Automatic market research of mobile health apps for the self-management of allergic rhinitis

Aram Antó1 | Bernardo Sousa-Pinto2,3 | Wienczyslawa Czarlewska4,5 | Oliver Pfaar6 | Sinthia Bosnic-Anticevich7,8 | Ludger Klimek9 | Paolo Matricardi10 | Salvatore Tripodi11 | João A. Fonseca2,3,12 | Josep M. Antó13,14,15,16 | Jean Bousquet17,18,19

1MASK-air, Montpellier, France
2MEDCIDS – Department of Community Medicine, Information and Health Decision Sciences, Faculty of Medicine, University of Porto, Porto, Portugal
3CINTESIS – Center for Health Technology and Services Research, University of Porto, Porto, Portugal
4Medical Consulting Czarlewska, Levallois, France
5ARIA, Montpellier, France
6Section of Rhinology and Allergy, Department of Otorhinolaryngology, Head and Neck Surgery, University Hospital Marburg, Philipps-Universität Marburg, Marburg, Germany
7Quality Use of Respiratory Medicine Group, Woolcock Institute of Medical Research, The University of Sydney, Sydney, New South Wales, Australia
8Sydney Local Health District, Sydney, New South Wales, Australia
9AG Molecular Allergology and Digital Health, Department of Pediatric Pneumology, Immunology and Critical Care Medicine, Charité Universitätsmedizin - Berlin, Berlin, Germany
10Center for Rhinology and Allergy, Wiesbaden, Germany
11Allergology Service, Policlinico Casilino, Rome, Italy
12MEDIDA-Medicina, EDucação, I&D e Avaliação, Lda, Porto, Portugal
13ISGlobal, Barcelona Institute for Global Health, Barcelona, Spain
14Universitat Pompeu Fabra (UPF), Barcelona, Spain
15IMIM (Hospital del Mar Medical Research Institute), Barcelona, Spain
16CIBER Epidemiología y Salud Pública (CIBERESP), Barcelona, Spain
17Institute for Allergology, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin, Berlin, Germany
18Allergology and Immunology, Fraunhofer Institute for Translational Medicine and Pharmacology ITMP, Berlin, Germany
19University Hospital of Montpellier, Montpellier, France

Correspondence
Jean Bousquet, 273 Avenue d’Occitanie, 34090 Montpellier, France.
Email: jean.bousquet@orange.fr

Abstract

Background: Only a small number of apps addressing allergic rhinitis (AR) patients have been evaluated. This makes their selection difficult. We aimed to introduce a new approach to market research for AR apps, based on the automatic screening of Apple App and Google Play stores.

Methods: A JavaScript programme was devised for automatic app screening, and applied in a market assessment of AR self-management apps. We searched the Google Play and Apple App stores of three countries (USA, UK and Australia) with the following search terms: "hay fever", "hayfever", "asthma", "rhinitis", "allergic rhinitis". Apps were eligible if symptoms were evaluated. Results obtained with the automatic programme were compared to those of a blinded manual search. As an example, we used the search to assess apps that can be used to design a combined medication score for AR.
INTRODUCTION

Allergic rhinitis (AR), despite not being potentially fatal, has a major impact on social life, school and work productivity.\(^1\) Many guidelines for AR management have been developed, and effective treatments exist.\(^2-4\) However, many patients still have uncontrolled disease. Numerous AR patients rely on self-medication, ignoring or completely bypassing healthcare professionals, which often leads to suboptimal medication.\(^5\) The smartphone is a powerful tool that can address this problem, due to its ubiquity and high potential to change user behaviour.\(^5\) This can be exploited to create new habits, helping patients to improve their self-medication. Mobile health (mHealth) apps represent an important evolution of health and care since they allow the collection of real-world data (including the identification of risk exposure and patients’ behaviours and habits), potentially prompting a profound impact on current guidelines and care pathways.\(^6-8\)

The number of AR apps has increased substantially over the past years, but few have been subject to formal research and evaluation. To the best of our knowledge, four recent studies have searched for AR apps,\(^9-12\) with one of them showing that the use of such apps has increased over the past years.\(^10\) The development of new apps or the absence of updates from existing ones may have rendered these studies outdated.

Quick outdating is one of the greatest challenges for studies reviewing apps, given the pace at which e-tools evolve. There is a lack of updated evidence-based reviews for app selection. The availability of too many AR apps may overwhelm patients when searching (with some important apps possibly appearing too low on the search results), and undermine the distinction between evidence-based versus other apps (this is a problem given the large number of apps publicizing pseudo-therapies, as well as the lack of quality control of the apps).

An important challenge for app review studies concerns the lack of automatic standardized search strategies. Although some automatic methods for screening purposes have been proposed for breast cancer apps,\(^13\) the identification of potentially relevant apps remains a time-consuming manual task. Such limitations could be overcome by the further development of automatic methods for app screening, which would also provide certain advantages: (i) allowing for the running of screening processes frequently and at an increased speed, and (ii) potentially identifying relevant apps whose name and icon are not explicit.

We aimed to perform an assessment of AR self-management apps in three countries, using a new approach consisting of an automatic screening algorithm of Apple App and Google Play stores. We report not only a description of apps for AR self-management, but also the development of a JavaScript programme for automatic app screening and its application to this specific context of identification of AR apps.

RESULTS

The automatic search programme identified 39 potentially eligible apps out of a total of 1593 retrieved apps. Each of the 39 apps was individually checked, with 20 being classified as relevant. The manual search identified 19 relevant apps (out of 6750 screened apps). Combining both methods, a total of 21 relevant apps were identified, pointing to a sensitivity of 95% and a specificity of 99% for the automatic method. Among these 21 apps, only two could be used for the combined symptom-medications score for AR.

CONCLUSIONS

The programmed algorithm presented herein is able to continuously retrieve all relevant AR apps in the Apple App and Google Play stores, with high sensitivity and specificity. This approach has the potential to unveil the gaps and unmet needs of the apps developed so far.

KEYWORDS

allergic rhinitis, app, automatic search, Google, Apple, JavaScript

Key Message

- Although it is essential to assess the quality of mobile health apps for rhinitis, this is the first automatic screening of Apple App and Google Play stores.
- Healthcare professionals will have an overview of apps selected for their clinical relevance.
- Among the 1593 retrieved apps, only 20 were considered as clinically relevant.
we performed a descriptive analysis on the number and distribution of apps identified as potentially relevant. Both automatic and manual searches were performed on the 20th of December 2020. The following countries of app availability were selected: the UK and the US (because the vast majority of apps are registered in these countries) as well as Australia (as this is another English-speaking country).

2.2 | Eligibility criteria of apps

We included all apps relating to the self-management of AR. Apps were included if they allowed users to track the evolution of their symptoms and aimed to help patients manage their disease (i.e., integrating information on risk exposure, symptoms and/or treatment). We excluded apps (i) that did not belong to the “Medical,” “Health & Fitness” or “Weather” app store categories (app stores classify apps according to their subject – e.g., apps for health monitoring may be classified in the “Medical” category, while apps assessing pollen levels or air pollution may be classified in the “Weather” category), (ii) whose description was not in English, (iii) that had not been published or updated within the last 4 years, (iv) that had solely informative or investigative purposes, or (v) whose aim was to promote alternative treatments (Tables 1 and 2).

In order to make sure that we searched and selected apps solely based on their relevance to rhinitis, and to avoid excluding potentially relevant apps, we opted not to define eligibility criteria in the selection process based on ratings and reviews.

In this study, the automatic method for app selection was not developed based on training against a manual classification. In fact, the automatic method and the manual search process were performed separately and in parallel – that is, the classification results were not used to build the training.

2.3 | Automatic search process

An app screening programme was developed using JavaScript,\textsuperscript{14} in order to perform the app search without any human intervention. JavaScript is a programming language that was created by Netscape in 1995.\textsuperscript{15} It is one of the core technologies of the world wide web,\textsuperscript{16,17} and is supported by all major web browsers on the market. It is also widely used for backend applications. One of its main strengths is the huge number of open-source packages freely available for use.\textsuperscript{18}

The screening programme builds upon two open-source packages\textsuperscript{19} that provide a data scraping capability for Apple App and Google Play stores. The packages obtain the app data from Apple and Google by sending requests to the underlying web API that is used by Apple and Google to build the web version of their app stores. The packages wait for the reply and save the apps’ data once a reply is received. In JavaScript, a package is defined as a self-contained set of functions that provide a specific capability.\textsuperscript{20} Data scraping is a technique in which a computer programme extracts data from a human-readable output coming from another programme.

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Apps’ exclusion criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Not belonging to one of the categories: “Medical,” “Health &amp; Fitness” or “Weather”</td>
</tr>
<tr>
<td>2</td>
<td>Description not written in English</td>
</tr>
<tr>
<td>3</td>
<td>Not published/updated within the last 3 years</td>
</tr>
<tr>
<td>4</td>
<td>Target is healthcare professionals</td>
</tr>
<tr>
<td>5</td>
<td>Informative/educational app</td>
</tr>
<tr>
<td>6</td>
<td>No symptom-tracking function available</td>
</tr>
<tr>
<td>7</td>
<td>Aim consists of promoting natural/alternative treatments</td>
</tr>
<tr>
<td>8</td>
<td>Food allergies</td>
</tr>
<tr>
<td>9</td>
<td>Special credentials required. Aim consists of data retrieval for medical studies</td>
</tr>
</tbody>
</table>

The Apple App and Google Play stores’ scraping functions are provided by the packages iterated over the following two input variables (Figure 1):

- Search terms: hay fever, hayfever, asthma, rhinitis, allergic rhinitis
- Countries: United Kingdom (UK), United States (US), Australia

Apple App and Google Play stores were automatically searched, using the search terms hay fever, hayfever, asthma, rhinitis, and allergic rhinitis. The search terms were proposed by a specialist in AR, and correspond to the most commonly used words to search for rhinitis and related terms. For each search term, automatic searches were performed in three different geographical locations: UK, US and Australia. Therefore, a total of 15 iterations for each store were performed (5 search terms × 3 geographical locations). The number of results retrieved at each iteration was limited to a maximum of 200 apps for the Apple App store and 250 for Google Play. These values were set by Apple and Google.

In some cases, for a given search term, certain apps may only be found in one of the two stores, even though they are also registered in the other. This is because the algorithms that determine app selection are different in the Apple App and Google Play stores. In order to address this issue, the programme goes through all the apps found in only one store and searches the other store using the app’s title as a search term.

After retrieving the search results from all iterations, the programme compares the titles of all the retrieved apps, and discards the duplicates. An example of this process is shown in Figures 2 and 3. After duplicate removal, an automatic selection algorithm, based on the defined eligibility criteria, is applied. According to this algorithm, apps were automatically excluded if they did not meet the following criteria:

- belonging to the categories “Medical,” “Health & Fitness” or “Weather.” The “weather” category typically includes apps providing weather forecasts and/or displaying information on pollen levels or air pollution;
- with an English written description;
- published/updated within the last 3 years.
In addition, the selection algorithm encompassed the application of a set of logical statements (we refer to these as “conditions,” defined in Tables 1 and 2), performing text analysis within the apps' title, description and category. This allowed for the automatic exclusion of:

- Apps not meeting condition 1
- Apps not meeting condition 2
- Apps meeting condition 3
- Apps meeting condition 4

For the apps not automatically excluded by the screening algorithm, the following properties were automatically saved for each app: title, app store(s) where available, number of downloads in each store, score in each store, and number of ratings in each store. These apps were manually checked by two independent reviewers who classified them, according to the predefined eligibility criteria, as “fully relevant,” “possibly relevant” or “not relevant.”

### 2.4 Manual search process

In parallel to the automatic search process, a manual search was performed using the same search terms, in Apple App and Google Play stores, in the same geographical locations, and applying the same eligibility criteria. Switching to a different country within the app store was accomplished by following the available instructions. A total of 30 searches were thus performed (5 search terms × 3 geographical locations × 2 app stores). Each search in Google Play retrieved up to 250 results, while each search in Apple App retrieved up to 200.

The manual search was conducted in two stages: (i) two authors (A. A. and B. S.-P) independently screened the titles and icons of the apps resulting from the search terms in each of the considered countries; (ii) potentially relevant apps were subsequently analysed, based on their description. The apps were then classified as “fully relevant,” “possibly relevant” or “not relevant” (Table 3), according to the predefined eligibility criteria (Table 1). All disagreements were solved by consensus.

### 2.5 Benchmarking of search methods

Peer-reviewed papers published on apps classified as “fully relevant” and “possibly relevant” were retrieved using PubMed. The terms used were “name of the app” (and) “app.”

In order to assess the performance of the automatic search process, its sensitivity, specificity and predictive values were calculated by comparing the final number of relevant apps with the number of apps retrieved by the automatic search programme.

In order to assess the performance of the automatic search process, we obtained a confusion matrix based on the number of apps retrieved by the automatic search programme, and allowing its comparison with the results of the manual search process. Based on the confusion matrix, we calculated the agreement between both methods (by estimating the kappa coefficient), as well as the sensitivity, specificity and predictive values of the automatic search programme (considering manual search results as the gold-standard).

### 2.6 Selection of apps for the design of a symptom-medication score

An ARIA-EAACI (Allergic Rhinitis and its Impact on Asthma – European Academy of Allergy and Clinical Immunology) Task Force has developed an optimal symptom-medication score for allergen immunotherapy with an app. The first step was to screen the apps that could be used. The apps retrieved by manual and automatic searches were examined. The following criteria were required:

### Table 2

<table>
<thead>
<tr>
<th>Condition</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5 or 6 is met</td>
</tr>
<tr>
<td>2</td>
<td>The word “symptom*” is present in the description</td>
</tr>
<tr>
<td>3</td>
<td>At least one of the words “home remed*,” “natural remed*,” “homeopath*,” “alternative med*,” “acupressure” is present in the title or description</td>
</tr>
<tr>
<td>4</td>
<td>6.1 is not met and 4.1 is met</td>
</tr>
<tr>
<td>5</td>
<td>5.1 and 5.2 are both met</td>
</tr>
<tr>
<td>6</td>
<td>5 or 6.1 is met</td>
</tr>
</tbody>
</table>

Text searches were always done without distinguishing between uppercase and lowercase letters. The * character is used as a wildmark to represent different morphological forms of a word (e.g., allerg*: allergy, allergies, allergen, allergens). The inclusive variant of the “or” logical operator is used here.

The **character is used as a wildmark to represent different morphological forms of a word (e.g., allerg*: allergy, allergies, allergen, allergens). The inclusive variant of the “or” logical operator is used here.
1. Multilingual with at least English and 3 of the 4 following languages (French, German, Italian or Spanish), given that the app needs to be used in many countries.

2. Sufficient number of users in several countries. From a previous study, it was found that 500–1000 users per country were needed to perform the analyses.\textsuperscript{22}

3. Published methodology (or available on the web) of the app with COSMIN guidelines,\textsuperscript{23,24} independency of results, quality of data, reliability and responsiveness of data.

4. A full list of OTC and prescribed medications customized for each country, since patients may incorrectly ascribe medications to a given category.

5. App updated within the past 2 years, as the list of medications is revised regularly.

3 | RESULTS

3.1 | Automatic search process

The automatic search programme retrieved 1593 apps, of which 39 were automatically identified as potentially eligible, according to the defined conditions (Figure 4). The 39 apps automatically identified were manually checked, with 20 (51\%) being considered as fully or possibly...
relevant (Table 4). Among the remaining apps, 9 were duplicates and 10 did not meet the eligibility criteria. Among the latter, 5 were excluded for being educational, 3 for not being specific to AR, 1 for not having a symptom-tracking function, and 1 for targeting only healthcare professionals. The entire process (running the automatic search programme and subsequent manual check) was found to take 20 min overall.
3.2 | Manual search process

During the manual search process, 6750 apps were screened, of which we read the description of 78 from the Google Play store and 133 from the Apple App store. From these apps, 19 were selected as fully or possibly relevant (Table 4). The manual search process is shown in Figure 4. A total of 10 person-hours were required for the manual search process.

3.3 | Comparison between automatic and manual search processes

There was a large difference in the number of apps identified by automatic and manual searches. The lower number of apps identified by the automatic search is caused by the apps being automatically aggregated when appearing in different searches (different keywords or geographical locations). In the manual search, this aggregation is done at a later stage and therefore the initial number of identified apps is higher.

Combining the automatic and manual search processes, we identified a total of 21 potentially relevant apps for AR self-management, of which 16 were considered to be “fully relevant” and 5 “possibly relevant.” All except one of the relevant apps identified by the manual search were also identified by the automatic search (DailyBreath). On the other hand, two of the apps identified as relevant by the automatic search were not identified by the manual search. All apps missed by one of the search processes were classified as “possibly relevant.”

We observed good agreement between the automatic and the manual search processes (kappa coefficient = 0.661; 95% CI = 0.522–0.800). When compared to the final combined results (Table 5), the automatic search method was found to have: (i) a sensitivity of 95% (only 1 relevant app out of 21 was missed by the automatic search), (ii) a specificity of 99% (considering a total of 1593 apps automatically screened – 991 available in Google Play and 941 in Apple App), (iii) a positive predictive value of 51% (i.e., 51% of the apps identified by the automatic search programme were considered relevant), and (iv) a negative predictive value >99%. When taking only “fully relevant” apps into account, a sensitivity of 100% was observed, along with a specificity of 99%, a positive predictive value of 41%, and a negative predictive value of 100%.

3.4 | Characteristics of potentially relevant apps for AR self-management

Seventeen apps were available on both the Apple App and Google Play stores (Table 4). Two apps were available only on Google Play, and two were available only on the Apple App store. Only one of the apps offered a paid version with additional features. Three apps contained advertising, and four offered in-app purchases.

The US and Australia were the two countries where most of the relevant apps were developed (8 and 5 respectively), followed by Italy (with 2). One app was developed in Austria, Belgium, Denmark, England, France, the Netherlands and Singapore. In terms of the type of organization behind the relevant apps, 15 were developed by private for-profit companies, 5 by Universities, and 1 by a non-profit company.

On the Google Play store, five of the apps had been downloaded between 100,000 and 500,000 times, one between 50,000 and 100,000 times and five between 10,000 and 50,000. The average score of relevant apps was 3.6 and 3.5 (out of 5 stars) for Google Play and Apple App stores, respectively, with a standard deviation of 1.0 and 1.4 stars, respectively. Regarding the number of ratings, on the Google Play store, one app had more than 7000, while five had between 1000 and 2500. On the Apple App store, the three most popular apps had between 1500 and 4000 ratings.

3.5 | Selection of apps for the design of a symptom-medication score

Of the 21 apps retrieved, 16 were in English only. One was supposed to be also in French, but the translation was inappropriate as indicated by a comment by one of the app’s users in the app store and assessed by JB. Four apps were multilingual. One app (Galenus) was apparently not available in France even though a French language version was claimed to exist, and had been used by 500 users only. It is currently available in Austria, Germany and the UK. The Pollen app (Austria) uses medications by classes and has no published methodology. Finally, only the MASK-air® app and AllergyMonitor were found to have all the required features (Table 6).

4 | DISCUSSION

4.1 | Main findings

There was no automatic method of searching for mobile apps in AR. Manual screening has never been systematic. We performed a market search on AR apps, applying an automatic screening programme. In this study, we identified 21 apps potentially relevant for AR management. We combined the automatic and manual search methods to identify the potentially relevant apps. For the automatic search, we developed an app screening programme using JavaScript. By using pre-defined criteria, the programme successfully selected the apps considered as relevant (with a sensitivity of 95% and a specificity of 99%). Only one app retrieved by the manual search was missing in the automatic search.

Using an automatic screening approach, it appears possible to substantially reduce the burden of screening the app stores when performing a systematic review (in this study, we observed a 97% reduction of screening time when using the automatic method), without compromising the comprehensiveness of results. The
<table>
<thead>
<tr>
<th>App title</th>
<th>Relevant</th>
<th>Identified in automatic search</th>
<th>Identified in manual search</th>
<th>App store(s) with availability</th>
<th>N installs in Google Play</th>
<th>Score in Google Play</th>
<th>Score in Apple app store</th>
<th>N ratings in Google Play</th>
<th>N ratings in Apple</th>
<th>Associated publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>My Pollen Forecast - Allergies (free and Pro)</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple</td>
<td>–</td>
<td>–</td>
<td>4.6</td>
<td>–</td>
<td>–</td>
<td>26,034</td>
</tr>
<tr>
<td>klarify - Pollen Count &amp; Forecasts for Hayfever US</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>100,000+</td>
<td>4.0</td>
<td>4.6</td>
<td>2068</td>
<td>531</td>
<td></td>
</tr>
<tr>
<td>Allergy Alert by Pollen.com</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>100,000+</td>
<td>3.2</td>
<td>2.8</td>
<td>1366</td>
<td>92</td>
<td></td>
</tr>
<tr>
<td>WebMD Allergy</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>100,000+</td>
<td>4.2</td>
<td>4.5</td>
<td>7274</td>
<td>3760</td>
<td></td>
</tr>
<tr>
<td>ZYRTEC® AllergyCast®</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>100,000+</td>
<td>2.8</td>
<td>3.3</td>
<td>2325</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Pollen (Austria)</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>100,000+</td>
<td>3.5</td>
<td>ND</td>
<td>1279</td>
<td>3</td>
<td>49, 50</td>
</tr>
<tr>
<td>Correlate - Health Diary and Life Journal</td>
<td>Possibly</td>
<td>+</td>
<td></td>
<td>Google</td>
<td>50,000+</td>
<td>4.1</td>
<td>–</td>
<td>1632</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Pollen Wise - What’s in your air, when and where</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>10,000+</td>
<td>4.0</td>
<td>4.9</td>
<td>110</td>
<td>390</td>
<td></td>
</tr>
<tr>
<td>Sensio Air, Pollen &amp; Pollution, allergy tracker</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>10,000+</td>
<td>3.3</td>
<td>4.0</td>
<td>120</td>
<td>152</td>
<td></td>
</tr>
<tr>
<td>Melbourne Pollen Count</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>10,000+</td>
<td>4.5</td>
<td>4.6</td>
<td>543</td>
<td>1678</td>
<td></td>
</tr>
<tr>
<td>MASK-air</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>10,000+</td>
<td>3.9</td>
<td>ND</td>
<td>127</td>
<td>2</td>
<td>6, 8, 22, 25–44</td>
</tr>
<tr>
<td>AirRater</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>10,000+</td>
<td>4.2</td>
<td>4.4</td>
<td>211</td>
<td>512</td>
<td>51</td>
</tr>
<tr>
<td>Hay Fever Relief</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>5000+</td>
<td>4.1</td>
<td>ND</td>
<td>0</td>
<td>2</td>
<td>45–50</td>
</tr>
<tr>
<td>AllergyMonitor</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>5000+</td>
<td>ND</td>
<td>ND</td>
<td>117</td>
<td>0</td>
<td>45–50</td>
</tr>
<tr>
<td>Canberra Pollen Count</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>1000+</td>
<td>3.8</td>
<td>4.7</td>
<td>53</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Sydney Pollen Count</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>1000+</td>
<td>3.3</td>
<td>4.7</td>
<td>20</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td>Pattern - Correlate, Health Diary, Mood-Tracker</td>
<td>Possibly</td>
<td>+</td>
<td></td>
<td>Google</td>
<td>1000+</td>
<td>3.8</td>
<td>–</td>
<td>34</td>
<td>–</td>
<td></td>
</tr>
<tr>
<td>Galenus Health</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>500+</td>
<td>ND</td>
<td>ND</td>
<td>7</td>
<td>6</td>
<td>–</td>
</tr>
<tr>
<td>KagenAir</td>
<td>Fully</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>500+</td>
<td>ND</td>
<td>ND</td>
<td>5</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>Britain Breathing</td>
<td>Possibly</td>
<td>+</td>
<td>+</td>
<td>Apple, Google</td>
<td>50+</td>
<td>ND</td>
<td>–</td>
<td>0</td>
<td>2</td>
<td>–</td>
</tr>
<tr>
<td>DailyBreath</td>
<td>Possibly</td>
<td>+</td>
<td></td>
<td>Apple</td>
<td>–</td>
<td>–</td>
<td>3.5</td>
<td>–</td>
<td>11</td>
<td>–</td>
</tr>
</tbody>
</table>

*a N ratings and score in Apple app store are averaged over the considered countries.

bWhen there were <10 evaluations, ranking was not provided (ND).

cNo publication for AR.
### Table 5: Automatic search performance metrics compared to manual search

<table>
<thead>
<tr>
<th></th>
<th>Relevant</th>
<th>Not relevant</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manual search method</td>
<td>20</td>
<td>1</td>
<td>21</td>
</tr>
<tr>
<td>Not relevant</td>
<td>19</td>
<td>1553</td>
<td>1568</td>
</tr>
<tr>
<td>Total</td>
<td>39</td>
<td>1554</td>
<td>1593</td>
</tr>
</tbody>
</table>

- Sensitivity: $\frac{20}{21} = 95\%$
- Specificity: $\frac{1553}{1568} = 99\%$
- Positive predictive value: $\frac{20}{39} = 51\%$
- Negative predictive value: $\frac{1554}{1554} = 99\%$

### Table 6: List of apps that can be used for the design of a symptom-medication score

<table>
<thead>
<tr>
<th></th>
<th>Allergy monitor</th>
<th>Galenus</th>
<th>MASK-air</th>
<th>Pollen (Austria)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Countries</td>
<td>14</td>
<td>3</td>
<td>30</td>
<td>7</td>
</tr>
<tr>
<td>Last update</td>
<td>2018</td>
<td>Jan 2021</td>
<td>Nov 2020</td>
<td>Nov 2020</td>
</tr>
<tr>
<td>Published validation</td>
<td>45–50</td>
<td>For rhinosinusitis only, not for rhinitis 56, 57</td>
<td>22, 25–28, 30–33, 39, 40, 58–62</td>
<td>51, 52</td>
</tr>
<tr>
<td>List of all medications</td>
<td>YES, by medication and dosage, customized by country</td>
<td>YES</td>
<td>YES, by medication, customized by country</td>
<td>YES, by drug class</td>
</tr>
</tbody>
</table>
number of potentially relevant missed apps was in fact higher with the manual search approach, possibly because app titles and icons are not always indicative of the apps’ contents and features. This screening method is of importance since many apps were listed but not methodologically screened, due to the lack of a harmonized and standardized tool. This has been emphasized as an unmet need by the European Academy of Allergy and Clinical Immunology. An important advantage of the automatic process is its reliability and the prevention of errors (that can occur in the manual search).

As the use and availability of mobile health apps have increased, a thorough, accessible framework is needed for app evaluation. The growing importance of mHealth requires the use of validated apps. However, the six most popular AR apps (>100,000 downloads) are not necessarily associated with published validation or scientific publications. Publications associated with potentially relevant apps were found in PubMed, and concerned four apps: MASK-air (23 papers retrieved), AllergyMonitor, Pollen, and Air Rater. This represents a problem, since many patients do not understand the needs and benefits of mHealth, and many worry about data privacy. On the other hand, too many patients over-rely on internet-based information and on untested mHealth solutions. This attitude may have dangerous implications since patients may receive an incorrect diagnosis or management strategy. More in-depth analysis, such as the American Psychiatric Association’s app evaluation model, may be needed as a way to critically assess an app by considering accessibility, privacy and security, clinical foundation, engagement and interoperability. However, since the number of reliable apps is low in rhinitis, this strategy may be developed when more apps become available. Another aspect of relevance is the customer perspective. Our study did not retrieve any apps, althoughMASK-air has been assessed twice for this aspect.

Such an automatic screening tool has multiple applications. One of them is the screening of apps that can be used to develop a symptom-medication score for AR. An ARIA-EAAcI Task Force has developed this tool. The first step was to screen the apps that can be used. The current automatic search made it possible to identify eligible apps and to immediately discard those that did not possess the necessary features, such as a sufficient number of users to perform analyses or apps available only in English (since a multilingual app is necessary).

The practical utility of the automatic search also extends to the clinical practice. In fact, in a time when mobile health tools are ever more clinically relevant, the process described in this study can be adopted as a first step for those interested in thoroughly assessing the quality of apps. Having immediate access to a list of relevant health apps is helpful to both patients and doctors since it allows them to identify available alternatives for each disease, assess the quality of the apps and select those with the highest scientific quality. It is also of great importance in research and in grant applications since there is a need to assess the competition.

This study is likely to impact guidelines. mHealth will represent an essential component of AR guidelines, and ARIA 2020 guidelines already combine mHealth real-world data with evidence-based medicine. Guideline developers need to know the full landscape of mHealth that can be provided by the automatic screening.

4.2 | Strengths and limitations

This study has several limitations: (i) A potential bias in the app selection is expected, due to the limited set of countries chosen, and only English language apps have been considered. However, this does not compromise the accuracy of the automatic method (only the comprehensiveness of the listed results). (ii) The search terms selected, and the limitation to three categories, may also contribute to the potential selection bias. (iii) Since Apple and Google do not provide support to read the data automatically from their app stores, automatic app screening has to resort to data scraping algorithms. These may stop working when the app store owners (Apple and Google in this case) update certain parts of their source code. (iv) Few qualitative studies were performed, including a systematic review on AR self-management apps, and a review for asthma self-management apps in the US. The qualitative review included an individual assessment of each app, with its functions and quality measured, based on a framework such as MARS.

This study also presents important strengths: (i) The app stores of multiple countries can be simultaneously screened, allowing a broader analysis. (ii) We combined automatic and manual methods. (iii) The app list generated by the search programme can be regularly updated to always reflect the current state of the app stores, and the programme may run automatically. (iv) The app screening programme aggregates the values of the variables “country,” “keywords” and “source” into one single value for each app. The final value is the total set of values obtained during all the iterations (the programme iterates over the considered countries, keywords and sources). This resulting data structure is easy to save. It overcomes the impossibility to export results to a dataset of the manual app search. Merging the results automatically into a single data structure significantly reduces the manual time spent analysing and presenting the results.

There are potential improvements. Even though the set of text analysis conditions developed for this study has achieved a good specificity/sensitivity trade-off, it can be improved by further refining the condition set. Another potential improvement is the use of additional data along with the app’s title and description, such as the metadata inside the app’s source code. The automatic screening of apps addressing conditions other than AR (based on the same methods), and even developing a front-end allowing physicians and patients to search for relevant apps for potentially any disease, will be a topic for future studies, taking into account the ever-growing importance of mobile apps for the management of chronic diseases. Finally, the impact on the results of a larger set of keywords in the search algorithm could be studied.
Future research should include a qualitative review of the apps and multimorbidity (e.g., asthma) as well as a quality label that was beyond the scope of the review.

5 | CONCLUSION

In this study, we identified 21 apps for AR self-management, with six having more than 100,000 downloads. Clinicians should be aware of the role that such apps may play in helping patients control their disease. While mHealth apps may be a helpful tool for devising patient-centred management strategies, one should note that only a minority of available apps have associated scientific evidence. For identifying potentially relevant apps, we have developed a programme which is able to screen app stores, and help identify relevant apps on a certain topic. We have applied this programme using a set of pre-defined conditions to identify relevant apps on AR, and have shown that sensitivity and specificity are high when compared to manual search. Taking into account the role that mHealth has been assuming in chronic diseases, we believe that this type of solution may be even more relevant, and will help clinicians and patients identify apps that can support the management of their disease.

ACKNOWLEDGMENTS

None.

CONFLICT OF INTERESTS

S. B.-A. reports grants from TEVA, personal fees from TEVA, AstraZeneca, Boehringer Ingelheim, GSK, Sanofi, Mylan.

J. B. reports personal fees from Chiesi, Cipla, Hikma, Menarini, Mundipharma, Mylan, Novartis, Sanofi-Aventis, Takeda, Teva, Uriach, other from KYomed-Innov, personal fees from Purina.

J. A. F. reports personal fees from AstraZeneca, Mundipharma, and is a cofounder of a SME involved in mHealth.

L. K. reports personal grants and fees from Allergopharma, MEDA/Mylan, LETI Pharma, Sanofi, personal fees from HAL Allergie, Allergy Therapeut, Cassella med, grants from ALK Abelô, Quintiles, ASIT biotech, Lofarma, AstraZeneca, GSK, Immunotek and Membership: AeDA, DGHNO, Deutsche Akademie für Allergologie und klinische Immunologie, HNO-BV, GPA, EAACI.

O. P. reports grants and personal fees from ALK-Abelô, Allergopharma, Stallergenes Greer, HAL Allergy Holding B.V./HAL Allergie GmbH, Bencard Allergie GmbH/Allergy Therapeutics, Lofarma, Anergis S.A., ASIT Biotech Tools S.A., Laboratorios LETI/LETI Pharma, Glaxo Smith Kline; grants from Biomay, Circassia, Pohl-Boskamp, Immunotek S.L., personal fees from MEDA Pharma/MYLAN, Mobile Chamber Experts (a GA2LEN Partner), Indoor Biotechnologies, Astellas Pharma Global, EUFOREA, ROXALL Medizin, Novartis, Sanofi-Aventis and Sanofi-Genzyme, Med Update Europe GmbH, streamedup! GmbH, John Wiley and Sons, AS, Paul-Martini-Stiftung (PMS), ST is cofounder of TPS Production that has developed AllergyMonitor.

AUTHOR CONTRIBUTIONS

AA conceived the paper, performed the JavaScript, and wrote the paper. Bernardo Sousa-Pinto discussed with AA the concept of the paper and the methods. Josep M. Antó and Jean Bousquet proposed the study, revised the concept of the paper, and wrote the paper. All authors analysed the study and reviewed the paper.

IRB APPROVAL

Not needed.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Aram Antó https://orcid.org/0000-0001-8180-5481
Bernardo Sousa-Pinto https://orcid.org/0000-0002-1277-3401
Oliver Pfaar https://orcid.org/0000-0003-4374-9639
Paolo Matricardi https://orcid.org/0000-0001-5485-0324

REFERENCES


**How to cite this article:** Antó A, Sousa-Pinto B, Czarlewski W, et al. Automatic market research of mobile health apps for the self-management of allergic rhinitis. *Clin Exp Allergy*. 2022;00:1-13. doi:10.1111/cea.14135