Computers & Graphics 117 (2023) 61–72

Contents lists available at ScienceDirect

Computers & Graphics

journal homepage: www.elsevier.com/locate/cag

Special Section on RAGI 2023

Self-report user interfaces for patients with Rheumatic and Musculoskeletal Diseases: App review and usability experiments with mobile user interface components

Francisco Nunes [a,*], Petra Rato Grego [b], Ricardo Araújo [a], Paula Alexandra Silva [b]

[a] Fraunhofer Portugal AICOS, Porto, Portugal
[b] Department of Informatics Engineering, Center of Informatics and Systems, University of Coimbra, Coimbra, Portugal

ARTICLE INFO

Article history:
Received 22 July 2023
Received in revised form 1 October 2023
Accepted 15 October 2023
Available online 17 October 2023

Keywords:
Self-report
Self-monitor
Self-tracking
Self-care technologies
Mobile app
Rheumatic and musculoskeletal diseases

ABSTRACT

Rheumatic and Musculoskeletal Diseases (RMDs) affect 120 million Europeans and are responsible for joint inflammation, stiffness, pain, and fatigue. Patient-Reported Outcome Measures (PROMs), essential to diagnosis and treatment adjustments, are expected to revolutionise rheumatology care if mobile apps reach clinical practice. However, patients often experience finger dexterity issues that can hinder their interaction with mobile apps. This paper investigates the interaction of patients with RMDs with mobile apps for self-report. We started by reviewing existing iPhone and Android apps for RMDs, to identify common user interface (UI) components, and conducted usability experiments with 20 patients with RMDs to record their performance. The usability experiments showed that in-line selectors are the best-performing UI component and that column selectors are considered the most usable by patients. Sliders perform worse than in-line selectors, with significant differences. Results also showed little difference between test conditions aligned with mobile UI design guidelines and those that provided larger or more spaced targets, leading us to conclude that following existing Apple Human Interface Guidelines and Android Material Design will lead to apps with UIs that are appropriate for patients with RMDs.

1. Introduction

Rheumatic and Musculoskeletal Diseases (RMDs) affect 120 million Europeans with a high prevalence among older adults (∼25% in some countries) [1–3]. Patients with RMDs experience a variety of symptoms that often include: joint inflammation, stiffness, pain, or fatigue [4–6], all of which can strongly impact mobility and finger dexterity [7]. Many RMDs do not have a cure, so patients need to learn to live with the condition and to engage in self-care [8]. The care of patients with RMDs usually consists of regular visits to the physician, where patients are examined and answer a number of Patient-Reported Outcome Measures (PROMs) that characterise the impact of symptoms in their everyday life [9].

Mobile apps have the potential to support patients' self-care, by facilitating the collection of PROMs, promoting self-reflection, and encouraging healthy behaviours [10,11]. In fact, collecting PROMs remotely at home is key to supporting the improvement of rheumatology care, as it can help bring patients' perspectives and preferences to treatment decisions [12]. However, patients with RMDs often experience finger dexterity issues that impact data input, coarse precision, and thus impact their ability to self-report symptoms [13].

This paper explores the interaction of patients with RMDs with mobile User Interfaces (UIs) for self-reporting symptoms. We started by reviewing mobile apps for iPhone and Android to understand the most used UI components in today's self-report apps for RMDs. Having gathered the most common UI components, we planned and conducted usability experiments to understand the interaction of patients with them.

Our study has two main contributions. First, it presents a systematic review of mobile apps for patients with RMDs, that focuses on self-monitoring features and the UI components used, and complements previous reviews focused on general features and content of rheumatology apps in Germany [14] or self-management apps used in scientific studies [15]. Second, the paper reports on usability experiments conducted with patients with RMDs, providing recommendations on UI design choices that can lead to better performance for this user group, and complementing prior work studying other population groups [16,17].

* Corresponding author.
E-mail addresses: francisco.nunes@fraunhofer.pt (F. Nunes), Uc2017277429@student.uc.pt (P. Rato Grego), ricardo.araujo@fraunhofer.pt (R. Araújo), paulasilva@dei.uc.pt (P.A. Silva).

https://doi.org/10.1016/j.cag.2023.10.009
0097-8493/

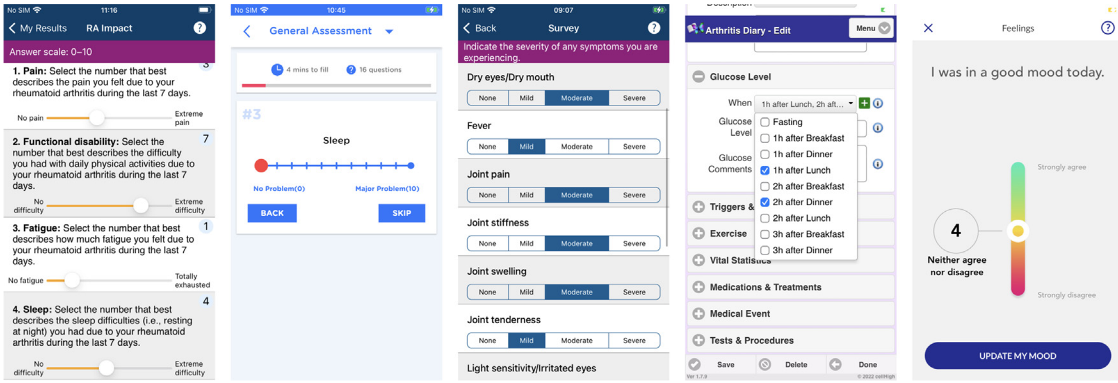


Fig. 1. Screenshots of four apps (RA Manager, Arthritis+Patient, Arthritis Diary, My Arthritis) for patients with RMDs that appeared in the app review. The examples show different data collected by patients, including PROMs, glucose level, exercise and others. Moreover, it is possible to observe different UI components used including horizontal sliders, vertical sliders, and in-line selectors.

This paper is structured into eight sections. Section 2 reviews the background of this study, including the impact of RMDs on touchscreen interaction, and mobile app design guidelines to deal with finger disability or dexterity issues. Section 3 describes the methods of this paper. Section 4 presents the results from the app review and Section 5 explains the UI component selection and validation. Section 6 details the results of the usability experiments. Section 7 discusses the results, contributions, and limitations of this work. Section 8 concludes this work and points to future work opportunities.

2. Background

2.1. Self-monitoring apps in RMDs

The European Alliance of Associations for Rheumatology argues that mobile apps can transform the mode and quality of healthcare, offering patients a more proactive role in their health and care [18]. Self-care apps for rheumatology can include symptom monitoring, education, or physical activity tracking features [12], similar to apps for other chronic conditions [19]. Using a health-related app might modify an individual's health beliefs and behaviours, improve knowledge and abilities in self-care, increase self-efficacy to manage symptoms, reduce health risk behaviours and lead to improved clinical outcomes [13]. Moreover, apps with self-monitoring features, here referred to as self-monitoring apps, have the potential to collect information about the patient's disease activity at home and in everyday life, to support care [20].

Most self-monitoring apps for RMDs enable the collection of PROMs, encouraging patients to classify the severity of their symptoms on a scale (Fig. 1 show different examples of apps collecting PROMs). These self-monitoring apps provide ways to record individual information that, once recorded over time and visually displayed, allow the identification of patterns [12]. Patterns and trends on symptom flare, treatment response and general day-to-day management provide helpful information for the individual in making future decisions regarding self-care. Individuals can also provide this data to their physician to inform discussions about their disease state and treatment adjustment decisions [21]. Furthermore, fluctuations and peaks in disease activity are easily missed or remain unnoticed, which can have severe repercussions in joint damage [22]. Self-monitoring apps have the potential to provide a better understanding of disease fluctuations in-between outpatient clinical appointments. However, patients with RMDs experience symptoms that affect their hand dexterity, consequently influencing their interactions with technology [13]. This warrants the need for research in this area.

2.2. The impact of RMDs on smartphone use

Patients with RMDs are more likely to experience pain and disability in their hands compared with the general population or other people affected by chronic illnesses [13]. For example, Rheumatoid arthritis causes joint stiffness, pain or swelling that persists for over a few weeks [23]. Osteoarthritis is linked to loss of cartilage, joint stiffness, inflammatory pain, and dysfunction [24]. Psoriatic arthritis is characterised by symptoms that include swollen and tender joints, and dactylitis (inflammation of an entire digit) [25,26].

Pain, stiffness, swelling, and limited movement all affect user dexterity and present constraints to smartphone use. Even though prior work [27] affirms that patients do not see their symptoms as barriers to smartphone data input, it is important to consider the input challenges of this user group, as smartphone screens are small and require simple mechanism for data input. People with dexterity issues may not be able to move their hand as quickly, may lack the flexibility of movement, may be unable to touch a button or to touch it involuntarily, or can have so much pain that their hand movement is prevented. Therefore, attention to design aspects of apps, such as button size, spacing, and design components, is critical to improving usability, adherence, and clinical outcomes [13].

Research in RMDs is scant regarding technology usability and accessibility. Apps in the market are rarely usability tested, which decreases their potential to be effectively used [20]. Multimodal input approaches, drawing on electroencephalography [28], eye tracking [29], or voice-based commands [30], which have supported other input-impaired groups, have not been explored with patients with RMDs. Moreover, and while studies stress the potential of digital health technologies for patients with RMDs (e.g., [20]), there is a lack of studies that investigated smartphone touchscreen interaction with this group.

2.3. Recommendations for dexterity and finger disability

We could not find specific recommendations for designing UIs for patients with RMDs, however, there is general advice that can apply to people with finger dexterity issues. In fact, the W3C [31], Google [32,33], and Apple [34] have developed guidelines to make user interfaces more inclusive of people with a broader range of abilities and levels of function. The W3C Input Modalities guidelines [31] mention that when interacting with touchscreen technologies like the smartphone, the finger is the pointer and that all functionalities should be accessible using finger gestures. For people with hand dexterity issues, whose hand precision can be compromised, the W3C Input Modalities

Guidelines recommends providing large enough targets for users to easily activate them, even if the user is accessing content on a small handheld touchscreen device [35].

The Android Material Design (AMD) guidelines 3 [33] recommend touch targets to be at least 48 × 48 dp, between 7 and 10 mm, or larger "to accommodate a larger spectrum of users". The target spacing should be "8 dp" or more to "promote balanced information density and usability".

According to Apple's Human Interface Guidelines (HIG) [34], an accessible app supports accessibility personalisations by design and gives everyone a great user experience, regardless of their physical abilities or how they use their iOS devices. In line with this, the HIG introduce two guidelines that are applicable when addressing challenges experienced by people with hand dexterity issues. First, opt for simplified gestures instead of complex ones (multi-finger, long press, or repeated press). Second, ensure all targets measure at least "44 × 44 pt", approximately 7 mm.

3. Methods

This section describes the steps we took in this research. We start by describing the systematic app review, to then explain the process of selection and validation of UI components. The section also explains the main decisions underlying the research as well as the procedures and participants involved.

3.1. Systematic mobile app review

Our app review was guided by the research question: *What are the UI components used in self-monitoring apps for patients with RMDs?*. Focusing on self-report and self-monitoring apps meant that we excluded apps that lacked symptom monitoring features, or that solely described the diagnosis/symptoms of RMDs. The focus on patient-facing apps excluded symptom assessment scales or instruments designed for clinicians.

3.1.1. Search expression

After multiple discussion sessions among the authors, we decided to use the following search expression, which combines the symptom 'arthritis' with the most common rheumatic conditions [1] (written in English or Portuguese):

> "arthritis" OR "artrite" OR "ankylosing spondylitis" OR "espondilite anquilosante" OR "rheumatoid arthritis" OR "artrite reumatóide" OR "psoriatic arthritis" OR "artrite psoriática" OR "osteoarthritis" OR "osteoartrose" OR "osteoporosis" OR "osteoporose" OR "polymyalgia rheumatic" OR "polimialgia reumática" OR "lupus" OR "lúpus" OR "sjögren" OR "myositis" OR "miosite" OR "scleroderma" OR "esclerodermia"

3.1.2. App store databases

We searched the Google Play Store and iOS App Store, the two leading mobile app stores, which account for 99% of the market [36]. To search for the apps and collate information about them, we used SerpAPI [37]. SerpAPI web app extracts all available information from Google Play Store and iOS App Store, is sensitive to boolean search operators, and exports the collected data as a JSON (JavaScript Object Notation) file. Our search was conducted on March 2022, yielding 759 records: 469 Android apps, 178 iOS apps, and 112 apps available for both Android and iOS.

3.1.3. App selection strategy

The screening of the apps was performed using Microsoft Excel. Following the search, we converted SerpAPI's JSON file into CSV (Comma-Separated Values) for analysis. Each app description was screened by one author, who consulted with the remaining authors in case of doubts. The first screening resulted in 89 apps to potentially include in the review. After installing and using these apps, we reached a final set of 18 apps that fulfilled the review criteria: (i) mobile apps should be designed for patients and include self-reporting of symptoms or other condition-related aspects (e.g., pain, medication intake, quality of life); (ii) the language of the mobile app should be Portuguese or English; and (iii) the app should be possible to install in Portugal. We excluded apps that were: (i) impossible to register or use without externally provided credentials, (ii) impossible to open or use due to errors, (iii) incompatible with iOS 15 or Android 12 (the current operating system versions at the time), (iv) repeated apps with a different name, or (v) developed for a purpose other than healthcare (e.g., games).

3.1.4. Analysis

The 18 applications were used for several days to self-report data and explore their features. During the experimentation period, the researcher took screenshots and noted down the characteristics of each of the apps on an Excel spreadsheet (18 rows × 94 columns at the end). The analysis method resembled a thematic analysis [38], where we used open and iterative coding to note down different characteristics of the apps. We used columns in Microsoft Excel to code different app characteristics, and we returned to previous apps each time a column was changed or added. Apps that appeared in both app stores were reviewed separately, to capture the UI elements used in the different versions. We also created a Mural[1] board with screenshots of the mobile apps, organised by self-report category and user interface element, to support visual analysis of the apps.

3.2. UI component selection and validation

The systematic app review identified six main UI components used to support self-report in mobile apps for patients with RMDs. Based on these results, we started to plan usability experiments, by contrasting the selected UI component configurations with the mobile UI design guidelines from iPhone and Android, and an expert review session with designers.

3.2.1. Contrast app screens with mobile UI design guidelines

The first step in selecting UI components for the usability experiments was to contrast the uncovered UI component configurations with the mobile UI design guidelines for iPhone (Apple Human Interface Guidelines [34]) and Android (Google Material Design [32,33]). This helped set minimum usability standards, ensuring all usability experiments' conditions complied with basic UI design guidelines from Google or Apple. The guidelines were not definitive in some areas, either because there was no guidance (e.g., how to display the value on horizontal sliders) or because iPhone and Android guidelines were in some disagreement (e.g., tick marks recommended for iPhone and optional for Android); which led us to test different configurations for the UI components. At the end of this phase, we had a plan for the usability experiments that we discussed with a group of designers.

[1] Mural is a virtual tool that enables teams to collaborate visually and brainstorm solutions to their problems or challenges. The web app is available at: https://www.mural.co/.

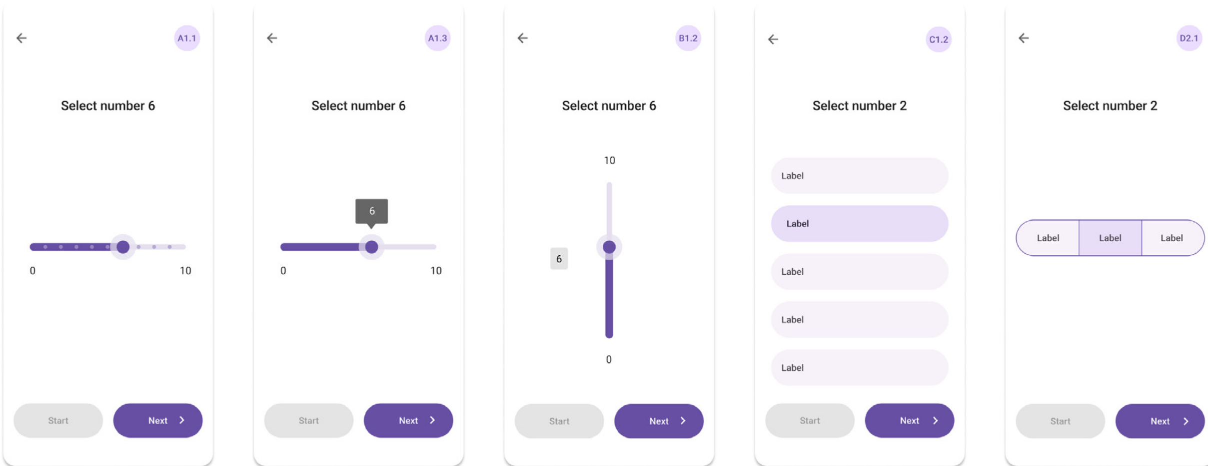


Fig. 2. Screenshots of different UIs used in the usability experiments.

3.2.2. Expert review with designers

The expert review was organised in two parts. First, we made a brief Powerpoint presentation outlining the main insights of the app review, with a particular focus on the UI components used to self-report symptoms. The presenter described the most common UI components and how they varied in different apps, as well as the rationale for selecting UI components for the usability experiments. Second, experts and researchers engaged in a discussion about the review, the UI components identified, and the rationale for selecting UI components to test. At the end of the session, there was a final list of components to test and insights regarding the test conditions.

The review engaged four designers with backgrounds in communication design and multimedia and a track record in designing user interfaces for mobile devices. All of them worked at Fraunhofer Portugal AICOS, but had not been previously involved in the study. Their experience spanned from 2 to 15 years, with most of them working in mobile UI design and user experience for over five years.

3.3. Usability experiments

We conducted usability experiments to assess the interaction of patients with RMDs with the most common UI components found in self-monitoring apps. The app review and validation session helped identify four main UI components – horizontal sliders, vertical sliders, column selectors, and in-line selectors – which we tested through usability experiments. Each of the four UI components was configured with different characteristics to understand whether changes on them resulted in changes in the performance of patients. For example, we tested horizontal sliders with and without tick marks, with different ways of displaying a value label (or with no value label), and with distinct thumb sizes.

3.3.1. Test conditions

Participants completed 28 tests with distinct UI component configurations, performing three or four tasks of value input each time. Table 3 lists all test conditions, detailing the UI component configurations, the numbers participants were asked to input, and example snapshots of some of the UI component configurations. Test conditions were designed to align with iOS and Android Material Design guidelines [32–34], and went beyond them in some tests to understand if patients with RMDs required additional accessibility accommodations.

3.3.2. Materials and apparatus

The usability experiments were enabled by a mobile app developed at Fraunhofer Portugal AICOS specifically for this purpose (see Fig. 2). Each app screen was composed of: instructions on top, navigation buttons (start, next) at the bottom, and the UI component to test in between. Pressing the *start button* began counting the time of the activity and enabled the participant to interact with the UI component. The *next button*, stopped time counting and directed the participant to the following screen. In addition to time, the app also recorded the number of taps to allow for the quantification of errors. Each session included 3–4 tasks of number input, and there were similar instructions in some of the tests to enable the direct comparison of results. The *next button* helped ensure tasks were completed as expected, as the button only became activated after the participant correctly performed the input task. After completing all activities, a detailed log of the session was created and saved to the smartphone storage.

The study was conducted on a OnePlus 7T, an Android smartphone that measures 160.9 × 74.4 mm and has a resolution of 1080 × 2400px at 402 ppi. The smartphone screen was recorded with an app and a GoPro HERO7 was attached to a tripod to film the hands of participants interacting with the smartphone (from above). All participants performed the tasks while sitting on a chair in front of a table, holding the device as they typically use their smartphone.

3.3.3. Procedure

Usability experiments were divided into three parts. First, we introduced participants to the session, the test activities, and how participants could interact with each task screen (start, task, next). Participants were encouraged to clarify doubts and let the researcher know if they felt pain or needed a break. Second, participants engaged in a series of sessions of number input, grouped by the UI component being tested. Once all the tasks of a specific UI component were completed, participants were asked to rate the most usable UI component configuration. After all test sessions, participants were asked to choose their (i) preferred UI component and the (ii) one that caused them the most pain.

3.3.4. Analysis

Statistical analysis was conducted in SPSS and organised in four steps. First, we used descriptive statistics such as averages and frequency distributions, to characterise the study sample, calculate average task completion time, and average number of errors, or the preferred UI components. Second, we used Wilcoxon

Table 1
Characteristics of participants from usability experiments.

No.	Age	Sex	Condition	Diagnosis[a]
P1	41–50	F	OA	3
P2	41–50	F	RA	45
P3	51–60	F	RA	8
P4	41–50	F	RA	40
P5	31–40	F	RA	4
P6	51–60	F	RA	6
P7	51–60	F	PsA	10
P8	41–50	M	PsA	7
P9	61–70	M	RA, OA	12, 5
P10	51–60	F	OA	1
P11	51–60	F	PsA	23
P12	41–50	M	RA	30
P13	71–80	F	PsA	50
P14	51–60	F	PsA	25
P15	41–50	F	RA	12
P16	61–70	M	PsA	2
P17	51–60	F	RA	10
P18	51–60	F	RA	11
P19	71–80	F	OA	8
P20	51–60	F	OA	3

[a] Number of years since. ***OA*** - osteoarthritis. ***RA*** - rheumatoid arthritis. ***PsA*** - psoriatic arthritis.

tests [39] to compare average completion times when completing the same input tasks (e.g., selecting number 2) on different UI components and UI component configurations, to determine the influence of different visual characteristics on performance. Third, we used the Friedman test [40] (with Bonferroni correction for multiple tests) to compare test conditions grouped by UI component. Fourth, we used Spearman's correlation coefficient [41] to assess correlations between UI component configurations and the task completion time or number of errors, to test the possible existence of an association between task completion time and the number of gesture interactions with screen elements.

3.3.5. Participants

We involved 20 participants with RMDs (see Table 1). Sixteen participants were female and four were male. Participant's age ranged from 31 to 80, with nine participants aged 51–60, and six aged 41–50. Ten participants had rheumatoid arthritis, six had psoriatic arthritis, and five had osteoarthritis. All participants were right-handed (dexterous) and used the smartphone multiple times per day.

Recruitment criteria focused on: (a) having a rheumatic condition, (b) being able to travel to the Fraunhofer Portugal AICOS offices to perform the usability experiments, and (c) using a smartphone regularly. Participants were recruited through a Facebook post of the Portuguese League Against Rheumatic Diseases, and through personal networks of the research team.

The study was approved by NOVA Medical School ethics board (no. 221/2021/CEFCM, 19/01/2022). All participants provided written informed consent. Potential participants were presented with the study's goals, an overview of the activities, and a summary of the data privacy practices. Having clarified doubts and asked questions, participants provided written informed consent.

4. Systematic mobile app review results

To determine the most common mobile UI components used in self-reporting apps for RMDs, we carried out a systematic review of available apps. The following sections describe general characteristics of the apps as well as their self-monitoring features and UI components used.

4.1. Characterisation

The review captured 18 apps (see Table 2). The older apps in the review were released in 2011 (DAS28 and Manage My Pain). 2012 and 2013 had no apps published, but starting in 2014, at least one app was published yearly. Most apps were available for both iOS and Android (11), four apps were only available for iOS, and three were only available for Android. All apps were free, except *Arthritis Diary*, which costed 5€.

The apps in the review targeted distinct RMDs. The most commonly targeted conditions were: rheumatoid arthritis (8 apps), psoriatic arthritis (6 apps), or spondyloarthritis (4). Osteoarthritis and Lupus had three apps each, and a few apps targeted other conditions like Scleroderma, Sjögren syndrome, or did not target any specific RMD.

The apps varied in terms of features.[2] All apps supported self-monitoring due to the inclusion criteria. 13 apps had a medication management module, with a medication list, reminders, or the ability to report intakes. 11 apps had a health education module, with information about symptoms, treatment, or news about one or more RMDs. Eight apps had an exercise management module allowing users to keep track of their exercise.

The usability of the apps was poor in general. Many apps featured small buttons, checkboxes, and slider thumbs, with little space between elements. While the apps were framed for patients with RMDs, the UI design did not usually comply with general UI design guidelines, e.g., button target size, which gave the idea that their designers and developers did not consider the target users' characteristics in the design of the apps.

4.2. Self-monitoring features

The apps in the review enabled patients to self-report various condition-related aspects. Symptom monitoring was the most common feature and was present in all of the apps. Patients could use the apps to report: mood or stress (11 apps), pain level (10), joint inflammation or pain (10), fatigue level (9), sleep quality (6), or skin problems (3). The user interfaces for symptom self-report contained a question (e.g., "How much pain have you felt?", "What best describes your pain?") or a statement (e.g., "Pain", "Pain intensity", "Pain score"), which was accompanied by a UI component for reporting that particular symptom (e.g., pain intensity) on a scale (e.g., 0–10, 1–10, or 1–3). The apps sometimes displayed multiple questions/statements and UI components, but most interfaces invited users to report one aspect per screen.

It was possible to report wellbeing and (physical) functionality in eight apps. Wellbeing reporting was similar to symptom self-report, asking users to rate how they felt, and displaying a UI component to collect answers. The reporting of functionality was more varied, with some apps asking one single question about functionality, while others presented multiple questions to obtain a functionality score. The least common self-reporting features were: quality of life (3 apps) or an open field to report any information they would like (5 apps).

4.3. Employed UI components

The apps employed six UI components to capture user self-report data (see Table 2). The most common UI component was the **horizontal slider**[3] (9 apps), which allowed users to select a value from a range provided on the screen (e.g., 1–10), by

[2] Due to space constraints we do not present a detailed list of features from the apps in the review, however, readers can refer to prior work for an overview of the kinds of apps available for RMDs [18].

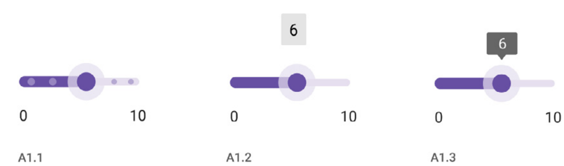
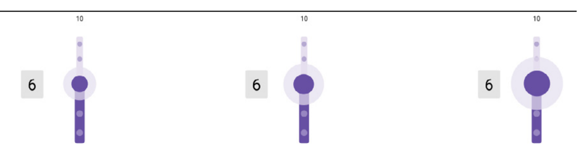
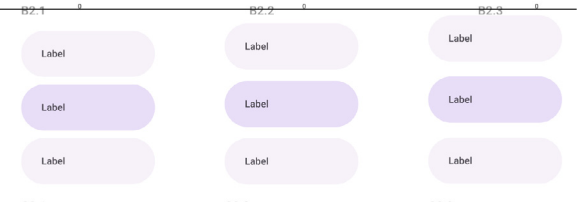
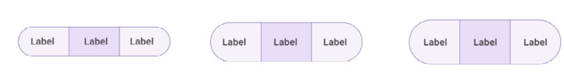
[3] The general preference for sliders can be connected with the prevalence of PROMs in rheumatology care, where physicians regularly ask patients at the clinic to rate their pain, fatigue, sleep, or functionality, on a scale from 1–10.

Table 2
Characteristics of the apps reviewed.

App name	App store			Conditions						Tracking					UI components used					
	Store	Launch year	Price (euros)	Rheumatoid Arthritis	Psoriatic Arthritis	Spondyloarthritis	Osteoarthritis	Lupus	Unspecific or other	Wellbeing	Functionality	Symptoms	Quality of life	Open field	Fatigue level	Skin	Sleep quality	Mental health	Pain level	Assess joint status
Arthritis Diary [42,43]	b	'14	5	–	–	–	–	–	x	x	–	x	–	–	–	–	CB	CB	CB	CB
Arthritis+ Patient [44]	b	'20	0	x	x	x	–	–	x	x	–	x	–	–	HS	–	HS	HS	–	BG
ArthritisPower [45,46]	b	'15	0	–	–	–	–	–	x	–	x	x	–	–	S	–	–	–	S	–
Chronic Insights [47]	g	'21	0	–	x	x	–	–	x	x	–	x	–	x	CS	–	–	CS	VS	–
cliexa-RA [48,49]	b	'16	0	x	–	–	–	–	–	–	–	x	–	–	–	–	–	–	–	BG
DAS28/ACR-EULAR [50]	i	'11	0	x	x	–	–	–	–	–	–	x	–	–	–	–	–	–	–	BG
Elsa [51,52]	b	'19	0	x	–	–	–	–	x	x	–	x	–	x	HS	BG	–	HS	HS	BG
GeoPain:Home [53]	g	'18	0	–	–	–	–	–	x	–	–	x	–	–	–	–	–	–	HS,VS	–
GRAPPA App [54]	b	'18	0	–	x	–	–	–	–	–	x	x	–	–	HS	HS	HS	HS	HS	–
Jointfully Osteoarthritis [55,56]	b	'16	0	–	–	–	x	–	–	x	–	x	–	–	–	–	–	S	–	BG,HS
LupusMinder [57,58]	b	'17	0	–	–	–	–	x	–	–	–	x	–	x	–	–	–	–	HS	–
Manage My Pain [59,60]	b	'11	0	–	–	–	–	–	x	–	x	x	–	–	–	–	–	S	–	–
My Arthritis [61,62]	b	'19	0	x	x	–	x	x	x	x	x	x	–	–	–	–	VS	VS	VS	–
MySpA [63,64]	i	'18	0	–	–	x	–	–	–	x	x	x	x	x	CB	S	–	–	–	BG
Pain Diary [65]	g	'18	0	–	–	–	–	–	x	–	–	x	–	–	–	–	–	–	HS	–
RA Manager [66]	i	'16	0	x	–	–	–	–	–	x	x	x	x	x	HS,S	–	HS	HS	HS	HS,S
RA Monitor [67,68]	b	'17	0	x	–	–	–	–	–	–	x	x	–	–	HS	–	–	HS	–	BG,CB
Rheumatic Monitor [69]	i	'21	0	x	x	x	x	x	x	–	x	x	x	–	S	–	S	S	S	BG,S

Store: g - Google Play Store, i - iOS App store, b - both Google Play Store and iOS App store. ***UI components***: BG - Body graphic, CB - Checkbox, CS - Circular Slider, HS - Horizontal Slider, S - Selector, VS - Vertical Slider.

Table 3
Usability experiments conditions, input, and UI screen examples.

ID	Tested UI component configurations	Input values	Test scenarios snapshot examples
A1.1	HS w/ tick marks	1, 2, 6, 8	
A1.2	HS w/ a static value label	1, 4, 8, 9	
A1.3	HS w/ a value label following the thumb	2, 4, 6, 9	
A2.1	HS w/ tick marks & static value label (centre)	2, 3, 6, 8	
A2.2	HS w/ tick marks & static value label (bottom)	1, 3, 7, 8	
A3.1	HS w/ a thumb 1x size of AMD2	2, 3, 6, 8	
A3.2	HS w/ a thumb 1.25x size of AMD2	2, 3, 6, 9	
A3.3	HS w/ a thumb 1.5x size of AMD2	1, 3, 7, 9	
B1.1	VS w/ tick marks	1, 2, 7, 9	
B1.2	VS w/ a static value label	1, 3, 6, 8	
B1.3	VS w/ a value label following the thumb	2, 3, 8, 9	
B2.1	VS w/ a thumb 1x size of AMD2	2, 4, 7, 8	
B2.2	VS w/ a thumb 1.25x size of AMD2	2, 3, 7, 9	
B2.3	VS w/ a thumb 1.5x size of AMD2	2, 4, 7, 8	
C1.1	CS w/ 1x size & 1x spacing of AMD3	5, 1, 3	
C1.2	CS w/ 1x size & 1.5x spacing of AMD3	2, 5, 1	
C1.3	CS w/ 1x size & 2x spacing of AMD3	1, 3, 5	
C2.1	CS w/ 1.25x size & 1x spacing of AMD3	1, 4, 2	
C2.2	CS w/ 1.25x size & 1.5x spacing of AMD3	3, 1, 4	
C2.3	CS w/ 1.25x size & 2x spacing of AMD3	4, 2, 1	
C3.1	CS w/ 1.5x size & 1x spacing of AMD3	4, 1, 3	
C3.2	CS w/ 1.5x size & 1.5x spacing of AMD3	1, 4, 2	
C3.3	CS w/ 1.5x size & 2x spacing of AMD3	3, 4, 1	
D1.1	IS at the centre of the screen	1, 3, 2	
D1.2	IS at the bottom of the screen	2, 3, 1	
D2.1	IS w/ 1x size of AMD3	1, 4, 2	
D2.2	IS w/ 1.25x size of AMD3	3, 1, 2	
D2.3	IS w/ 1.5x size of AMD3	2, 1, 3	

HS - Horizontal Slider, VS - Vertical slider, CS - Column Selector, IS - In-line Selector. AMD2/3 - Google Material Design version 2 or 3.

swiping their finger from left to right. In the reviewed apps, the left extremity was associated with the lower range value (e.g., 1) and the right with the higher range value (e.g., 9). The horizontal sliders in the UIs varied in terms of the screen position, presence of tick marks, presence and behaviour of value labels, or the width of the slider thumb (name of the element that the user drags throughout a slider).

The **vertical sliders** were similar to horizontal sliders, but users went from bottom to top to select a value. In this type of slider, the lowest value appeared at the bottom end, while the highest appeared at the top. This UI component appeared in three apps and, similarly to horizontal sliders, varied regarding the position, presence of tick marks, presence and behaviour of value labels, or the width of the slider thumb.

Another variation was the **circular slider** which appeared in *Chronic Insights*. The interaction with the circular slider component started at the top of the circle, and as the user swiped the finger down the circle, the value selected advanced.

The **body graphic**, a map representation where users report the area where they feel pain, inflammation, or skin issues, was used in eight apps. The illustration of the body varied, with some apps adopting a more medical representation, faithful to the anatomical body characteristics (e.g., Chronic Insights), and others resorting to cartoon-like representations of the body that exaggerated areas of interest, such as the hands or feet (e.g., Elsa). In some cases, the body map used showed the whole body (e.g., GeoPain:Home), while in others, it focused on a specific region of the body such as the hand (e.g., cliexa-RA). The apps that recorded the inflammation in the joints amplified the hand when the user selected it, to select the specific joints of the hand where the user felt pain.

Selectors, sets of buttons on the screen to report different values or answers to the self-monitoring prompt, appeared in seven apps. This UI component had two different types: (i) **column selectors**, where a series of buttons were displayed vertically (e.g., RA Manager), or **in-line selectors**, where a series of buttons were displayed horizontally (e.g., Rheumatic Monitor). The UI component varied from app to app, regarding the size of each button and spacing between buttons.

Checkboxes – the square buttons with a check to indicate their state – appear in three apps of the review. Similar to selectors, checkboxes appeared together in columns (e.g., Arthritis Diary) and in-line (e.g., MySpA) sets of buttons. The UI components vary in terms of the position on the screen, target size box, or position of the text label.

5. UI component selection and validation results

Having identified the most common UI components, we contrasted the app screens with the UI design guidelines from AMD2 and AMD3 [32,33] and Apple HIG [34]. This helped pre-select the UI components and devise test conditions, which we later discussed with a panel of design experts. Finished the expert review, we had a list of component configurations to test.

5.1. Results from contrasting screens with UI design guidelines

Considering the guidelines on **horizontal sliders**, we planned to test four different aspects. First, test horizontal sliders with and without *tick marks*, because while Apple HIG recommends the use of tick marks to increase the clarity and accuracy of horizontal sliders [34], AMD2 considers it as optional [32]. Second, test *different placements for the value* chosen in the horizontal slider, because there is no recommendation of whether the value should be static or move with the thumb. Third, test *different slider thumb sizes*, assuming the AMD2 minimum thumb radius of 20 dp [32], as well as 1.25x and 1.5x its size, as patients with RMDs might need more generous thumb sizes. Fourth, we would test horizontal sliders centred and at the bottom of the screen, as guidelines do not recommend one or the other. Moreover, we followed the minimum slider height of 6 dp and inactive track minimum height of 4 dp recommended in AMD2 [32].

The **vertical sliders** would be tested under similar conditions, except for the onscreen position. Although vertical sliders are not included in AMD2/3 [32,33] or Apple HIG guidelines [34], this UI component was included since it appeared in several reviewed apps. The reason for excluding the onscreen position is that it was already tested in the horizontal sliders.

We excluded **circular sliders** because they appear in only one app and are not mentioned in either AMD2/3 or Apple HIG.

Assessing the guidelines applicable to **column selectors**, we planned to test two different aspects. One aspect to test was the *selector size*, which according to AMD3 should have a minimum height of 56 dp [33]. We used this value (56 dp) as the minimum height and tested it against 1.25x and 1.5x of that size, assuming patients with RMDs might need the additional button height. Another aspect to test was the *selector spacing*, which should enable users to distinguish the selector from neighbouring elements and information [34]. We used the minimum in-between button spacing of 12 dp [33,34], and tested it against 1.5x and 2x of that space.

The **in-line selectors** should be tested similarly to column selectors but with two exceptions. Selector spacing does not apply to in-line selectors, as all buttons appear in the same line. Moreover, we needed to test centre and bottom onscreen positions because guidelines do not recommend for one or the other.

We excluded **checkboxes** from the usability experiments because guidelines recommend treating checkboxes as selectors where users can click the square button or the label text [32]. Since we already tested two types of selectors (column, in-line), testing another similar UI component was deemed unnecessary.

The **body graphic** was also excluded from the usability experiments. It would be possible to test target size of body region selectors, but we already tested selectors did not expect to learn more from testing selectors in a body graphic. We asked participants to choose the most appropriate body map representations, but that qualitative analysis should be published elsewhere.

In summary, we proposed to take the horizontal sliders, vertical sliders, column selectors, and in-line selectors to the usability experiments.

5.2. Expert review validation results

After contrasting the screens with existing guidelines and preselecting UI components to test, we organised an expert review session with designers to assess the usability experiments plan. The session started with a Powerpoint presentation, showing insights from the systematic app review and the rationale for the pre-selection. Discussion ensued about the conducted work.

The experts validated the presented plan – to test horizontal sliders, vertical sliders, column selectors, and in-line selectors – and made two recommendations. The first recommendation was that the usability experiments focus on the touchscreen input and avoid any need for content interpretation to reduce the impact on response time and potential confounding noise that could emerge due to the need of interpretation. Even though testing (original) app screens would provide more context for users, it could be detrimental to measurement and comparison; thus we chose to create screens from scratch, where participants only saw one simple instruction to select an input value. The second recommendation was to use navigation buttons on the screen to support measurement and ensure the task was effectively completed successfully. Following this recommendation, we included buttons at the bottom to signal the start and the conclusion of the task, which enabled us to clock the time of the task and account for errors in entering input.

6. Usability experiments results

This section presents the results of the usability experiments. Our analysis looked at the average task completion times of the different test conditions, contrasting it with the subjective preference of participants regarding the most usable test condition interface or UI components (e.g., horizontal slider vs. selector). We also planned to analyse the average number of errors in the

Table 4
Usability experiments results.

Test	Input	Input 1 results		Input 2 results		Input 3 results		Input 4 results		Total avg	
ID	values	Time[a]	Err[b]	Time	Err	Time	Err	Time	Err	Time[c]	Err[d]
A1.1	1, 2, 6, 8	4,60 ± 2,61	0,00	3,01 ± 1,02	0,00	9,09 ± 4,61	0,90	4,53 ± 2,27	0,00	5,31 ± 2,63	0,26
A1.2	1, 4, 8, 9	5,11 ± 3,88	0,50	5,46 ± 2,04	0,20	5,36 ± 2,42	0,65	4,45 ± 2,13	0,25	5,10 ± 2,62	0,40
A1.3	2, 4, 6, 9	3,76 ± 1,51	0,15	4,02 ± 1,37	0,15	3,99 ± 1,81	0,25	3,95 ± 1,32	0,30	3,93 ± 1,50	0,21
A2.1	2, 3, 6, 8	3,15 ± 1,37	0,10	3,06 ± 1,16	0,00	3,65 ± 1,12	0,00	3,24 ± 1,00	0,00	3,27 ± 1,16	0,03
A2.2	1, 3, 7, 8	2,62 ± 0,74	0,00	2,89 ± 1,2	0,00	3,62 ± 1,12	0,00	3,33 ± 1,46	0,00	3,33 ± 1,13	0,00
A3.1	2, 3, 6, 8	3,15 ± 1,37	0,10	3,06 ± 1,16	0,00	3,65 ± 1,12	0,00	3,24 ± 1,00	0,00	3,27 ± 1,16	0,03
A3.2	2, 3, 6, 9	2,92 ± 1,10	0,00	2,96 ± 0,93	0,00	3,42 ± 1,22	0,00	3,85 ± 1,45	0,20	3,29 ± 1,18	0,05
A3.3	1, 3, 7, 9	2,81 ± 0,97	0,10	3,34 ± 2,00	0,00	3,38 ± 1,66	0,10	3,11 ± 1,18	0,00	3,16 ± 1,45	0,05
B1.1	1, 2, 7, 9	3,74 ± 2,06	0,20	6,53 ± 3,77	0,45	7,82 ± 5,04	0,15	3,53 ± 1,69	0,00	5,40 ± 3,14	0,20
B1.2	1, 3, 6, 8	3,42 ± 1,56	0,00	4,26 ± 2,02	0,70	4,15 ± 1,94	0,40	3,86 ± 1,02	0,00	3,92 ± 1,64	0,28
B1.3	2, 3, 8, 9	3,54 ± 1,69	0,35	3,40 ± 0,95	0,00	3,98 ± 1,67	0,20	3,28 ± 1,25	0,00	3,55 ± 1,39	0,14
B2.1	2, 4, 7, 8	2,58 ± 0,8	0,00	2,66 ± 1,25	0,00	4,10 ± 1,87	0,00	3,10 ± 0,73	0,00	3,11 ± 1,16	0,00
B2.2	2, 3, 7, 9	2,71 ± 1,06	0,00	3,06 ± 1,62	0,00	3,51 ± 1,07	0,00	3,37 ± 1,27	0,20	3,16 ± 1,26	0,05
B2.3	2, 4, 7, 8	2,78 ± 1,14	0,10	3,14 ± 1,14	0,10	3,32 ± 0,96	0,00	2,96 ± 0,89	0,00	3,05 ± 1,03	0,05
C1.1	5, 1, 3	2,03 ± 0,64	0,00	1,89 ± 0,77	0,00	2,35 ± 1,18	0,00			2,09 ± 0,86	0,00
C1.2	2, 5, 1	1,77 ± 0,60	0,00	1,89 ± 0,83	0,00	1,93 ± 0,89	0,00			1,91 ± 0,77	0,00
C1.3	1, 3, 5	2,04 ± 0,91	0,00	1,73 ± 0,79	0,00	1,88 ± 0,79	0,00			1,88 ± 0,83	0,00
C2.1	1, 4, 2	1,98 ± 0,68	0,00	1,77 ± 0,72	0,00	1,96 ± 0,76	0,00			1,91 ± 0,72	0,00
C2.2	3, 1, 4	1,92 ± 0,71	0,00	1,83 ± 0,66	0,00	1,68 ± 0,64	0,00			1,81 ± 0,67	0,00
C2.3	4, 2, 1	1,96 ± 1,14	0,00	1,90 ± 0,85	0,00	2,08 ± 0,92	0,00			1,98 ± 0,97	0,00
C3.1	4, 1, 3	1,99 ± 1,04	0,00	1,80 ± 1,00	0,00	2,10 ± 0,79	0,00			1,96 ± 0,94	0,00
C3.2	1, 4, 2	1,83 ± 0,71	0,00	1,81 ± 0,69	0,00	1,73 ± 0,54	0,00			1,79 ± 0,65	0,00
C3.3	3, 4, 1	1,82 ± 0,76	0,00	1,67 ± 0,63	0,00	1,82 ± 1,10	0,00			1,77 ± 0,83	0,00
D1.1	1, 3, 2	1,83 ± 0,64	0,00	1,67 ± 0,68	0,00	1,72 ± 0,62	0,00			1,74 ± 0,65	0,00
D1.2	2, 3, 1	1,37 ± 0,71	0,00	1,88 ± 0,92	0,00	1,58 ± 0,76	0,00			1,61 ± 0,80	0,00
D2.1	1, 4, 2	1,83 ± 0,64	0,00	1,67 ± 0,68	0,00	1,72 ± 0,62	0,00			1,74 ± 0,65	0,00
D2.2	3, 1, 2	1,74 ± 0,88	0,00	1,43 ± 0,52	0,00	2,25 ± 1,57	0,00			1,81 ± 0,99	0,00
D2.3	2, 1, 3	1,46 ± 0,53	0,00	1,55 ± 0,56	0,00	1,61 ± 0,69	0,00			1,54 ± 0,59	0,00

[a] Average time to select input (e.g., '1') using specific UI test condition (e.g., A1.1).
[b] Average number of errors in selecting input (e.g., '1') using specific UI test condition (e.g., A1.1).
[c] Average time to complete session of multiple inputs using specific UI test condition (e.g., A1.1).
[d] Average number of errors in completing session of multiple inputs using specific UI test condition (e.g., A1.1).

tasks, however, the number of errors never came close to 1 and the differences between test conditions were too low to be useful to support comparisons.

Table 4 provides an overview of usability experiments' results. The first two columns refer to the test conditions, including the test ID, and the input values the participant entered in those interfaces. The following columns (3–10) report the average completion time and number of errors for each of the three or four sessions with the same test conditions. The last two columns show the average number of errors in completing a session of multiple inputs with a specific UI test condition.

6.1. Comparison between different UI components

The UI component with the **best overall performance** was the in-line selector (see Table 5). Grouping the results of all test conditions with the same UI component, we conclude that in-line selectors are the UI component with the shortest task completion time (M[4] = 1,69 s). Column selectors (M = 1,90 s), vertical sliders (M = 3,69 s) and horizontal sliders (M = 3,83 s) achieve worse performance. Differences are significant between the in-line selector and vertical sliders ($p = .042$) and the in-line selector and horizontal slider ($p = .004$).

Performance results are not entirely aligned with participants' subjective preferences. When asked about the easier UI component, 17 participants referred column selectors, one referred in-line selectors, and two had no presence. These results could encourage choosing column selectors over in-line selectors, as the differences between column and in-line selectors and not statistically significant. Moreover, we also inquired participants about the UI components that cause them more pain to input the numbers. 18 participants referred that they felt no difference between test conditions, and two participants referred column selectors caused them more pain. In any case, column selectors perform as one of the best UI component options.

Table 5
Average task completion time grouping all test conditions from the same UI component.

	Horizontal slider (A)	Vertical slider (B)	Column selector (C)	In-line selector (D)
AVG	3,83	3,69	1,90	1,69
SD	0,91	0,82	0,09	0,1

In general, selectors are better than sliders for restricted answer types. Considering the above-mentioned results, in-line and column selectors will probably obtain a lower average completion time (and preference) than sliders. However, selectors can only fit 3–5 options for button size to fulfil the guidelines. Sliders will thus be useful for answering PROMs that ask participants to enter values on a scale from one to ten, something which would not work as easily with selectors due to the limited screen size. The differences in average task performance of horizontal and vertical sliders are not significant ($p > .05$), so it is possible to use the one that best fits the user interface.

6.2. Horizontal sliders

The **horizontal slider configuration** that achieved the best performance was the horizontal slider with a value following the thumb (A1.3: M = 3,93 s), which was 35% faster than the slider with tick marks (A1.1: M = 5,31 s), and 30% faster than the slider with the static value (A1.2: M = 5,10 s). However, these differences were not significant ($p > .05$).

[4] M stands for Mean.

Regarding the horizontal **slider position**, there was no difference. The average completion time of the task with the horizontal slider at the centre (A2.1) or the bottom (A2.2) was essentially the same (3,27 s vs. 3,33 s). No significant differences were found ($p > .05$).

The results were also very similar concerning the **thumb size**. The largest thumb size, 1.5x AMD2 recommended size (30 dp or 4.8 mm), tested in A3.3 achieved the best performance with an average completion time of 3,16 s. However, the average completion times of the tasks testing the remaining thumb sizes were only 0,11 s (A3.1) or 0,13 s (A3.2) higher, which is negligible ($p > .05$).

Regarding **preference**, participants referred they saw no difference in ease of use between distinct thumb sizes or onscreen positions of the horizontal slider ($N = 20$). When asked about the slider value, seven participants preferred sliders to display a static value, four preferred a value moving with the thumb, four preferred sliders not to display a value, and five referred they had no preference. Moreover, multiple participants mentioned during the tests that it would be beneficial to have tick marks as a way to support them in selecting the value faster.

6.3. Vertical sliders

The **vertical slider configuration** (B1.1, B1.2, B1.3) that achieved the best performance was the slider with a value following the thumb (B1.3), with an average completion time of 3,55 s, which is 52% faster than the slider with tick marks (B1.1: $M = 5{,}40$ s), and 10% faster than the slider with static value. However, the difference was not significant ($p > .05$).

Regarding the **thumb size**, there was no difference. The largest thumb size (B2.3), which is 1.5x AMD2 recommended size (30 dp or 4.8 mm), achieved the best performance with an average completion time of 3,05 s; nevertheless, the other test conditions obtained essentially the same results (B2.1: $M = 3{,}11$ s, B2.2: $M = 3{,}16$).

In terms of **preference**, most participants preferred the slider's value to be displayed ($N = 15$) rather than not ($N = 4$). Nine participants preferred the slider value moving with the thumb, while six participants preferred the value to appear in a static position. Moreover, one participant did not have a preference. Regarding the thumb slider sizes, 15 participants did not have a preference of slider thumb size. From the ones who had a preference, four selected the 1.5x AMD2 size (30 dp or 4.8 mm), and 1 preferred the 1.25x AMD2 size (25 dp or 3.9 mm).

6.4. Column selectors

The column selectors (C1.1-C3.3) did not achieve very different results. The best-performing column selector, C3.3, was the column selector with 1.5x size (84 dp, 13.3 mm) and 2x AMD2 spacing (24 dp, 3.8 mm), achieved an average completion time of 1,77 s. The worst-performing column selector, C1.1, was the 1x size (56 dp, 8.9 mm) and 1x AMD2 spacing (12 dp, 1.9 mm), which achieved an average completion time of 2,09 s, only 18% slower than the faster one. The differences observed in the averages are 0.2 s or smaller and were not significant ($p > .05$).

As for perceived ease of use, most participants saw no difference between the tested conditions. When asked to choose between 1x size and the spacing conditions of 1x, 1.5x, and 2x AMD3, 17 participants did not notice any difference, and one participant preferred each of the different spacing conditions. When asked to choose between 1.25x size and spacing conditions of 1x, 1.5x, and 2x AMD3, 15 participants did not notice any difference, two participants chose 1x spacing, two participants chose 2x spacing, and one participant chose the 1.5x spacing. Finally, when asked to choose between 1.5x size and spacing conditions of 1x, 1.5x, and 2x AMD3, 13 participants did not notice any difference, five preferred 1x spacing, one preferred 1.5x spacing, and one preferred 2x spacing. In summary, most participants saw no difference between test options and the ones that did chose distinct options, leading one to conclude that participants did not feel a condition was radically better than the other test conditions.

6.5. In-line selectors

The in-line selectors (D1.1-D2.3) also achieved very similar results. The in-line selector with the best performance, D2.3, the in-line selector with 1.5x size of AMD3 (84 dp or 13.3 mm), achieved an average completion time of 1,54 s, only 0,2 s better than the worst-performing in-line selectors (D1.1 and D2.1). These differences are negligible (non-significant: $p > .05$) and mean that the position and size of the in-line selector did not influence the end performance in the conditions tested.

As for perceived ease of use, there was some dispersion. When asked to choose between the bottom or centre-aligned options, 10 participants said the centred option was easier to use, seven did not see a difference, and three considered the slider positioned at the bottom of the screen as easier. Regarding the in-line selector size, 18 participants did not express a preference and two preferred in-line selectors with a size 1.25x the Google recommended (70 dp or 11.1 mm).

7. Discussion

This paper explored the interaction of patients with RMDs with mobile UIs for self-reporting symptoms, through systematic app review and usability experiments testing 28 UI component configurations. Our results complement previous reviews that analysed the features and content of rheumatology apps in Germany [14] and self-management apps used in scientific studies [15], by analysing the UI components used in self-monitoring apps in the iOS App Store and the Google Play Store. The usability experiments followed general guidelines from Apple HIG, Google AMD, and W3C, and enabled us to confirm that these guidelines are sufficient for this group. Moreover, the usability experiments complement previous work studying other population groups (e.g., [16,17]), with results that can guide the design of user interfaces for patients with RMDs.

The systematic app review revealed that apps for patients with RMDs had a variety of features, including self-monitoring, medication management, health education, or exercise management, which aligns with previous work [14,15]. In terms of self-monitoring aspects, our review uncovered that apps mainly focused on tracking symptoms, from mood to pain, and less on wellbeing, functionality, or quality of life. Self-report was made through a diversity of UI components that included sliders (horizontal, vertical, circular), column and in-line selectors, checkboxes, and body graphics.

We observed a variety of usability issues in the mobile apps in the review, including small buttons and spacing that would fail to meet mobile UI design guidelines. The lack of user involvement, identified in prior work [14], might have led to these usability issues, which is why Yuqing and Hong [70] argues that healthcare professionals, patients, and developers should collaborate to develop high-quality, evidence-based apps that meet the needs of patients and the perspectives of health professionals.

Results from the usability experiments showed that in-line selectors are the best-performing UI component and that column selectors are considered the most usable by patients. The horizontal and vertical sliders perform worse than in-line selectors, with significant differences.

Slider thumb test conditions are not significantly different, meaning the smaller slider thumb size used, the AMD2 [32] recommend size (20 dp, 3.2 mm) can be employed. This thumb value is aligned with previous work (14 mm square) that tested swipe target sizes with older adults [16] and people with Parkinson's disease [17].

Column and in-line selectors test conditions were also very similar, meaning the smaller tested target selector and spacing can be used. The target size of 56 dp (8.9 mm) of AMD3 [33] is aligned with the recommendation from Apple HIG (44 dp), and Leitão and Silva [16] (14 mm square). The spacing size of 8 dp (1.3 mm) of AMD3 [33] is also aligned with previous work [16] (0 to 10.5 mm).

7.1. Recommendations for design

The first and main recommendation that emerges from this work is to follow Apple HIG and AMD guidelines. Participants were able to use all the UI component configurations tested, and did not benefit substantially from the accessibility adaptations made. Yet, to obtain the best input interaction performance, we suggest that the following three recommendations are applied.

Design sliders (horizontal or vertical) that show the value, with a thumb size of at least 20 dp, and include tick marks if possible. Usability experiments' results did not show significant differences between showing the slider's value on a static or moving position, but these test conditions outperformed sliders without displaying the value, which aligned with participants' preferences. The tested thumb sizes achieved similar results, thus it is possible to use the AMD2 recommended thumb size (20 dp, 3.2 mm) for this user group. Participants mentioned on multiple occasions that tick marks made easier the task of choosing a value in the slider.

Design column selectors with at least 56 dp size and 12 dp in-between spacing. Usability experiments showed that the different column selector test conditions performed very similarly, and most participants considered all test conditions to be equally usable. These results mean that the AMD3 recommended size can be used with this user group.

Design in-line selectors with at least 56 dp size, positioned at the centre if possible. The results from usability experiments showed similar results for the different in-line selector test conditions. Half of the participants referred that they preferred in-line selectors at the centre, which might be a reason for positioning the element in that way.

7.2. Implications for medical practice or industry

We believe our study has two implications for practice or industry. Regarding healthcare delivery, the study suggests that patients with RMDs should be able to use self-monitoring apps that respect existing design guidelines. Even though patients with RMDs face finger dexterity issues that would encourage larger target sizes and spaces [13], applying existing general guidelines from Apple HIG or AMD was enough to guarantee that our participants could attain appropriate performance. This confirms that healthcare professionals can safely prescribe (usable) apps to patients, without expecting major accessibility issues in self-reporting. However, before prescribing an app, we suggest hospitals or clinics first to conduct usability tests with the apps, or to engage designers or researchers in assessing whether existing guidelines are respected in a specific app (in line with [71]). Having done this, it should be possible to prescribe and use the app without significant accessibility barriers.

Results are also positive for industry developing mobile apps for RMDs. Acknowledging that patients with RMDs could perform appropriately with user interfaces that followed existing general design guidelines means that companies might be able to reuse patterns and UIs created and tested for other groups of users. Moreover, we believe our study will help companies choosing between different UI component configurations, as we have contrasted these in the usability experiments.

7.3. Limitations

Our app review has three main limitations. First, we reviewed apps in English and Portuguese, excluding apps available in other languages from being captured. Second, we were restricted to the apps published in the Portuguese iPhone App Store and Google Play Store, excluding apps from vendors that chose to publish their apps only for a specific region or country. Third, our search focused on the official stores for the iPhone and Android. These limitations might have restricted the amount of apps we uncovered and consequently the UI components identified. In future work, it would be interesting to conduct the review in different countries and languages, ensuring that more local apps would be captured and considered as part of the review.

The usability experiments also had three main limitations. The first issue is related to the homogeneity of our study sample. Our results were derived from a sample of 20 right-handed participants, unevenly distributed between sexes, from one country in the south of Europe, with low disease activity (medicated/controlled RMDs), who regularly use a smartphone. Whilst ours was a sizeable sample, it would have been ideal to involve a larger number of participants, evenly distributed between right- and left-handed, and men and women. Second, we did not perform the usability experiments with a control group, which would have been helpful to rule out confounding factors (e.g., age), and help understand if specific results would also apply to more mainstream user groups that do not have RMDs. Third, we did not use real-world mobile app screens, which reduced the ecological validity of the study. We opted to use the screens we designed to have more detailed metrics and less noise in the average completion time, however, in future work, it would be interesting to see how the results translate in real app user interfaces when screens have more context, distractions, and options.

8. Conclusions and future work

This paper examined the interaction of RMDs' patients with mobile user interfaces for self-reporting symptoms. By carrying out a mobile app review, this study identified the most common mobile components used in self-reporting apps for RMD conditions: horizontal sliders, vertical slides, circular slides, body graphics, selectors and checkboxes. Building upon these findings and following a validation through an expert review, 28 tests with distinct UI component configurations were designed to assess user performance with the different UI components and configurations. The tests were divided into four main test conditions, each focusing on a different type of UI component – horizontal sliders, vertical sliders, column selectors, or on in-line selectors – that 20 patients with RDMs tested through usability experiments.

Our results showed that in-line selectors are the best-performing UI component and that column selectors the most usable according to patients. Sliders perform worse than in-line selectors, with significant differences. Existing Apple Human Interface Guidelines and Android Material Design are appropriate to use, as the larger targets and spacing conditions tested do not lead to significantly better results.

The results from our work suggest that medical professionals should be able to prescribe mobile apps to patients without expecting them to encounter significant accessibility issues, provided that apps respect existing design guidelines. Industry developing mobile apps for patients with RMDs should be able to draw on previous patterns and designs created for other population groups, and benefit from knowing the UI component configurations that best work for patients with RMDs that we uncovered in this study.

Given the limitations of our sample, in future usability experiments, it would be valuable to recruit a more diverse and balanced sample of participants, not only in terms of handedness and sex but also in what concerns disease activity. Adding to this, to eliminate confounding factors and enhance the reliability of the findings, it would be important to have a control group. Furthermore, to improve the quality of the evidence and the ecological validity of the study, it would be worth investigating user performance with a more comprehensive list of UI components and configurations from real-world mobile app screens. All in all, further studies would yield more robust and lend increased generalisability to the results.

CRediT authorship contribution statement

Francisco Nunes: Conceptualization, Methodology, Writing – original draft, Supervision, Writing – review & editing. **Petra Rato Grego:** Methodology, Investigation, Formal analysis, Writing – original draft, Writing – review & editing. **Ricardo Araújo:** Formal analysis, Writing – review & editing. **Paula Alexandra Silva:** Conceptualization, Methodology, Supervision, Writing – review & editing.

Declaration of competing interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome. We confirm that the manuscript has been read and approved by all named authors and that there are no other persons who satisfied the criteria for authorship but are not listed. We further confirm that the order of authors listed in the manuscript has been approved by all of us. We confirm that we have given due consideration to the protection of intellectual property associated with this work and that there are no impediments to publication, including the timing of publication, with respect to intellectual property. In so doing we confirm that we have followed the regulations of our institutions concerning intellectual property.

Data availability

Data will be made available on request.

Acknowledgements

We would like to thank Nuno Cardoso and Ricardo Graça, who developed the mobile app to support the usability experiments. This study was supported in part by the project COTIDIANA (AAL2020-7-146-CP) funded under the AAL Joint Programme, by the European Commission and the National Funding Authorities of Portugal, Austria, and Switzerland.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.cag.2023.10.009.

References

[1] EULAR. EULAR website. Tech. rep, EULAR; 2011, URL: www.eular.org.
[2] Branco JC, Rodrigues AM, Gouveia N, Eusébio M, Ramiro S, Machado PM, da Costa LP, Mourão AF, Silva I, Laires P, Sepriano A, Araújo F, Gonçalves S, Coelho PS, Tavares V, Cerol J, Mendes JM, Carmona L, Canhão H, study group E. Prevalence of rheumatic and musculoskeletal diseases and their impact on health-related quality of life, physical function and mental health in Portugal: results from EpiReumaPt- a national health survey. RMD Open 2016;2(1):e000166. http://dx.doi.org/10.1136/rmdopen-2015-000166.
[3] Sebbag E, Felten R, Sagez F, Sibilia J, Devilliers H, Arnaud L. The world-wide burden of musculoskeletal diseases: a systematic analysis of the World Health Organization Burden of Diseases Database. Ann Rheum Dis 2019;78(6):844–8. http://dx.doi.org/10.1136/annrheumdis-2019-215142.
[4] van Tuyl LH, Lems WF, Boers M. Measurement of stiffness in patients with rheumatoid arthritis in low disease activity or remission: a systematic review. BMC Musculoskelet Disord 2014;15:28. http://dx.doi.org/10.1186/1471-2474-15-28.
[5] Nicassio PM, Wallston KA. Longitudinal relationships among pain, sleep problems, and depression in rheumatoid arthritis. J Abnorm Psychol 1992;101(3):514.
[6] Griffith J, Carr A. What is the impact of early rheumatoid arthritis on the individual? Best Pract Res Clin Rheumatol 2001;15(1):77–90.
[7] Luurssen-Masurel N, Weel AE, Hazes JM, de Jong P, tREACH group investigators. The impact of different (rheumatoid) arthritis phenotypes on patients' lives. Rheumatology 2021;60(8):3716–26.
[8] Nikiphorou E, Santos EJF, Marques A, Böhm P, Bijlsma JW, Daien CI, Esbensen BA, Ferreira RJO, Fragoulis GE, Holmes P, McBain H, Metsios GS, Moe RH, Stamm TA, de Thurah A, Zabalan C, Carmona L, Bosworth A. 2021 EULAR recommendations for the implementation of self-management strategies in patients with inflammatory arthritis. Ann Rheum Dis 2021;80(10):1278–85. http://dx.doi.org/10.1136/annrheumdis-2021-220249.
[9] Fautrel B, Alten R, Kirkham B, de la Torre I, Durand F, Barry J, Holzkaemper T, Fakhouri W, Taylor PC. Call for action: how to improve use of patient-reported outcomes to guide clinical decision making in rheumatoid arthritis. Rheumatol Int 2018;38(6):935–47. http://dx.doi.org/10.1007/s00296-018-4005-5.
[10] Solomon DH, Rudin RS. Digital health technologies: opportunities and challenges in rheumatology. Nat Rev Rheumatol 2020;16(9):525–35. http://dx.doi.org/10.1038/s41584-020-0480-7.
[11] Kataria S, Ravindran V. Digital health: a new dimension in rheumatology patient care. Rheumatol Int 2018;38(11):1949–57. http://dx.doi.org/10.1007/s00296-018-4127-y.
[12] Mollard E, Michaud K. Self-management of rheumatoid arthritis: Mobile applications. Curr Rheumatol Rep 2021;23. http://dx.doi.org/10.1007/s11926-020-00968-7.
[13] Mollard E, Michaud K. Mobile apps for rheumatoid arthritis: Opportunities and challenges. Rheum Dis Clin North Am 2019;45:197–209. http://dx.doi.org/10.1016/j.rdc.2019.01.011.
[14] Knitza J, Tascilar K, Messner E-M, Meyer M, Vossen D, Pulla A, Bosch P, Kittler J, Kleyer A, Sewerin P, Mucke J, Haase I, Simon D, Krusche M. German mobile apps in rheumatology: Review and analysis using the mobile application rating scale (MARS). JMIR Mhealth Uhealth 2019;7(8):e14991. http://dx.doi.org/10.2196/14991, URL: https://mhealth.jmir.org/2019/8/e14991/.
[15] Najm A, Gossec L, Weill C, Benoist D, Berenbaum F, Nikiphorou E. Mobile health apps for self-management of rheumatic and musculoskeletal diseases: Systematic literature review. JMIR Mhealth Uhealth 2019;7(11):e14730. http://dx.doi.org/10.2196/14730, URL: https://mhealth.jmir.org/2019/11/e14730.
[16] Leitão R, Silva PA. Target and spacing sizes for smartphone user interfaces for older adults: Design patterns based on an evaluation with users. In: 19th conference on pattern languages of programs. 2012, URL: https://mural.maynoothuniversity.ie/6045/.
[17] Nunes F, Silva PA, Cevada J, Correia Barros A, Teixeira L. User interface design guidelines for smartphone applications for people with Parkinson's disease. Univers Access Inf Soc 2016;15(4):659–79. http://dx.doi.org/10.1007/s10209-015-0440-1.
[18] Najm A, Nikiphorou E, Kostine M, Richez C, Pauling JD, Finckh A, Ritschl V, Prior Y, Balážová P, Stones S, Szekanecz Z, Iagnocco A, Ramiro S, Sivera F, Dougados M, Carmona L, Burmester G, Wiek D, Gossec L, Berenbaum F. EULAR points to consider for the development, evaluation and implementation of mobile health applications aiding self-management in people living with rheumatic and musculoskeletal diseases. RMD Open 2019;5. http://dx.doi.org/10.1136/rmdopen-2019-001014.
[19] Nunes F, Verdezoto N, Fitzpatrick G, Kyng M, Grönvall E, Storni C. Self-care technologies in HCI: Trends, tensions, and opportunities. ACM Trans Comput-Hum Interact 2015;22(6). http://dx.doi.org/10.1145/2803173.

[20] Solomon DH, Rudin RS. Digital health technologies: opportunities and challenges in rheumatology. Nat Rev Rheumatol 2020;16(9):525–35. http://dx.doi.org/10.1038/s41584-020-0461-x.

[21] Colls J, Lee YC, Xu C, Corrigan C, Lu F, Marquez-Grap G, Murray M, Suh DH, Solomon DH. Patient adherence with a smartphone app for patient-reported outcomes in rheumatoid arthritis. Rheumatology (United Kingdom) 2021;60:108–12. http://dx.doi.org/10.1093/rheumatology/keaa202.

[22] Welsing PM, Landewé RB, Riel PLV, Boers M, Gestel AMV, Linden SVD, Swinkels HL, Heijde DMVD. The relationship between disease activity and radiologic progression in patients with rheumatoid arthritis: A longitudinal analysis. Arthritis Rheum 2004;50:2082–93. http://dx.doi.org/10.1002/art.20350.

[23] Sparks JA. In the Clinic® rheumatoid arthritis. Ann Int Med 2019;170:ITC1–ITC15. http://dx.doi.org/10.7326/AITC201901010.

[24] Kean WF, Kean R, Buchanan WW. Osteoarthritis: symptoms, signs and source of pain. Int Immopharmacol 2004;12:3–31.

[25] Boehncke WH, Qureshi A, Merola JF, Thaçi D, Krueger GG, Walsh J, Kim N, Gottlieb AB. Diagnosing and treating psoriatic arthritis: An update. Br J Dermatol 2014;170:772–86. http://dx.doi.org/10.1111/bjd.12748.

[26] Duarte GV, Faillace C, Carvalho JFD. Psoriatic arthritis. Best Pract Res: Clin Rheumatol 2012;26:147–56. http://dx.doi.org/10.1016/j.berh.2012.01.003.

[27] Grainger E, Al. Patient and clinician views on an app for rheumatoid arthritis disease monitoring: Function, implementation and implications. Int J Rheum Dis 2020.

[28] Castiblanco Jimenez IA, Gomez Acevedo JS, Olivetti EC, Marcolin F, Ulrich L, Moos S, Vezzetti E. User engagement comparison between advergames and traditional advertising using EEG: Does the user's engagement influence purchase intention? Electronics 2023;12(1). http://dx.doi.org/10.3390/electronics12010122.

[29] Dewi DS, Ilmi N, Dewi RS. Applying eye tracker technology for analyzing user interface design of online transportation applications (case study: grab application). IOP Conf Ser: Mater Sci Eng 2020;722(1):012038. http://dx.doi.org/10.1088/1757-899X/722/1/012038.

[30] Jian W-S, Wang J-Y, Rahmanti AR, Chien S-C, Hsu C-K, Chien C-H, Li Y-C, Chen C-Y, Chin Y-P, Huang C-L. Voice-based control system for smart hospital wards: a pilot study of patient acceptance. BMC Health Serv Res 2022;22(1):287. http://dx.doi.org/10.1186/s12913-022-07668-1.

[31] W3C. Understanding guideline 2.5: Input modalities. 2021, URL: https://www.w3.org/WAI/WCAG21/Understanding/input-modalities.

[32] Google. Material design 2. 2014, URL: https://material.io/.

[33] Google. Material design 3. 2018, URL: https://m3.material.io/.

[34] Apple. iOS human interface guidelines. 2020, URL: https://developer.apple.com/design/human-interface-guidelines/guidelines/overview/.

[35] W3C. Understanding success criterion 2.5.5: Target size. 2021, URL: https://www.w3.org/WAI/WCAG21/Understanding/target-size.

[36] GlobalStats. Mobile operating system market share worldwide | statcounter global stats. 2021, URL: https://gs.statcounter.com/os-market-share/mobile/worldwide.

[37] SerpApi. SerpApi playground. 2021, URL: https://serpapi.com/playground?engine=apple_app_store&term=rheumatic&country=us&lang=en-us&num=15&page=0.

[38] Braun V, Clarke V. Thematic analysis. American Psychological Association; 2012.

[39] Wilcoxon F. Individual comparisons by ranking methods. Biom Bull 1945;1(6):80–3. http://dx.doi.org/10.2307/3001968.

[40] Friedman M. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. J Amer Statist Assoc 1937;32(200):675–701.

[41] Spearman C. The proof and measurement of association between two things. Am J Psychol 1904;15(1):72–101. http://dx.doi.org/10.2307/1412159.

[42] cellHigh LLC. Arthritis diary - Apps on google play. 2014, https://play.google.com/store/apps/details?id=com.cellhigh.ArthritisDiary&hl=en&gl=US, (Accessed on 07/04/2023).

[43] cellHigh. Arthritis diary on the app store. 2014, https://apps.apple.com/us/app/arthritis-diary/id945349814?platform=iphone, (Accessed on 07/04/2023).

[44] Azaya Technologies Inc. Arthritis+ patient on the app store. 2020, https://apps.apple.com/us/app/arthritis-patient/id1508917640?platform=iphone, (Accessed on 07/04/2023).

[45] Global Healthy Living Foundation. ArthritisPower - Apps on google play. 2015, https://play.google.com/store/apps/details?id=com.csdg.uab.arpower&hl=en&gl=US, (Accessed on 07/04/2023).

[46] Global Healthy Living Foundation. ArthritisPower on the app store. 2019, https://apps.apple.com/us/app/arthritispower/id1004705508?platform=iphone, (Accessed on 07/04/2023).

[47] Chronic Insights Ltd. Chronic insights symptom diary - Apps on google play. 2021, https://play.google.com/store/apps/details?id=com.chronicinsights, (Accessed on 07/04/2023).

[48] cliexa, Inc. Cliexa-RA - Apps on google play. 2016, https://play.google.com/store/apps/details?id=com.cliexa.cliexara&hl=en&gl=US, (Accessed on 07/04/2023).

[49] CN4CE, LLC. Cliexa-RA on the app store. 2017, https://apps.apple.com/us/app/cliexa-ra/id1077589407, (Accessed on 07/04/2023).

[50] Matsui K. DAS28/ACR-EULAR criteria on the app store. 2011, https://apps.apple.com/us/app/das28-acr-eular-criteria/id443707029?platform=iphone, (Accessed on 07/04/2023).

[51] Elsa Science AB. Elsa — Your rheumatic companion - Apps on google play. 2019, https://play.google.com/store/apps/details?id=science.elsa.app, (Accessed on 07/04/2023).

[52] Elsa Science AB. Elsa – Påverka din reumatism on the App Store. 2021, https://apps.apple.com/app/id1455833439, (Accessed on 07/04/2023).

[53] MoxyTech, Inc. GeoPain:Home - 3D chronic pain & symptom manager. 2018, https://www.apkmonk.com/app/com.MoxyTechInc.GeoPain.Home.Release/, (Accessed on 07/04/2023).

[54] Group for Research and Assessment of Psoriasis and Psoriatic Arthritis. GRAPPA app on the app store. 2018, https://apps.apple.com/us/app/grappa-app/id1346646781, (Accessed on 07/04/2023).

[55] MicroHealth Inc. Jointfully osteoarthritis - Apps on google play. 2016, https://play.google.com/store/apps/details?id=com.jointfully&hl=en&gl=US, (Accessed on 07/04/2023).

[56] MicroHealth LLC. Jointfully osteoarthritis on the app store. 2016, https://apps.apple.com/app/jointfully-osteoarthritis/id1068422271?platform=iphone, (Accessed on 07/04/2023).

[57] Hospital for Special Surgery. LupusMinder - Apps on google play. 2017, https://play.google.com/store/apps/details?id=com.hss.lupusminder&hl=en_US, (Accessed on 07/04/2023).

[58] Hospital for Special Surgery. LupusMinder on the app store. 2017, https://apps.apple.com/pt/app/lupusminder/id1287778388?l=en&platform=iphone, (Accessed on 07/04/2023).

[59] ManagingLife, Inc. Manage my pain - Apps on google play. 2011, https://play.google.com/store/apps/details?id=com.lcs.mmp.lite&hl=en&gl=US, (Accessed on 07/04/2023).

[60] ManagingLife. Manage my pain on the app store. 2021, https://apps.apple.com/us/app/manage-my-pain/id1444320523?platform=iphone, (Accessed on 07/04/2023).

[61] Ampersand Health. My arthritis - Apps on google play. 2019, https://play.google.com/store/apps/details?id=nhs.ibd.com.rheumatology&hl=en_US, (Accessed on 07/04/2023).

[62] Ampersand Health Limited. My arthritis on the app store. 2020, https://apps.apple.com/gb/app/my-arthritis/id1431862637, (Accessed on 07/04/2023).

[63] Earthware Ltd. MySpA - Apps on google play. 2018, https://play.google.com/store/apps/details?id=com.earthware.whippscross&hl=en&gl=US, (Accessed on 07/04/2023).

[64] Barts Health NHS Trust. MySpA na app store. 2018, https://apps.apple.com/pt/app/myspa/id1378899849?platform=iphone, (Accessed on 07/04/2023).

[65] Roobol S. Pain diary - Pain management log. 2018, https://www.apkfollow.com/app/pain-diary-pain-management-log/com.samantharoobol.painlog/, (Accessed on 07/04/2023).

[66] @Point of Care. RA manager on the app store. 2016, https://apps.apple.com/us/app/ra-manager/id840972490?platform=iphone, (Accessed on 07/04/2023).

[67] RPM Healthcare. RA monitor - Apps on google play. 2017, https://play.google.com/store/apps/details?id=com.healthmonitor.ramonitor&hl=en&gl=US, (Accessed on 07/04/2023).

[68] RPM Healthcare, LLC. RA monitor on the app store. 2017, https://apps.apple.com/us/app/ra-monitor/id1312851551, (Accessed on 07/04/2023).

[69] BGU Medical Informatics Research Center. Rheumatic monitor on the app store. 2021, https://apps.apple.com/us/app/rheumatic-monitor/id1545357629, (Accessed on 07/04/2023).

[70] Yuqing, Hong. Evaluating Chinese mobile health apps for ankylosing spondylitis management: Systematic app search. JMIR Mhealth Uhealth 2021;9:e27234. http://dx.doi.org/10.2196/27234.

[71] Vollenwyder B, Petralito S, Iten GH, Brühlmann F, Opwis K, Mekler ED. How compliance with web accessibility standards shapes the experiences of users with and without disabilities. Int J Hum-Comput Stud 2023;170:102956. http://dx.doi.org/10.1016/j.ijhcs.2022.102956.