7   Greenaway D. *Shape of training. Securing the future of excellent patient care.* London: Shape of Training, 2013.
8   British Medical Association. *Little need for more generalism.* London: BMA, 2013. Available online at http://bma.org.uk/news-views-analysis/news/2013/february/little-need-for-more-generalism [Accessed 5 November 2015].
9   Goddard A, Evans T, Phillips C. Medical Registrars in 2010: experience and expectations of future consultant physicians of the UK. *Clin Med* 2011;11:532–5.
10   Royal College of Physicians. *The medical registrar: empowering the unsung heroes of patient care.* London: RCP, 2013.
11   Federation of the Royal Colleges of Physicians of the United Kingdom. *Census of consultant physicians and medical registrars in the UK, 2013–14: data and commentary.* London: RCP, 2014.
12   Joint Royal College of Physicians Training Board. *Specialty training curriculum for general internal medicine.* London: JRCPTB, 2009 (updated 2012). Available online at www.jrcptb.org.uk/sites/default/files/2009%20GIM%20(amendment%202012).pdf. [Accessed 5 November 2015].

Address for correspondence: Dr A Goddard, Royal College of Physicians, 11 St Andrews Place, Regent's Park, London, NW1 4LE, UK.
Email: andrew.goddard@rcplondon.ac.uk

What makes a good clinical app? Introducing the RCP Health Informatics Unit checklist

Authors: Jeremy C Wyatt,[A] Harold Thimbleby,[B] Paul Rastall,[C] Jan Hoogewerf,[D] Darren Wooldridge[E] and John Williams[F]

ABSTRACT

Doctors increasingly rely on medical apps running on smart phones or tablet computers to support their work. However, these apps vary hugely in the quality of their data input screens, internal data processing, the methods used to handle sensitive patient data and how they communicate their output to the user. Inspired by Donabedian's approach to assessing quality and the principles of good user interface design, the Royal College of Physicians' Health Informatics Unit has developed and piloted an 18-item checklist to help clinicians assess the structure, functions and impact of medical apps. Use of this checklist should help clinicians to feel more confident about using medical apps themselves, about recommending them to their staff or prescribing them for patients.

KEYWORDS: Medical apps, mHealth, quality assessment checklist, Donabedian's structure, process, outcome, health informatics, clinical use of technology

Authors: [A]leadership chair in eHealth research, Leeds Institute of Health Sciences, Leeds, UK, and clinical advisor on new technologies, HIU, RCP, London, UK; [B]professor of computer science, Swansea University, Swansea, UK, and expert advisor on IT, HIU, RCP, London, UK; [C]clinical research fellow, Neath Port Talbot Hospital, Port Talbot, UK, and clinical fellow, HIU, RCP, London, UK; [D]programme manager, HIU, RCP, London, UK; [E]project manager, HIU, RCP, London, UK; [F]professor of health services research, Swansea University, Swansea, UK, and director, HIU, RCP, London, UK

Analysis and problem statement

Smart phone apps are potentially very useful additions to clinical practice and are widely used by junior and senior doctors to support their work. Preliminary results of a 2015 survey of 1,104 Royal College of Physicians (RCP) members and fellows (response rate after two reminders 42% of the 2,658 members of the RCP research panel) show that 586 (54%) use apps to support their clinical work and of these, 42% believed the apps were 'essential' or 'very important' to their work. However, the survey also reveals that 43% – nearly half – of respondents were concerned about some aspects of app quality. The concern shown by RCP members and fellows about app quality is very appropriate, as several studies have shown that the quality of some apps varies too much for safe clinical use without prior assessment. For example, a study of 23 calculators for converting opioid drug dose equivalents,[1] found dangerously large variations in calculated doses. Conversion of a 1 mg dose of oral morphine to methadone resulted in a dose ranging from 0.05 to 0.67 mg methadone (a 13:1 range), with fewer than half the apps recognising that the conversion formula used should depend on the actual dose as well as on the drugs concerned. Thimbleby et al[2] found that the delete key on many apps does not work correctly, so ironically if a user tries to correct an error they notice, that correction may cause an error they do not notice.

Proposed solution

Building on Donabedian's classical 1978 analysis of the factors that determine the quality of medical care[3] and with input

Jeremy C Wyatt, Harold Thimbleby, Paul Rastall *et al*

App name and version: _____ For iPhone / Android / other:_____

Date of filling out this checklist: _____

1. **Who developed the app, and what's inside it?**

    a) Is it clear who this app is for and how it should be used? Yes / No / Don't know

    b) Is it clear which problem the app is designed to alleviate or what outcome it helps to promote? Yes / No / Don't know

    c) Do the app developer and sponsor seem well informed about this problem or outcome, and likely to be unbiased in their approach to it? Yes / No / Don't know

    d) Have they located sound, relevant, up-to-date evidence, images, video etc to use in their app? Yes / No / Don't know

    e) Do the app screens look well designed, is text clear? Not applicable / Yes / No / Don't know

    f) Is it clear what data the app needs from the user with units defined, out of range detection and a 'clear last patient' button? Not applicable / Yes / No

    g) Does the app collect any identifiable patient information? Yes / No / Unclear

    h) Does it seem to keep user and patient data secure and private? Yes / No / Don't know

    i) If the app is designed to support any medical task,[*] is it CE marked? Not applicable / Yes / No / Unclear

2. **How well does the app work?**

    a) Is the app fast and easy to use in clinical settings? Yes / No / Don't know

    b) Does the app give the user usable answers or advice quickly? Not applicable / Yes / No / Unclear

    c) Do the answers, advice or calculated risks appear to be correct? Yes / No / Unclear

    d) Is there a way to feed back user comments to the app developer? Yes / No / Don't know

3. **Is there any evidence that the app does actually alleviate the problem?**

    a) Have any studies been carried out to measure the impact of using the app on clinical or patient knowledge, actions or (preferably) patient outcomes? Yes / No / Don't know

    b) Were these studies independently conducted, well designed, large enough, and applicable to the user? Not applicable / Yes / No / Don't know

    c) Did any study also examine health resource use, potential harms caused by the app, or quantify cost effectiveness? Not applicable / Yes / No / Don't know

    d) Overall, do the benefits of using this app seem likely to outweigh inconvenience and costs to the user? Yes / No / Don't know

    e) Is there any specific clinical scenario or patient subgroup in which using the app seems particularly likely to be useful? Yes - Which? _____ / No / Unclear

Fig 1. RCP Health Informatics Unit clinical app quality checklist. [*]Screening, diagnosis, test ordering or interpretation, the choice or titration of treatment and prognosis. RCP = Royal College of Physicians.

---

from app developers, the RCP Health Informatics Unit (HIU) has developed a short checklist (see Fig 1) to support clinicians wishing to assess the quality of apps. We recommend that you start by assessing the app structure, including who is it for, who published it, and whether it incorporates relevant, up to date, evidence-based content. Another key aspect is the design of screens that communicate with the user and ask for data to be input: are they well designed, with clear and unambiguous text? When patient data are requested, are the units and methods to collect the data clear? Is the patient's name, NHS number or other identifiers captured, and if so, where are these sent, or are they stored on the device? If the app collects data, does it have a 'new patient' button to clear all previous entries or does it (as many do) carry over some data from the previous patient? If the data are not stored in encrypted form this constitutes a privacy risk when the device is lost or borrowed. You also need to check that the app is CE marked, which shows that it complies with the relevant medical devices and software law, recently summarised in HIU guidance for RCP members and fellows.[4] This applies if the app is used to support medical tasks such as diagnosis, test ordering, interpretation of test results, risk assessment or the choice or titration of therapy. If the aim of the app is not medical – for example, to aid the public with lifestyle decisions such as tracking exercise, weight loss or smoking cessation – CE marking is generally not needed.[5]

If the app passes this long list of screening questions, the next step is to check if it works well – the equivalent of

Donabedian's process measures. How does the app handle error? Try making a few mistakes (does the app reject them?) and try correcting them. Does it appear to generate sensible advice or output for a wide variety of input data items? Does it signal invalid or extremely unlikely data values? Equally, does it request all data likely to be relevant: we know of apps that ignore an age over 75 or provide no field to enter that a patient is diabetic – key information if we need to calculate cardiac risk. Is the output presented to the user in a usable and actionable format avoiding jargon, undefined abbreviations or an obscure risk score that anyone would find hard to interpret at 3am? For example, risks are best communicated using a '1 in N patients' scale rather than a probability between zero and one or a percentage probability.[6]

Assuming that the app appears to work well when tested on sample data, the final hurdle corresponds with Donabedian's outcome: does it actually help us manage patients? The best test is a large, well-designed randomised trial,[7] but so far few of these have been carried out. Trials of apps are perfectly possible, as shown by a study[8] which randomised 128 people wanting to lose weight to the MyMealMate app, a website or a paper diary. At three months, weight loss was highest in the app group and least in the diary group. However, only 20 app randomised controlled trials (RCTs) have so far been published (source: PubMed clinical queries search carried out on 4 September 2015 using 'mobile applications' or 'smartphone applications'), so insisting on an RCT would severely limit the number of apps you could recommend to your colleagues. As an alternative, you could potentially test a small number of selected apps on a series of your own patients. This would entail making your initial decision about each patient, then using each app in turn to see if it changes your decision (and ideally subsequent actions and even patient outcome) for the better. Unsafe apps may be reported to the MHRA hotline (www.mhra.gov.uk/safetyinformation/ reportingsafetyproblems/devices), using data captured on the RCP checklist. This process may also help you to identify typical patients whom the app seems to help, and other patients in which the app seems less useful – helpful when it comes to sharing your findings with your staff. ∎

References

1 Haffey F, Brady RR, Maxwell S. A comparison of the reliability of smartphone apps for opioid conversion. *Drug Saf* 2013;36:111–7.
2 Thimbleby H, Oladimeji P, Cairns P. Unreliable numbers: error and harm induced by bad design can be reduced by better design. *J Royal Soc Interface* 2015, in press.
3 Donabedian A. The quality of medical care. *Science* 1978; 200: 856–64.
4 Health Informatics Unit. Guidance on apps for physicians. London RCP, 2015. Available online at www.rcplondon.ac.uk/resources/ using-apps-clinical-practice-guidance [Accessed 8 September 2015].
5 MHRA. Medical devices – software applications. Available online at www.gov.uk/government/publications/medical-devices-software-applications-apps [Accessed 8 September 2015].
6 Hoffrage U, Lindsey S, Hertwig R, Gigerenzer G. Communicating statistical information. Science 2000:290:2261–2.
7 Liu JLY, Wyatt JC. The case for randomized controlled trials to assess the impact of clinical information systems. *J Am Med Inform Assoc* 2011;18:173–80.
8 Carter MC, Burley JV, Nykjaer C, Cade EJ. Adherence to a smart-phone application for weight loss compared to website and paper diary: pilot randomized controlled trial. *J Med Internet Res* 2013;15:e32.

Address for correspondence: Prof J Wyatt, Leeds Institute of Health Sciences, University of Leeds, 101 Clarendon Road, Leeds LS2 9LJ, UK.
Email: j.c.wyatt@leeds.ac.uk