Evaluation of generic medical information accessed via mobile phones at the point of care in resource-limited settings

Hayley Goldbach,1,2 Aileen Y Chang,2,3 Andrea Kyer,4 Dineo Ketshogileng,5 Lynne Taylor,1,6 Amit Chandra,7 Matthew Dacso,8 Shiang-Ju Kung,7,9 Taatske Rijken,7 Paul Fontelo,10 Ryan Littman-Quinn,2 Anne K Seymour,2,4 Carrie L Kovarik1,2,3

ABSTRACT

Objective Many mobile phone resources have been developed to increase access to health education in the developing world, yet few studies have compared these resources or quantified their performance in a resource-limited setting. This study aims to compare the performance of resident physicians in answering clinical scenarios using PubMed abstracts accessed via the PubMed for Handhelds (PubMed4Hh) website versus medical/drug reference applications (Medical Apps) accessed via software on the mobile phone.

Methods A two-arm comparative study with crossover design was conducted. Subjects, who were resident physicians at the University of Botswana, completed eight scenarios, each with multi-part questions. The primary outcome was a grade for each question. The primary independent variable was the intervention arm and other independent variables included residency and question.

Results Within each question type there were significant differences in ‘percentage correct’ between Medical Apps and PubMed4Hh for three of the six types of questions: drug-related, diagnosis/definitions, and treatment/management. Within each of these question types, Medical Apps had a higher percentage of fully correct responses than PubMed4Hh (63% vs 13%, 33% vs 12%, and 41% vs 13%, respectively). PubMed4Hh performed better for epidemiologic questions.

Conclusions While mobile access to primary literature remains important and serves an information niche, mobile applications with condensed content may be more appropriate for point-of-care information needs. Further research is required to examine the specific information needs of clinicians in resource-limited settings and to evaluate the appropriateness of current resources in bridging location- and context-specific information gaps.

BACKGROUND

As mobile technology has become touted as a democratizing tool in the quest to improve access to healthcare worldwide, it is important to evaluate which tools may be beneficial to clinicians in resource-limited settings. Innovative solutions have emerged to improve information access, patient compliance, and physician/patient communication.1 2 One of the most widely adopted uses of mobile technology is point-of-care access to health information.

Swift access to relevant updated information is a critical part of healthcare delivery.3–4 What types of information do clinicians need and how can they be addressed by mobile technology? Lancaster and Warner identified three general types of information needs: (1) background information on a topic; (2) information to ‘keep up’ with new advances in a given subject area; and (3) to help [solve] a certain problem or [make] a decision.5 Prior to the internet age, this information could be found using condensed pocket texts, accessing books and journals at libraries, or consultation with colleagues. However, in resource-limited settings, these avenues are often not available6 as libraries are not well-stocked and specialist access is very limited. Theoretically, the advent of the internet should have democratized this information but, in reality, the internet has failed to close this information gap7 and has exacerbated the disparity, creating what has been referred to as the international digital divide.8 The creation of this digital divide has been multifactorial in nature. Infrastructure is perhaps the most striking barrier as the price of computers and internet service remains prohibitively high in developing countries.9 Reliable incoming bandwidth is also an issue, and speeds can vary within the same region. For example, Botswana’s incoming bandwidth is 14 megabits/s,10 but only a few African countries have a better bandwidth than this.11 Its neighbor, South Africa, is reported to have about 80 times more bandwidth for data communications than Botswana.12 Lack of familiarity with computer technology and lack of trained librarians or information specialists further exacerbate the problem. In developing countries where access is not a limiting factor, providers are experiencing the information paradox of the developed world—namely, being inundated with mostly irrelevant information and unable to find answers to the clinical questions at hand.8

Mobile technology has been identified as a potential vehicle for increasing access to healthcare information and narrowing the digital divide.13–15 Basic mobile phones are already widely used in the developing world,13–15 and the cost of smartphones has been declining. Smartphones represent a less costly alternative to computers as they do not require expensive fixed broadband internet connections. Mobile health education (mHealthEd), which overlaps mobile health (mHealth) and mobile learning (mLearning), harnesses mobile technology...
to broaden global access to healthcare information. The first International Mobile Technology for Education and Development (m4Ed4Dev) symposium was held in August 2011,\textsuperscript{16} popular media has covered the subject extensively,\textsuperscript{17–20} and mLearning trials with smartphones and other handheld devices have been reported among physicians and healthcare workers in Peru,\textsuperscript{21} Tanzania,\textsuperscript{22} Kenya,\textsuperscript{23} and others. In both resource-limited and resource-rich countries, current research in this area has largely relied on descriptive analysis of mLearning technology. Most researchers have used usage data or satisfaction surveys as outcome measures.\textsuperscript{24–29} Even large-scale randomized controlled trials have largely relied on self-report as an outcome measure.\textsuperscript{30} Fewer studies have attempted to quantify the performance of mobile technology in clinical decision-making. There are also relatively few data comparing tools within the umbrella of mobile technology, especially in resource-limited settings. For example, smartphones can be loaded with many different applications, some of which might be more or less helpful in facilitating the dissemination of healthcare information and assisting in point-of-care decisions.\textsuperscript{28}

Moreover, applications which may be useful to physicians in the USA or Europe may not be useful for physicians practicing in other countries.

With the expanding emphasis on evidence-based care, narrowing the digital divide and providing improved access to information for providers in developing countries is becoming increasingly important. The Institute of Medicine has named ‘using information technology’ and ‘practicing evidenced-based medicine’ as two of their five core competencies for health professional education.\textsuperscript{31} The advent of information resources on a mobile phone has created excitement over the prospect of clinicians in resource-limited settings further incorporating evidence-based medicine into their practice, even without access to a reliable computer. Some have suggested that, for clinicians hoping to practice evidence-based medicine, a search of PubMed abstracts may be fruitful even within a limited timeframe.\textsuperscript{32,33}

As such, comparisons of mobile resources are increasingly important as the availability of mobile resources continues to grow rapidly. This study compares the performance of University of Botswana resident physicians answering clinical scenario prompts using PubMed abstracts, accessed via the PubMed for Handhelds (PubMed4Hh) website, with the use of medical/drug reference applications (Medical Apps) accessed via locally loaded software on the mobile phone.

**MATERIALS AND METHODS**

A two-arm comparative study with a crossover design was conducted to compare the use of PubMed4Hh with Medical Apps in answering clinical scenarios. Both of these clinical decision support interventions (PubMed4Hh and Medical Apps) were accessed via smartphones.

**Study subjects and phone**

Subjects were first-year residents (physicians in postgraduate training) in internal medicine, pediatrics, emergency medicine, and family medicine programs at the University of Botswana School of Medicine. At the time of the study, residents had been using the phone for 3 months. Residents were required to participate in this study in order to receive the phone. If residents did not participate in the study, they were required to return the phone.

The smartphones used were myTouch 3G Slide HTC Android phones. Data-enabled subscriber identification module cards allowed access to mobile internet and thus the web and email. Locally loaded Medical Apps included Medscape, Unbound Medicine (which includes 5-Minute Clinical Consult, 5-Minute Pediatric Clinical Consult, A to Z, Drug Facts, Clinical Evidence 2e, Cochrane Abstracts, Communicable Diseases, Drug Interaction Facts, Emergency Medicine Manual, Evidence-Based Medicine Guidelines, EE+POEM Archive/EE+POEM Daily, The Merck Manual, Red Book, Review of Natural Products, Taber’s, 21st Edition), Skyscape (which includes MedAlert, Archimedes, Dynamed, Outlines in Clinical Medicine, Rx Drugs), and ePocrates Rx. Medscape is an application with both disease and drug information. UCentral and Skyscape are medical software, each containing multiple applications including drug references and disease references. ePocrates Rx is a stand-alone drug reference application. Medscape and ePocrates are available through free subscriptions. UCentral and Skyscape both require paid subscriptions, which were donated by their respective companies. Based on the available resources, the study team aimed to provide the subjects with a wide variety of applications that were meant to reflect an array of resources readily available on the market at the time of the study design. At the time of phone distribution, a training session was conducted in which usage of the phone and its information resources, including its Medical Apps and the PubMed4Hh website (http://pubmedhh.nlm.nih.gov or go.usa.gov/xFb), were discussed. Of note, there exists a PubMed4Hh application which provides users with several options for searching and viewing abstracts from PubMed. The PubMed4Hh website, not the application, was used in this study. PubMed4Hh requires mobile internet connectivity. Participants connected through the 3G or Edge network, depending on their location in the country. The Medical Apps were locally loaded on the phone and did not require connectivity for usage during the study. Once the Medical Apps and their content have been downloaded onto the device, they are functional without internet connectivity although multimedia features and software updates require connectivity. All software was updated prior to conducting this study.

**Study design**

Residents completed eight clinical scenarios generated by residency program directors, each with multi-part questions related to the scenario (see online supplementary appendix). Within each training program, residents were randomized by order of presentation to the study room to use either PubMed4Hh or Medical Apps to answer the first four clinical scenarios. They then switched arms to answer the last four clinical scenarios. Residents were asked to defer using existing knowledge and were instructed to answer the scenario questions using only the available resources within their arm, regardless of prior knowledge.

They were also asked to list either the PubMed article identification number or the medical/drug information application used. No resident was asked the same question twice. All questions were answered at least once in both arms. Within each residency program, each resident received a unique version of the clinical scenarios form (same questions, different order). Residents were given a 5 min time limit to answer each question, which was deemed a reasonable and appropriate amount of time to locate information at the point of care.

The primary outcome was a grade for each question: fully incorrect/blank=0, partially correct=1, or fully correct=2. Answers were de-identified. The graders were the physicians who created the questions, and grades were assigned based on content and verification that the written answer exists in the...
listed PubMed abstract or medical/drug information application. If the content of an answer was graded as fully correct or partially incorrect, but not available in the listed resource, a grade of fully incorrect was assigned. The primary independent variable was intervention arm (PubMed4Hh or Medical Apps). Other independent variables included residency type (family medicine, internal medicine, emergency medicine, or pediatrics) and question type (drug-related, diagnosis/definrivions, treatment/management, pathophysiology, epidemiology, or prevention).

Sample size
Nineteen people participated in this study, each answering 14–25 questions. This resulted in 423 person-question observations. One person was excluded for not answering at least three questions in both intervention arms, which then resulted in 18 final participants (table 1) and 409 person-question observations.

Clinical scenarios were created by the different residency program directors working in Botswana and were intended to represent clinical queries that they felt were representative of the information needs of resident physicians in their respective field. Each scenario contained multi-part questions and varied in the number of questions. Thus, participants from different residency programs answered a different number of questions.

Statistical analysis
Graphics and univariate analyses were conducted to examine the primary outcome of grade and the distribution of study variables (intervention arm, resident type, and question type). A χ² test was used to assess equal distribution of residency type and question type within the intervention arms.

In assessing the association between grade and intervention arm, χ² tests, Mann–Whitney–Wilcoxon tests, analysis of variance (ANOVA) using general linear models (GLM), and mixed models were conducted. Using person-question as the unit of analysis (n=409), we performed several analyses treating the primary outcome of grade as a categorical, ordinal, and continuous variable. Treating grade as categorical, cross-tabulation, χ² and Cochran–Mantel–Haenszel statistics of general association were used to assess the association between ‘percent correct’ and intervention arm. Treating grade as ordinal, the Mann–Whitney–Wilcoxon test was used to assess the association between ‘levels of correctness’ and intervention arm. Treating grade as continuous, ANOVA (using GLM) was used to assess the relationship between ‘correctness’ and intervention arm.

These relationships were further examined using person as the unit of analysis (n=18), treating grade as continuous. Mixed models were used to examine the relationship between grade and intervention arm, taking into account the dependency or ‘clusters of observations’ within persons. The person clusters were handled by adding a subject random intercept component to the model and specifying compound symmetry as the working correlation structure.

In all analyses the association between grade and arm was examined, both overall and by residency type and question type. Statistical analyses were performed using SAS V9.3.

This study was approved by the Institutional Review Board at the University of Pennsylvania in the USA and the Ministry of Health in Botswana.

RESULTS
Grade differs by intervention arm
χ² results revealed statistically significant differences in the percentage of correct responses between the PubMed4Hh arm and Medical Apps arm (χ²=41.27, p<0.0001). Overall, Medical Apps had a higher percentage of fully correct responses than PubMed4Hh (36% vs 14%, respectively; figure 1A,B). When

Table 1 General characteristics of the 18 participants included in the analysis

<table>
<thead>
<tr>
<th>Age</th>
<th>Sex</th>
<th>Residency</th>
<th>Years since MD</th>
</tr>
</thead>
<tbody>
<tr>
<td>33</td>
<td>M</td>
<td>EM</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>M</td>
<td>EM</td>
<td>4</td>
</tr>
<tr>
<td>27</td>
<td>M</td>
<td>EM</td>
<td>2.5</td>
</tr>
<tr>
<td>27</td>
<td>F</td>
<td>EM</td>
<td>2</td>
</tr>
<tr>
<td>42</td>
<td>M</td>
<td>FM</td>
<td>12</td>
</tr>
<tr>
<td>33</td>
<td>M</td>
<td>FM</td>
<td>5</td>
</tr>
<tr>
<td>30</td>
<td>F</td>
<td>FM</td>
<td>4</td>
</tr>
<tr>
<td>30</td>
<td>M</td>
<td>FM</td>
<td>3</td>
</tr>
<tr>
<td>29</td>
<td>M</td>
<td>FM</td>
<td>2.7</td>
</tr>
<tr>
<td>29</td>
<td>M</td>
<td>FM</td>
<td>3</td>
</tr>
<tr>
<td>28</td>
<td>F</td>
<td>FM</td>
<td>2</td>
</tr>
<tr>
<td>No data</td>
<td>F</td>
<td>IM</td>
<td>No data</td>
</tr>
<tr>
<td>31</td>
<td>M</td>
<td>IM</td>
<td>4</td>
</tr>
<tr>
<td>30</td>
<td>M</td>
<td>IM</td>
<td>4</td>
</tr>
<tr>
<td>31</td>
<td>F</td>
<td>Peds</td>
<td>5</td>
</tr>
<tr>
<td>31</td>
<td>M</td>
<td>Peds</td>
<td>5.5</td>
</tr>
<tr>
<td>30</td>
<td>M</td>
<td>Peds</td>
<td>3</td>
</tr>
<tr>
<td>28</td>
<td>F</td>
<td>Peds</td>
<td>2</td>
</tr>
</tbody>
</table>

Age (years), sex, residency program, and years since obtaining medical degree (MD) are reported. EM, emergency medicine; FM, family medicine; IM, internal medicine; Peds, pediatrics.

Figure 1 Percentage breakdown of grades (fully correct, partially correct, incorrect) for subjects using (A) PubMed4Hh and (B) Medical Apps.
grade was analyzed as an ordinal variable, the Mann–Whitney–Wilcoxon test also demonstrated a statistically significant difference between Medical Apps (median grade=1, partially correct) and PubMed4Hh (median grade=0, fully incorrect) (p<0.0001). When grade was analyzed as a continuous variable, the ANOVA (using GLM) tests also demonstrated a statistically significant mean grade difference between Medical Apps and PubMed4Hh (mean=1.0 vs mean=0.47; table 2). Using the mixed model which adjusts for clustering by looking at 18 unique people rather than 409 unique person-questions, there was also a statistically significant mean grade difference between Medical Apps and PubMed4Hh (p<0.0001).

There was no association between intervention arm and residency type or intervention arm and question type. Therefore, further statistical adjustment for effect modification by residency type or question type was not warranted.

**Difference in grade is similar across residency types**

Within each residency type, the Medical Apps arm had a significantly higher percentage of correct responses than the PubMed4Hh arm (family medicine: 24% vs 15%, p=0.0071; internal medicine: 27% vs 5%, p=0.05; emergency medicine: 52% vs 24%, p=0.0021; pediatrics: 46% vs 11%, p <0.0001). The Mann–Whitney–Wilcoxon test also demonstrated a significant difference in ’levels of correctness’ within each residency type. Similarly, ANOVA (using GLM) results showed that the magnitude of the difference between the intervention arms was similar across residency types. After adjusting for clusters of observations by using mixed models, the results similarly showed non-significant interaction effects between arm and residency type.

**Difference in grade varies based on question type**

Within each question type there were significant differences in ‘percentage correct’ between Medical Apps and PubMed4Hh for three of the six types of questions: drug-related, diagnosis/definitions, and treatment/management. Within each of these question types, Medical Apps had a higher percentage of fully correct responses than PubMed4Hh (63% vs 13%, 33% vs 12%, and 41% vs 13%, respectively). When analyzing grade as an ordinal variable, the Mann–Whitney test also demonstrated statistically significant arm differences in the level of correctness for these three question types.

The ANOVA (using GLM) results showed that the magnitude of the difference between the intervention arms varied by question type. Medical Apps had higher means than PubMed4Hh for all question types except epidemiology (0.17 vs 0.59; table 2). The mixed model results revealed significant interaction effects between arm and question type (p=0.01), which similarly suggests that the magnitude of the grade difference between the intervention arms varies by question type. Limited sample size prohibited post hoc comparisons.

**DISCUSSION**

This study demonstrates that the performance of medical/drug information applications is superior to PubMed abstracts for answering questions related to clinical scenarios. Residents received higher mean grades using Medical Apps than with PubMed4Hh. This was true across all residency groups and for all question types except epidemiologic questions, for which PubMed4Hh yielded higher grades.

For those who have used a mobile phone, these results may not be surprising. Medical Apps are specifically created to be user-friendly. They create a general aggregate of the ‘prevailing knowledge’ about a given topic and present that information in an easily digestible format. This type of format works well for point-of-care decisions about topics such as drug dosing and conventional treatment regimens. This study primarily tested access to information that would be considered as ‘background’ using Lancaster and Warner’s classification. While many questions were related to management, which is ‘decision-making,’ these queries were restricted to standard of care information that could be accessed using applications.

Subjects did not perform as well when using PubMed abstracts as the sole source of information. Although the link to the PubMed4Hh website appeared on the phone’s home screen as an icon, like an ‘application’, it is really a distinct entity. PubMed functions as a search engine of primary literature and does not aggregate or consolide information, nor does it have editorial function to grade the evidence or assimilate it into the existing body of knowledge. As a point-of-care information tool, it has several disadvantages. Primary literature needs to be approached with a critical eye, which is a challenging task when using PubMed4Hh. While the format of the PubMed4Hh is easy to use (enter in keywords and/or author, journal title, or article title), it returns a mix of articles that range from review

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Mean scores (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PubMed for handhelds (Abstracts)</td>
</tr>
<tr>
<td>Overall</td>
<td>0.47 (0.37 to 0.57)</td>
</tr>
<tr>
<td>Residency type</td>
<td></td>
</tr>
<tr>
<td>Family medicine</td>
<td>0.47 (0.31 to 0.63)</td>
</tr>
<tr>
<td>Internal medicine</td>
<td>0.32 (0.13 to 0.52)</td>
</tr>
<tr>
<td>Emergency medicine</td>
<td>0.60 (0.33 to 0.86)</td>
</tr>
<tr>
<td>Pediatrics</td>
<td>0.47 (0.27 to 0.67)</td>
</tr>
<tr>
<td>Question type</td>
<td></td>
</tr>
<tr>
<td>Drug-related</td>
<td>0.38 (0* to 1.0)</td>
</tr>
<tr>
<td>Diagnosis/definitions</td>
<td>0.48 (0.31 to 0.65)</td>
</tr>
<tr>
<td>Treatment/management</td>
<td>0.44 (0.30 to 0.58)</td>
</tr>
<tr>
<td>Pathophysiology</td>
<td>0.27 (0* to 0.59)</td>
</tr>
<tr>
<td>Epidemiology</td>
<td>0.59 (0.19 to 0.99)</td>
</tr>
<tr>
<td>Prevention</td>
<td>1.0 (0* to 2.2)</td>
</tr>
</tbody>
</table>

Answers to clinical scenarios were graded as 0 (fully incorrect or blank), 1 (partially correct), or 2 (fully correct).

*Actual calculated CI extended into negative range.
articles to pilot studies of experimental treatments to genetic studies with minimal clinical relevance. These results are ranked by reverse order added instead of relevance. In addition, the sheer volume of search results may be overwhelming to a new user. Also, study subjects only had access to abstracts and not to full articles, unless access was free. Furthermore, access to PubMed4Hh requires internet connectivity whereas the Medical Apps used in this study did not require connectivity—and the ability to function without a reliable internet connection should not be understated, especially in a resource-limited context.

Undeniably, PubMed4Hh offers a tremendous opportunity to interface with evidence-based information, but it requires the clinician to be a first-line consumer of primary literature. At the point of care it is unrealistic to expect a clinician to have the time to perform a critical analysis. Indeed, several have theorized that meaningful interaction with primary literature takes days to weeks. These interactions can be facilitated best by specially trained third parties such as ‘informationists’ who can help clinicians digest an often overwhelming volume of primary information and integrate it into clinical practice.34 35 Given constraints on human resources, the incorporation of informationists is unlikely to be a realistic solution in developing countries. However, a basic tutorial led by informationists to introduce skills of primary literature appraisal is feasible and could be directed towards clinicians and librarians. Notably, certain medical information applications (eg, DynaMed) provide aggregated information that includes primary literature (which is hyperlinked and can be launched by a web browser), as well as treatment options that include evidence-based analyses.26 This reviewed content is periodically updated by DynaMed editors. It also contains a section entitled ‘Updates’ which includes recent PubMed articles that are relevant to the disease but have yet to be reviewed by the editorial team. This format allows the clinician more information about the sources of evidence but still maintains an editorial analysis of the information incorporated into its content.

As seamless as Medical Apps appear, they do sacrifice a degree of breadth and depth of information to achieve a streamlined interface. Consequent limitations are most apparent when using the search function, which often relies on the user knowing how to query the database. For example, Medscape can be searched only by disease type and not by symptom, treatment, or a combination. ‘Albuterol for bronchitis’ will return no results. The user must search ‘Bronchitis’ and then select the ‘Treatment and Management’ tab to see if the editors included albuterol in their summary of treatment recommendations. In general, when using a Medical App, one must search for a specific drug or disease whereas PubMed4Hh allows the user to create a more specific and personalized search such as ‘drug-resistant hypertension’ or ‘levetiracetam and retinal toxicity’. Similarly, while the edited content of Medical Apps allows for synthesis of information, the content has been selected by the editorial boards of the application’s creators. This point becomes very relevant when considering the use of a Medical App in an international resource-limited setting. Currently available applications are mostly created in the USA/Europe and are geared towards users in a similar setting. Thus, Medical Apps often recommend a specific standard of care which may not be possible or appropriate in Botswana. One can only expect that this problem would be exacerbated in countries with even more limited resources.

Even for disease-specific information, Medical Apps sometimes have a restricted scope, especially for diseases that are rarely encountered in developed countries. For example, Medscape has no entries for African tick bite fever whereas PubMed4Hh returns 117 articles (accessed July 2012). Medical Apps also cull data on epidemiology that is invariably USA/Europe-centric, with some entries including worldwide prevalence data but many only discussing the incidence/prevalence within the USA/Europe. Furthermore, not all applications are created equal, and no one application can truly serve as a ‘one-stop shop’. Some are better for drug-related information while others are better for management strategies, forcing the user to juggle several different platforms to find the most appropriate answer.

Quality assurance is another issue with Medical Apps. Unlike journal publications which have relatively transparent editing procedures and set standards for peer review, Medical Apps are not subject to the same extrinsic scrutiny. The lack of regulatory procedures is cause for concern, and ongoing efforts are being made to create certification procedures to establish minimum quality standards for Medical Apps.37 38

While PubMed4Hh did not function as well for point-of-care health information, it has an invaluable role to play in decision support. Returning to Lancaster and Warner’s delineation of information needs, access to primary literature certainly fulfills the need to ‘keep up with advances’.1 However, primary literature can also help with the second category of information need—namely, decision support. Many physicians report using primary literature, especially abstracts, for time-sensitive decision-making.32 33 The use of primary literature, despite the availability of easier to use pre-appraised resources, points to information gaps that may be filled by PubMed. This principle was illustrated in our study in which the performance of PubMed4Hh was superior only on questions related to epidemiology. This may be explained by the fact that these questions required the most specific information. For example, one question asked about the percentage of HIV-positive patients who experience immune reconstitution inflammatory syndrome. This type of question represents a request for relatively specific information, which is better suited to the more comprehensive search strategy of PubMed4Hh. However, one could argue that these questions were least representative of point-of-care information needs. Nonetheless, they demonstrate that PubMed4Hh can subserve a unique information niche that can be seen as complementary to the information niche served by Medical Apps, especially in a developing country where trained professionals such as informationists may not be readily available to assist clinicians with a thorough review of primary literature.

Limitations of the study
In this study the question types were sufficiently different to provoke different information needs. Data were collected on which Medical App was used for each question, but no distinction was made with regard to the specific Medical App within a ‘suite’ of Medical Apps. For example, MedAlert, Archimedes, Dynamed, Outlines in Clinical Medicine, Rx Drugs were all coded as Skyscape. Additionally, the relatively small sample size meant that the study was insufficiently powered to provide data regarding the performance of individual Medical Apps within each question type category.

As such, the manner of data collection and sample size prohibited analysis of which Medical App produced the highest score for each question type (ie, head-to-head comparison between Medical Apps). UCentral was the app with the most ‘uses’ (n=61), followed by Medscape (n=43), Skyscape (n=31), a combination (n=24), and ePocrates (n=1), but this usage data does not reflect the type of information sought by the study.

participants, nor does it reflect which aspect of the app or app ‘suite’ was used to generate an answer.

Further research could elucidate, for example, whether drug-related questions such as dosing and toxicities are better suited for ‘drug reference’ applications such as ePocrates or RxDrugs (part of the Skyscape suite). Furthermore, as stated above, PubMed4Hh was better suited for answering epidemiology questions than Medical Apps, perhaps hinting at an information need that was unmet by medical applications. Further research might elucidate which applications are superior for which types of information needs and may also be able to uncover information needs not adequately addressed by these applications.

CONCLUSIONS
In summary, the use of Medical Apps produced higher mean scores in graded clinical scenarios. Medical information applications provide broad stroke overviews of different diseases and treatment protocols which lend themselves well to point-of-care clinical information needs. PubMed4Hh, on the other hand, requires evaluation of primary literature which does not seem ideal for point-of-care decision-making. However, PubMed4Hh may be better suited for accessing specific clinical and epidemiologic information that cannot readily be found within Medical Apps.

Common sense dictates that a ‘one size fits all’ solution to the healthcare informatics divide is neither practical nor effective. Further research is required to examine the specific information needs of clinicians in resource-limited settings and to evaluate the appropriateness of current resources in bridging location- and context-specific information gaps.

Acknowledgements The authors thank the University of Botswana resident physicians in internal medicine, pediatrics, family medicine, and emergency medicine, Dr Gordana Cavic, Dr Andy Kessler, Dr Loeto Mazhani, Dr Luise Parsons, Dr Sunanda Ray, Orange Foundation in Botswana, EBSCO Publishing, and Unbound Medicine for their continued support. They also thank the Office of the Provost at the University of Pennsylvania for International Initiatives funding.

Funding This study was funded in part by the National Library of Medicine at the National Institutes of Health, USA.

Contributors HG, AYC, AK, DK, LT, AC, MD, S-JK, TR, PF, RL-G, AS, and CK each substantially contributed to (1) conception and design, acquisition of data, or analysis and interpretation of data; (2) drafting the article or revising it critically for important intellectual content; and (3) final approval of the version to be published.

Conflicts of interest None of the authors has any relevant conflicts of interest or financial disclosures.

Ethics approval Ethics approval was obtained from the University of Pennsylvania and Ministry of Health Botswana.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES
13 Kaplan WA. Can the ubiquitous power of mobile phones be used to improve health outcomes in developing countries? Global Health 2006;2:9.