design of mobile ICT devices and their related applications for people with cognitive disabilities*. https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx

Dawe, M. (2007). Understanding mobile phone requirements for young adults with cognitive disabilities. *ASSETS'07: Proceedings of the Ninth International ACM SIGACCESS Conference on Computers and Accessibility*, 179–186. https://doi.org/10.1145/1296843.1296874

Dias, A. L., Fortes, R. P. D. M., & Masiero, P. C. (2012). Increasing the quality of web systems: By inserting requirements of accessibility and usability. *Proceedings - 2012 8th International Conference on the Quality of Information and Communications Technology, QUATIC 2012*, 224–229. https://doi.org/10.1109/QUATIC.2012.33

Eler, M. M., Rojas, J. M., Ge, Y., & Fraser, G. (2018). Automated Accessibility Testing of Mobile Apps. *Proceedings - 2018 IEEE 11th International Conference on Software Testing, Verification and Validation, ICST 2018*, 116–126. https://doi.org/10.1109/ICST.2018.00021

Fan, D., Fay Siu, A., Sung-ho Kim, G., Vazquez, X., Greco, L., Follmer, S., Vazquez, X., Follmer, S., Siu, A. F., Rao, H., Rao, H., & Sung-Ho Kim, G. (2023). The Accessibility of Data Visualizations on the Web for Screen Reader Users: Practices and Experiences During COVID-19. *ACM Transactions on Accessible Computing*, *16*(1). https://doi.org/10.1145/3557899

Ghosh, R., Roy, D., Dubey, S., Das, S., & Benito-León, J. (2022). Movement Disorders in Multiple Sclerosis: An Update. *Tremor and Other Hyperkinetic Movements*, *12*(1). https://doi.org/10.5334/TOHM.671

Google. (2022). *Build more accessible apps | Android Developers*. https://developer.android.com/guide/topics/ui/accessibility

Google. (2023a). *Get started with Accessibility Scanner - Android Accessibility Help*. https://support.google.com/accessibility/android/answer/6376570?hl=en

Google. (2023b). *Test your app's accessibility | App quality | Android Developers*. https://developer.android.com/guide/topics/ui/accessibility/testing

Google. (2023c). *Use automatic captioning - YouTube Help*. Youtube Help. https://support.google.com/youtube/answer/6373554?hl=en#zippy=%2Cautomatic-captions-on-long-form-videos-and-shorts

Hasselbring, W. (2021). Benchmarking as Empirical Standard in Software Engineering Research. *ACM International Conference Proceeding Series*, 365–372. https://doi.org/10.1145/3463274.3463361

Hoff, J. M., Dhayalan, M., Midelfart, A., Tharaldsen, A. R., & Bø, L. (2019). Visual dysfunction in multiple sclerosis. *Tidsskrift for Den Norske Legeforening*, *139*(11). https://doi.org/10.4045/TIDSSKR.18.0786

Instagram. (2023). *Accessibility | Instagram Help Centre*. https://help.instagram.com/308605337351503?cms_id=308605337351503

Jain, M., Diwakar, N., & Swaminathan, M. (2021). Smartphone usage by expert blind users. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3411764.3445074

Kalb, R., Beier, M., Benedict, R. H. B., Charvet, L., Costello, K., Feinstein, A., Gingold, J., Goverover, Y., Halper, J., Harris, C., Kostich, L., Krupp, L., Lathi, E., LaRocca, N., Thrower, B., & DeLuca, J. (2018). Recommendations for cognitive screening and management in multiple sclerosis care. *Https://Doi.Org/10.1177/1352458518803785*, *24*(13), 1665–1680. https://doi.org/10.1177/1352458518803785

Kistowski, J. V., Arnold, J. A., Huppler, K., Lange, K. D., Henning, J. L., & Cao, P. (2015). How to build a benchmark. *ICPE 2015 - Proceedings of the 6th ACM/SPEC International Conference on Performance Engineering*, 333–336. https://doi.org/10.1145/2668930.2688819

Kounev, S., Lange, K.-D., & Kistowski, J. von. (2020). Benchmarking Basics. *Systems Benchmarking*, 3–21. https://doi.org/10.1007/978-3-030-41705-5_1

Lippincot, B., Thompson, N., Morris, J., Jones, M., & DeRuyter, F. (2020). Survey of user needs: Mobile apps for mhealth and people with disabilities. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *12377 LNCS*, 266–273. https://doi.org/10.1007/978-3-030-58805-2_32/TABLES/4

Morris, M. R., Zolyomi, A., Yao, C., Bahram, S., Bigham, J. P., & Kane, S. K. (2016). "With most of it being pictures now, I rarely use it": Understanding Twitter's evolving accessibility to blind users. *Conference on Human Factors in Computing Systems - Proceedings*, 5506–5516. https://doi.org/10.1145/2858036.2858116

Nicolau, H., & Jorge, J. (2012). Elderly text-entry performance on touchscreens. *ASSETS'12 - Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility*, 127–134. https://doi.org/10.1145/2384916.2384939

Niman, B. Von, Böcker, M., Floratos, N., Martinez-Normand, L., Pluke, M., Schneider, M., & Whitney, G. (2015). Requirement and input collection: Development of guidelines to allow people with cognitive disabilities to exploit the full potential of mobile ICT. *MobileHCI 2015 - Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services Adjunct*, 1024–1029. https://doi.org/10.1145/2786567.2794317

Official Statistics of Finland (OSF). (2021). *Internetin käyttö eri laitteilla 2021, %-osuus väestöstä*. https://www.stat.fi/til/sutivi/2021/sutivi_2021_2021-11-30_tau_012_fi.html

Regional State Administrative Agency. (2022). *Legislation and standards on web accessibility - Web Accessibility Requirements*. https://www.webaccessibility.fi/requirements-of-the-act-on-the-provision-of-digital-services/

Richens, E. P., Williams, G. F., Knight, J., Mathews, M., & Nancarrow, R. (2022a). *Error messages and correction - Accessibility for Products - BBC*. https://www.bbc.co.uk/accessibility/forproducts/guides/mobile/error-messages-and-correction/

Richens, E. P., Williams, G. F., Knight, J., Mathews, M., & Nancarrow, R. (2022b). *Mobile Accessibility Guidelines - Accessibility for Products - BBC*. https://www.bbc.co.uk/accessibility/forproducts/guides/mobile/

Ruzic, L., & Sanford, J. A. (2017). Usability of Mobile Consumer Applications for Individuals Aging with Multiple Sclerosis. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10277 LNCS*, 258–276. https://doi.org/10.1007/978-3-319-58706-6_21

Shiver, B., & Wolfe, R. (2015). Evaluating alternatives for better deaf accessibility to selected web-based multimedia. *ASSETS 2015 - Proceedings of the 17th International ACM SIGACCESS Conference on Computers and Accessibility*, 231–238. https://doi.org/10.1145/2700648.2809857

Siebra, C., Gouveia, T. B., Macedo, J., B da Silva, F. Q., M Santos CIn, A. L., Correia, W., Penha, M., & Anjos Fabiana Florentin, M. (2017). *Toward Accessibility with Usability: Understanding the Requirements of Impaired Uses in the Mobile Context*. https://doi.org/10.1145/3022227.3022233

Silva, C., Medeiros Eler, M., & Fraser, G. (2018). *A survey on the tool support for the automatic evaluation of mobile accessibility*. *8*, pages. https://doi.org/10.1145/3218585.3218673

Statista. (2018). • *Total mHealth market size worldwide 2025 forecast*. https://www.statista.com/statistics/938544/mhealth-market-size-forecast-globally/

Stewart, C. (2022). *Global femtech market size 2030 | Statista*. https://www.statista.com/statistics/1333181/global-femtech-market-size/

Stewart, C. (2023). *The femtech industry - statistics and facts | Statista*. https://www.statista.com/topics/10267/femtech/#topicOverview

Tomlinson, B. J., Schuett, J. H., Shortridge, W., Chandran, J., & Walker, B. N. (2016). "Talkin'' about the weather": Incorporating TalkBack functionality and sonifications for accessible app design." *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2016*, 377–386. https://doi.org/10.1145/2935334.2935390

W3C. (1999). *Web Content Accessibility Guidelines 1.0*. https://www.w3.org/TR/WAI-WEBCONTENT/

W3C. (2009). *How WCAG 2.0 Differs from WCAG 1.0 | Web Accessibility Initiative (WAI) | W3C*. https://www.w3.org/WAI/WCAG20/from10/diff.php

W3C. (2018). *Web Content Accessibility Guidelines (WCAG) 2.1*. https://www.w3.org/TR/WCAG21/

W3C. (2020). *What's New in WCAG 2.1 | Web Accessibility Initiative (WAI) | W3C*. https://www.w3.org/WAI/standards-guidelines/wcag/new-in-21/

W3C. (2022a). *Cognitive Accessibility at W3C | Web Accessibility Initiative (WAI) | W3C*. https://www.w3.org/WAI/cognitive/

W3C. (2022b). *Mobile Accessibility at W3C | Web Accessibility Initiative (WAI) | W3C*. https://www.w3.org/WAI/standards-guidelines/mobile/

W3C. (2022c). *WCAG 3 Introduction | Web Accessibility Initiative (WAI) | W3C*. https://www.w3.org/WAI/standards-guidelines/wcag/wcag3-intro/

W3C. (2022d). *Web Content Accessibility Guidelines (WCAG) 2 Overview | Web Accessibility Initiative (WAI) | W3C*. https://www.w3.org/WAI/standards-guidelines/wcag/

W3C. (2023). *What's New in WCAG 2.2 Draft | Web Accessibility Initiative (WAI) | W3C*. https://www.w3.org/WAI/standards-guidelines/wcag/new-in-22/

Whittaker, R. (2012). Issues in mHealth: Findings From Key Informant Interviews. *J Med Internet Res 2012;14(5):E129 Https://Www.Jmir.Org/2012/5/E129*, *14*(5), e1989. https://doi.org/10.2196/JMIR.1989

Yeratziotis, A., Achilleos, A., Koumou, S., Thibodeau, R. A., Vanezi, E., Geratziotis, G., Papadopoulos, G. A., Iasonas, I., & Kronis, C. (2022). Accessible System and Social Media Mobile Application for Deaf Users. *ACM International Conference Proceeding Series*, 39–47. https://doi.org/10.1145/3524458.3547234

Yu, D., Parmanto, B., & Dicianno, B. (2019). An mHealth App for Users with Dexterity Impairments: Accessibility Study. *JMIR Mhealth Uhealth 2019;7(1):E202 Https://Mhealth.Jmir.Org/2019/1/E202*, *7*(1), e9931. https://doi.org/10.2196/MHEALTH.9931

Yu, D. X., Parmanto, B., Dicianno, B. E., Watzlaf, V. J., & Seelman, K. D. (2014). Accessible mHealth for patients with dexterity impairments. *ASSETS14 - Proceedings of the 16th International ACM SIGACCESS Conference on Computers and Accessibility*, 235–236. https://doi.org/10.1145/2661334.2661402

Zhou, L., Saptono, A., Setiawan, I. M. A., & Parmanto, B. (2020). Making Self-Management Mobile Health Apps Accessible to People With Disabilities: Qualitative Single-Subject Study. *JMIR Mhealth Uhealth 2020;8(1):E15060 Https://Mhealth.Jmir.Org/2020/1/E15060*, *8*(1), e15060. https://doi.org/10.2196/15060