NATIONAL INFORMATION BOARD

Personalised Health and Care 2020

Work Stream 1.2

Roadmap Direction

Providing citizens with access to an assessed set of NHS and social care 'apps'

The high level objective of workstream 1.2 is to provide citizens with access to a set of NHS and care digital applications which have been *assessed* by the health and care system to enable them to make the right health and care choices

Contents

1 Introduction

1.1 Purpose and objectives

The high level objective of workstream 1.2 is to provide citizens with access to a set of NHS and care digital applications which have been assessed by the health and care system to enable them to make the right health and care choices**.**

The assumption underlying the objective is that an assessment framework for applications will enable health and care professionals to recommend the use of safe and effective digital applications and give greater confidence to patients and citizens to select and use them.

1.2 Terminology

Where the words “app” or “apps” are used in this document it should be understood that this is shorthand for digital applications, which may include either a mobile app, a web-based application or in certain cases a digital service.

1.3 Approach

The development of an assessment framework for digital applications is worth pursuing if it leads to better outcomes for service users and more efficient use of health and care system, and personal resources. Our current knowledge about apps and their benefits means that any evaluative process will need to proceed developmentally, with the assumptions on which it is based being tested at intervals as our knowledge improves.

We intend to trial the proposed assessment framework by running a series of carefully selected pilot apps through early prototype versions, challenging and improving the various parts of the emerging framework iteratively. We will also validate the proposal by undertaking user research. We will continue to do this until the model is ready to be adopted and released to the wider app developer community.

It is important to note that assessment of apps is an entirely new endeavour which does not have the benefit of the decades-long research and development foundation of conventional health technology assessment. It may not yield the benefits referred to earlier or it may need to be significantly modified to enable it to do so.

1.4 Assumptions

A number of assumptions underpin the approach we are proposing to take.. The evidence we have gathered is supportive but more is required to establish if an assessment framework for applications can add value and if so what shape it should take. These assumptions are that:

- Some apps can improve outcomes;
- An assessment system will help improve the confidence of health and care professionals in their use, and promote the development of good apps;

- Careful selection and promotion of apps will improve the use of health and care system, and personal resources;
- An assessment framework can be developed which will allow the NHS and the care system to exercise a judgement as to whether certain apps can be promoted and recommended for use;
- Developers are interested in improving the quality of their apps and are willing to participate in an assessment process on the basis that is effective, efficient and can help them access the market (and generate revenue).

1.5 Structure of this paper

The remaining sections of this paper cover the following topics.

- **Evidence Base** – summarises the evidence and its implications on these proposals
- **Assessment Framework (proposal)** – outlines a proposed approach and framework for the assessment of apps to stimulate discussion and input. This section also outlines the next steps involved in pursuing this work.
- **Benefits** – summarises the key benefits that are assumed
- **Costs** – summarises anticipated costs based on a number of assumptions and principles that are expected to be validated and updated during Summer/Autumn 2015

2 Evidence Base

2.1 What we have done to inform our proposals

A key component of the initial phase of workstream 1.2 ('Discovery' phase from January to June 2015) has been the collection of evidence to inform the development of our proposals. The evidence gathering activities have included desk-based 'structured' reviews and engagement with a range of stakeholders. These activities are summarised below.

2.1.1 Structured reviews

- **Literature review –** considered what existing literature tells us about the current acceptability and adoption of health and wellbeing apps and personal health records. It also reviewed the evidence regarding the impact of these digital technologies on health outcomes, focusing where relevant and possible on the workstream's four discovery phase pilot areas (diabetes, mood monitoring, smoking cessation and dementia). By reflecting on barriers to adoption and concerns associated with safety and quality, this research has informed the development of our proposed model for the assessment of apps.
- **State of the apps market review –** assessed the current state of the UK health app market to understand the nature of UK's most popular apps (segmentation by condition and type of applications) but also to understand the level of availability of

medical apps for specific health conditions, focusing again on the four aforementioned pilot areas.

- **Regulatory framework review -** reviewed any existing regulatory frameworks applicable to digital applications, including guidelines, existing models of endorsement and assessment, and regulation. This also covered other sectors and countries.
- **Clinical effectiveness review –** focused on exploring existing models for assessing the clinical effectiveness of digital applications and the challenges that arise from attempting such an assessment.

2.1.2 Stakeholder engagement

Since this project started, a broad range of experts have attended the workstream's weekly delivery team meetings to share their knowledge and experience in this area. Stakeholder engagement has also been conducted via informed expert, industry and NHS representative bodies including international organisations.

To facilitate future engagement with stakeholders, two advisory groups were formed. Membership of the groups includes NHS and industry experts as well as app developers. The main role of the advisory groups will be to provide key strategic advice and insights to the workstream following publication of the assessment framework proposals in June 2015.

2.1.3 App developer engagement

The advisory groups include selected app developers but in addition to this we are working directly with developers in the four selected pilot areas. Developers of apps from these four areas have, or will be asked to test the emerging assessment model.

2.2 What we have discovered

A number of themes have emerged from the evidence gathering. These are summarised below. More detailed information and references will be available as part of the structured reviews due to be published at the end of June 2015.

2.2.1 Large volume of apps

Over 100,000 health apps are currently available in Europe and this number is growing rapidly. The target users are primarily consumers, however 9 – 14% of the apps in the 'Medical' category target professionals. The vast majority of apps are intended to improve health and wellbeing. Apps focusing on the management of specific conditions ('medical' apps as opposed to 'health and wellbeing' apps) are the minority but are also developing quickly. For example, one study tells us the growth in the rate of publishing diabetes apps has been exponential. In 2009 there were 60 apps available, in 2011 there were 260 and by 2013 there were 650.

Design implication(s):

- **Scalability** – the large size and rapid growth of the apps market might necessitate the first steps of this assessment framework to provide a self-service assurance

against agreed criteria and then a crowd-sourced prioritisation to remove any potential health-system resource overload and act as an effective filter to help focus more robust evaluations on the highest quality apps.

2.2.2 Different approaches to segmenting the digital application market

No one categorisation of digital health applications has emerged from our reviews. Defining categories of apps is made more difficult by the emergence of new types of digital technologies such as wearables. Segmentations most often described are based on risk (to patients), functionality (information only, tracking or intervention), end user (citizens or professionals), connection to health system data records (or not), medical vs. health and wellbeing apps.

Design implication(s):

- **Risk profiling** - assessment must recognise the risk profile of apps both in terms of functional and contextual risk and their potential benefit/utility. High risk / high potential apps should be subjected to higher scrutiny and are more likely to be 'recommended' if their utility is confirmed/proven.

2.2.3 Consumer choice / discovery is linked to the most popular apps

Helping target audiences find relevant and high quality apps is a challenge. Despite the volume of health apps available, downloads are heavily skewed toward a small number of apps. The top 10 apps in any category account for 90%+ of downloads but some evidence suggests that quality is not necessarily linked to popularity.

There are also issues with longevity and sustainability of apps as a lot of apps are developed but they are not maintained.

Nevertheless, the evidence shows there is enthusiasm and willingness for using health apps. If current trends stay the same, by 2017 3.4 billion people worldwide will own a smartphone and half will use health apps.

Design implication(s):

- **Community participation** - assessment should provide a mechanism for "the crowd" to indicate a preference; filtering and surfacing apps that are well designed, safe, actively maintained by the developer and have strong support from end users and other interested stakeholders such as professionals or charities.
- **Sustainability and longevity** - the assessment framework must for example be able to screen out these apps which are prototypes, 'project's and will not typically be maintained beyond their first release.

2.2.4 The majority of apps are focused on health and wellbeing and have simple functionality

The areas of fertility/pregnancy and exercise/sports are the most popular in the mobile health app market in the UK. The top 100 apps in the App Store and Google Play included few apps from our four condition pilot areas (diabetes, smoking cessation, dementia and mood monitoring).

A report by IMS Health provides evidence that despite the large number of healthcare apps developed, the majority have only very simple functionality and do little more than provide

information. For certain conditions such as diabetes, this means patients often need to use multiple apps to manage their condition.

There is a growing number of apps with more sophisticated functionality, including tracking or capturing of user entered data, communication, or statistical analysis but these are in the minority.

Design implication(s):

- **Open but needs led** - assessment should be 'open' to help filter the higher quality apps across the market but should also help steer/direct the market to develop applications where there is a strong health and care need or gap in the market.

2.2.5 Concerns over the safety of apps

The literature indicates great variability in the quality of apps and digital services and there is concern over the risk of harm posed by medical apps in particular. Recent advice by the Royal College of Physicians is that clinicians and trusts should establish if the apps they use meet the medical device definition and if so not use these if they do not carry a CE mark.

Numerous safety and security issues have been highlighted in medical apps used by clinicians. This includes apps that provide out of date information, incorrect algorithms, and inaccurate drug dosage information. Also, patients have expressed numerous privacy concerns regarding mobile apps.

Design implication(s):

- **Strong assurance where appropriate** – assessment must be flexible and an appropriate level of assurance should be applied to high risk applications.

2.2.6 Usability of apps can be a key challenge

The ease of use of an health app will have an impact on its acceptability, adoption and usage - for example, the choice of platform on which the app is available or poor choices in the design of an app (e.g. poor navigation, the use of unclear icons or small typography).

It is also important to be mindful of the 'digital divide'. There is a need to consider the individual needs of an ageing population with regards to smartphones and mobile technology. Although, as discovered by the USA's Veterans Association, greater problems of mobility in older generations mean there can be greater benefits derived from appropriate digital health services.

Design implication(s):

- **Usability** – assessment must consider usability as part of the assessment but validating usability with end users should also be included in the process.
- **Access** - steps must be taken to ensure that those who might benefit most from an app can have access to the technology

2.2.7 There is mixed evidence regarding the impact/utility of apps

The literature review identified some promising evidence of positive impact for apps, for example in the smoking cessation and diabetes areas. However, many studies concluded that digital technologies, including apps, had limited or no impact.

Generally, the evidence on the cost-effectiveness of health apps was also mixed. Some studies have demonstrated efficiency gains, for example a reduction in hospital nights and re-hospitalization for patients with Chronic Obstructive Pulmonary Disease. However other studies have demonstrated on the contrary that health apps lead to increase in demand for health services from the 'worried well'.

Design implication(s):

- **Impact/Utility** –assessment must include a mechanism for confirming (or rejection) the utility of apps before they can be recommended. However, this level of validation is likely to be high cost and not scalable and therefore should be undertaken where the claimed benefit/impact is significant.

2.2.8 New research methods needed

A number of challenges have been identified regarding the nature of research into the impact and benefits of health apps. These include the lack of longitudinal studies looking at long term impact, the involvement of developers in the research, the use of non-validated outcome measures, recruitment bias and drop-out rates issues.

Randomised Control Trials (RCTs) are rightly the "gold standard" for assessing the impact of medical apps on quality (e.g. for patient-centred care, costs of care delivery, impact on resource use, cost effectiveness) and patient outcomes, as they address unknown variation across patients. However, the pace at which the health apps market has expanded has meant that the traditional approaches to evaluating the effectiveness and safety of healthcare interventions are not always feasible or appropriate for all apps.

Certain apps can lend themselves to rapid large-scale evaluation and digital technology allows for the development of new research models.

Whist apps have been involved in a range of study types, a clear alternative methodology for conducting robust studies of digital technologies was not identified as part of the literature review.

Design implication(s):

- **Impact/Utility** – developing a suitable assessment framework may require the development of new evaluation methodologies suitable for digital applications.

2.2.9 Entering the NHS market

Entrepreneurs consider it difficult to access the health app market due to the lack of a clear regulatory framework, interoperability and common quality criteria. App developers and investors are also reluctant to enter the NHS market as the commercial incentives for doing so are not clear.

The preferred app business model has shifted significantly from a single download payment to initially providing the app for free and then attempting to charge repeatedly for additional extra services (the "freemium" model). In many cases apps are just provided for free. These

revenue models raise questions regarding the ability of developers to maintain the quality of their apps. Many developers are therefore likely to rely on other sources of funding such as advertising or the resell of data.

Within the NHS, there is a lack of understanding of how the NHS can pay for and provide apps to patients. If commissioners are going to pay for new digital technologies, they need to see value in making an app available and this information is not always available or validated.

Design implication(s):

- **Revenue Generation** – assessment must facilitate the adoption of proven digital technologies and will need to integrate with existing commissioning and prescribing mechanisms to support revenue generation.

2.2.10 Regulatory Position

The health apps market is a new area of relatively low regulation. According to a review by PWC published at the start of 2013, over 150 countries had yet to develop regulatory frameworks or guidance. Whilst this situation is expected to have improved over the last 2 years it underlines the embryonic state of regulation for health apps.

The European Commission issued a Green Paper, in April 2014, on health apps for consultation and nearly half of the respondents asked for more patient safety and transparency of information, by means of certification schemes or quality labelling of lifestyle and wellbeing apps.

In the UK, we identified strong demand for an assessment framework for apps from a range of stakeholders. Some stakeholders consider that the regulatory stance adopted by the MHRA only covers a very small subsection of all health/medical apps available leaving many potentially unsafe apps in the public domain. In this respect, there are important links to be made with the consumer protection legislation. How it applies to the market of apps and the types of protection it provides will need to be considered too.

Design implication(s):

- **Confirmation of gap in system**
- **Clarity of positioning -** strong links with MHRA and existing regulatory frameworks such as consumer protection legislation will be required

2.2.11 Scepticism regarding the concept of a centralised assessment model for apps

Whilst many stakeholders approved of the idea of introducing an assessment framework for apps, many simultaneously feared it would create a barrier to innovation, would not scale well, would take much too long to provide relevant guidance and would have a high cost and therefore poor return on investment.

Some app developers believe an endorsement model will help differentiate their better applications from the mediocre ones, but many motivated app developers also told us they did not believe in the value of a centralised assessment model and emphasised the value of local endorsement through adoption. App developers, Small and Medium Enterprises or charities, told us they wanted airtime with their local NHS customers (CCGs or hospitals) and the best endorsement they believe in was ‘a pilot, a sale and then praise from local

NHS CEOs who are using their product'. They also challenged that central endorsement would make the local NHS more likely to buy. There also expressed a concern that the risks of seeking and failing formal central endorsement could outweigh the benefits of getting it.

These challenges from stakeholders resonated with user research conducted on the current Health Apps Library pilot. This demonstrated the difficulty in providing a sustainable albeit light touch solution that can change app adoption behaviours. Many patients assumed that a more rigorous certification process underpinned the Apps Library than was in fact the case, primarily because of the use of the NHS branding. Concern was also raised over the level of resource required to run and maintain the site. Also, it was not clear that the Apps Library changed the approach of users in selecting apps. People did not look for any 'kitemark' or certification process for any other apps downloaded (although admittedly such a 'kitemark' was not available to use). Many people would only visit the Health Apps Library when looking up symptoms or conditions they were experiencing rather than regularly using it as an advice platform for choosing health apps.

Design implication(s):

- **Use iterative and open approach - r**esistance to assessment proposals should be expected, fully understood and used as a robust challenge to the creation of this assessment framework.
- **Cost effectiveness of the framework**- the cost of running the assessment framework must be proportionate to the value the market can expect from it

2.3 Evidence still required

More evidence is still required to establish if an assessment framework for applications is needed and if so what shape it should take. Further evidence collection is planned as part of workstream 1.2., as follows:

- **User research with the public, professionals and commissioners:** User research is needed and this must be significant and on-going. It will help us to understand people's behaviours and needs, motivations and attitudes, and frustrations - when trying to choose or recommend apps. This will inform the design of the assessment framework and give an idea of the value different users would place on app assessment: Does it help them select apps in the first place? What would they expect assessment to consider in order to give them confidence in using, recommending or funding an app?
- **Further app developer engagement:** This is needed to test the detailed questions we would consider as part of the self-assessment component of the proposed framework as well as to test the appetite for the overall approach and its benefit case.
- **Testing the model with stakeholders:** So far we have taken the views of stakeholders on what is needed. We now need to take their feedback on our emerging proposals.

3 Proposed Assessment Model (Vision)

3.1 Reconciling NIB objectives and emerging evidence

The assumption underlying the objective is that an assessment framework for applications will enable health and care professionals to recommend the use of safe and effective digital applications and give greater confidence to patients and citizens to select and use them.

On balance, our review of the evidence collected to date suggest that an assessment model may be welcome by some users as it could support them in identifying the better apps within an increasingly crowded market - although more user research is required to understand the problems users face and how to solve them. The evidence collected so far also informs the dimensions of apps 'quality' which an assessment model would need to consider. These include safety, privacy, effectiveness / evidence of impact, sharing data onwards, accessibility and inequalities, interoperability, usability, and technical stability

The evidence review has however highlighted a number of challenges that the overall framework needs to consider.

If the assessment model is to be used by the NHS to formally recommend, brand and fund the use of digital applications, there is an expectation that this assessment should be very robust - independently delivered and based on good quality evidence of impact. It should 'put the bar high'. However such a robust assessment is unlikely to be available at scale, at speed or at low cost. There is a sense that this level of assessment could only be applied to categories of applications where the potential upside in adoption is most promising and demanded by the health and care system. But such a model of assessment could only consider a small number of apps and would exclude the vast majority of apps available to citizens.

In order to open the assessment framework to a much wider base of digital applications, then a model based on self-assessment and/or community participation would need to be used. But by their very nature, these types of assessment cannot be expected to carry the same signal of 'trust' from the NHS. A self-assessment and/or crowd sourcing model could help identify apps that meet basic criteria - and/or are the most popular with a community of users, but they cannot guarantee that using these apps will improve outcomes or be good value for money for citizens or care providers/commissioners.

Our initial app assessment framework proposals address the original NIB objectives and take account of the challenges identified in our research. It seeks to build an affordable as well as credible solution. The proposed framework is a staged approach to assessment which blends the use of self-assessment and community participation to identify apps with promising utility, with a more formal evaluative process, in which certain groups of digital applications can be assessed leading to more 'formal' recommendations for use, based on independently evaluated evidence of impact.

3.2 Emerging design principles

Our proposals adopt a number of design principles that address the challenges identified by stakeholders during the Discovery phase:

- **The model should be 'open' to support new entrants and innovation.** The first stages of the model can be open to all digital applications. However, there is also an

opportunity to focus assessment, at least initially, on disease and condition specific apps with the greatest claims of clinical utility

- **The model should be low cost and must deliver value to the health and care system.** The bulk of the apps considered will go through a low cost self-assessment stage. Higher levels of investment will only be considered where there is an expectation of positive return on investment from the category of applications subjected to a higher degree of assessment;
- **The model should be scalable.** By being digitally supported, the model will be scalable for the early stages of assessment where most of the volume might be expected;
- **The model should involve health and care community participation.** Community participation will be encouraged and integrated into the assessment process;
- **The model should involve testing claimed utility/benefits.** This type of impact/utility validation is likely to be high cost and not scalable and therefore should only be undertaken where the claimed benefit/impact is significant
- **The model should build on existing standards** where possible (for example the new PAS 277 from the British Standards Institute) and connect to existing regulatory frameworks where applicable, for example the MHRA;
- **The approach should be flexible and proportionate**, adapting the assessment components to the nature of each application – relative risk, benefit, cost of each app;
- **The model should be attractive to the industry**. Each stage of the assessment will deliver value to applicants.

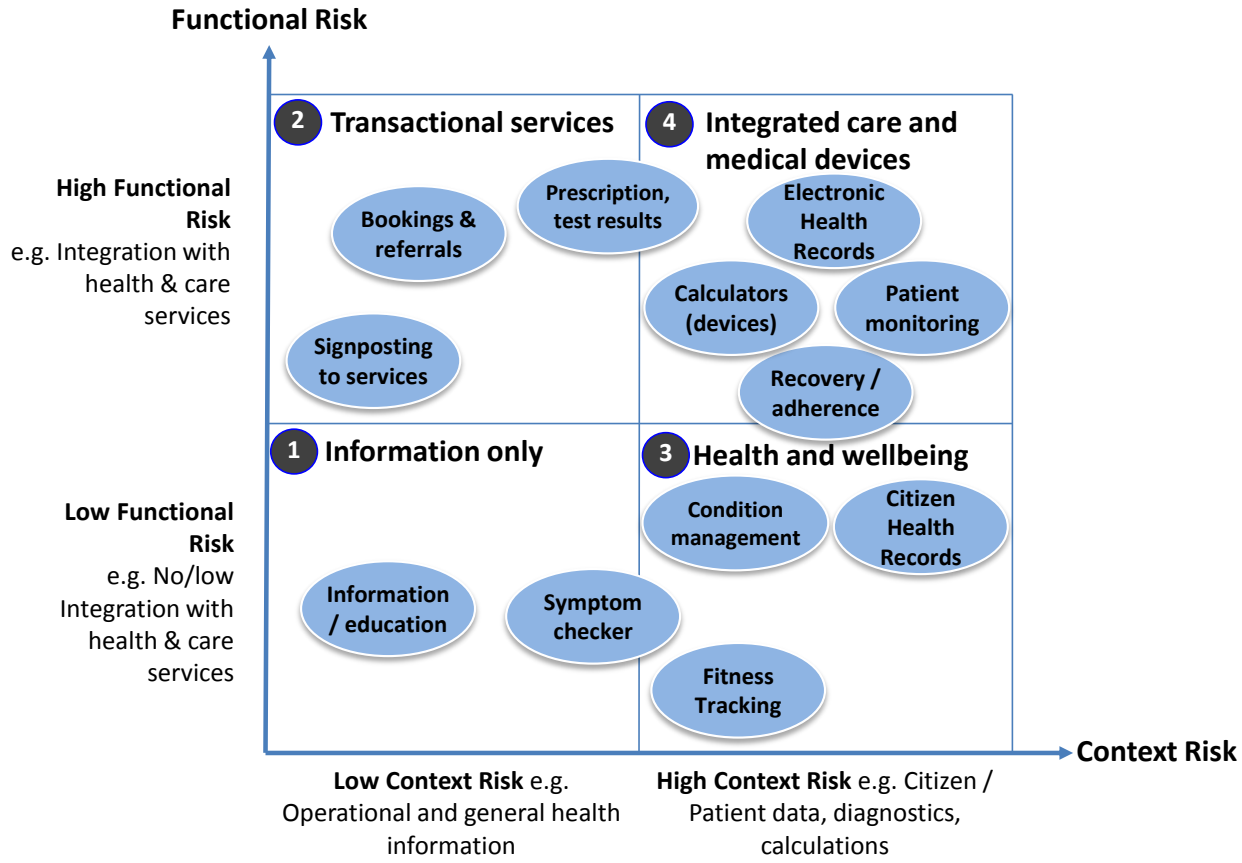
3.3 Risk based classification of apps – linked to assessment scrutiny

Categorisation of apps can help in the judgement of risk associated with the app or in how helpful the app can be to the health service. However, the apps market can be broken down in many ways to serve different purposes and needs. For the purposes of this paper a risk-based classification has been applied to better understand the risk profile (and potential impact) of apps. This classification segments the market based on the sensitivity of the information held or processed and the complexity of the app functionality. We may choose to adopt a different risk classification if further research indicates an alternative method is more appropriate.

Figure 1 below provides a summary of the risk classification model with the bottom left being low risk and top right representing the highest risk apps. A few example app types based on their purpose have been mapped to the classification model for illustrative purposes.

Apps with a higher risk profile will require more scrutiny both in terms of assessing their quality and in terms of validating any claimed utility/benefit.

Figure 1: risk classification model for digital applications



3.4 Overview of proposed assessment framework

We have developed a high level outline of the proposed assessment framework, available under figure 2. The four stages of the proposed assessment framework are also summarised below.

It is important to note that the model assumes that only a small number of apps will go through the entire process and that apps do not need to go through the entire assessment in order to benefit from the process. Visibility, for example exposure on NHS.UK (NHS Choices), will arise from completing Stage 2. However, strong positive recommendations to use and to fund applications will require independent evaluation that will occur in Stages 3 and 4. It is likely this level of assessment will be relevant to only a sub-set of applications.

3.4.1 Decision to apply

The assessment process will be open to all apps. As such, most registrations for assessment are expected to be driven by app developers hearing about the process and choosing to enter the assessment. In addition, the health and care system may choose to encourage certain groups of applications to participate in the assessment based on strategic or policy priorities, local priorities or public trends.

In this illustration, we have assumed that 10,000 apps might eventually be submitted for a Stage 1 self-assessment.

3.4.2 Stage 1: Self-assessment against agreed criteria

This stage is designed to screen the large number of apps that might apply and to identify those apps that meet some agreed criteria, based on self-assessment.

Apps which undertake self-assessment will follow a structured set of questions organised against key 'quality' dimensions. These dimensions will include safety, privacy, sharing data onwards, accessibility, usability, technical stability and interoperability. Self-assessment will also use proxy questions to uncover any available evidence of impact.

The intention is for Stage 1 to be supported by a digital service. This will lead developers through the assessment and provide guidance and links to best practice where appropriate.

High transparency and the use of random audits will be used to keep answers honest and prevent some entrants from 'gaming' the self-assessment process.

Some apps may be identified, through the responses given, as higher risk apps. These will be required to undertake additional external assessments - for example, if they are likely to qualify as a medical device, to register with the MHRA and self-certify as a class1 device or employ a Notified Body (for higher class devices) to approve CE certification.

Only those apps that meet the necessary criteria will be progressed into Stage 2 and be available for community evaluation. For illustrative purposes, out of an assumed 10,000 apps entering Stage 1, only 2,000 may be invited to move into Stage 2.

3.4.3 Stage 2: Community evaluation

Similar to many online up-voting websites, this stage is designed to use the wisdom of an engaged community of professionals, commissioners or end-users to evaluate and therefore support the better apps emerging from self-assessment. The 'crowd' would be asked to give opinion on functionality, usability and share anecdotal evidence of impact and utility. Much of the feedback may arise from local pilots and the evidence of impact from local implementation projects. Early adopters in the clinical community will be provided with a platform to share their successes and challenges with health apps.

Current models for crowdsourcing can be open to commercial bias or are too unstructured to help clinicians recommend an app so this stage will be designed to take this into account.

Apps that are evaluated by the 'crowd' through Stage 2 will benefit from exposure on NHS.UK. Exposure may also be available through other channels e.g. Public Health England's Stopober campaign. Exposure in this way will not constitute a formal recommendation for use.

The presentation of the apps may reflect the relative performance of apps across Stages 1 and 2 with only the top scoring apps being visible on NHS UK.

For illustration purposes, up to 100 apps per year emerge as 'evaluated' by the community though this phase of the process.

3.4.4 Stage 3: Preparing a benefit case

Stages 3 and 4 of the framework are about taking selected apps through a more robust assessment process. This level of assessment will be required to ensure that apps that are

eventually recommended by the NHS, reimbursed or possibly prescribed have the evidence to support their claims.

It is likely these stages of the assessment will be reserved for a small number of apps from distinct priority areas with strong business case for efficiency or cost effectiveness.

At Stage 3, apps will be enrolled in a range of activities to enable further assessment. This may include data sharing and end user participation, advice on study design.

The type of study and support needed in Stage 3 will depend on the type of applications under consideration. This is where the chosen method of apps categorisation presented in section 3.3 will be of particular importance.

3.4.5 Stage 4: Independent impact evaluation

Stage 4 involves an independent impact evaluation. As in Stage 3, different approaches may be required for different categories of apps with clinical interventions likely to require the most robust evaluation process. Some categories of apps may be assessed much faster than others. These approaches still need to be defined (see next steps section).

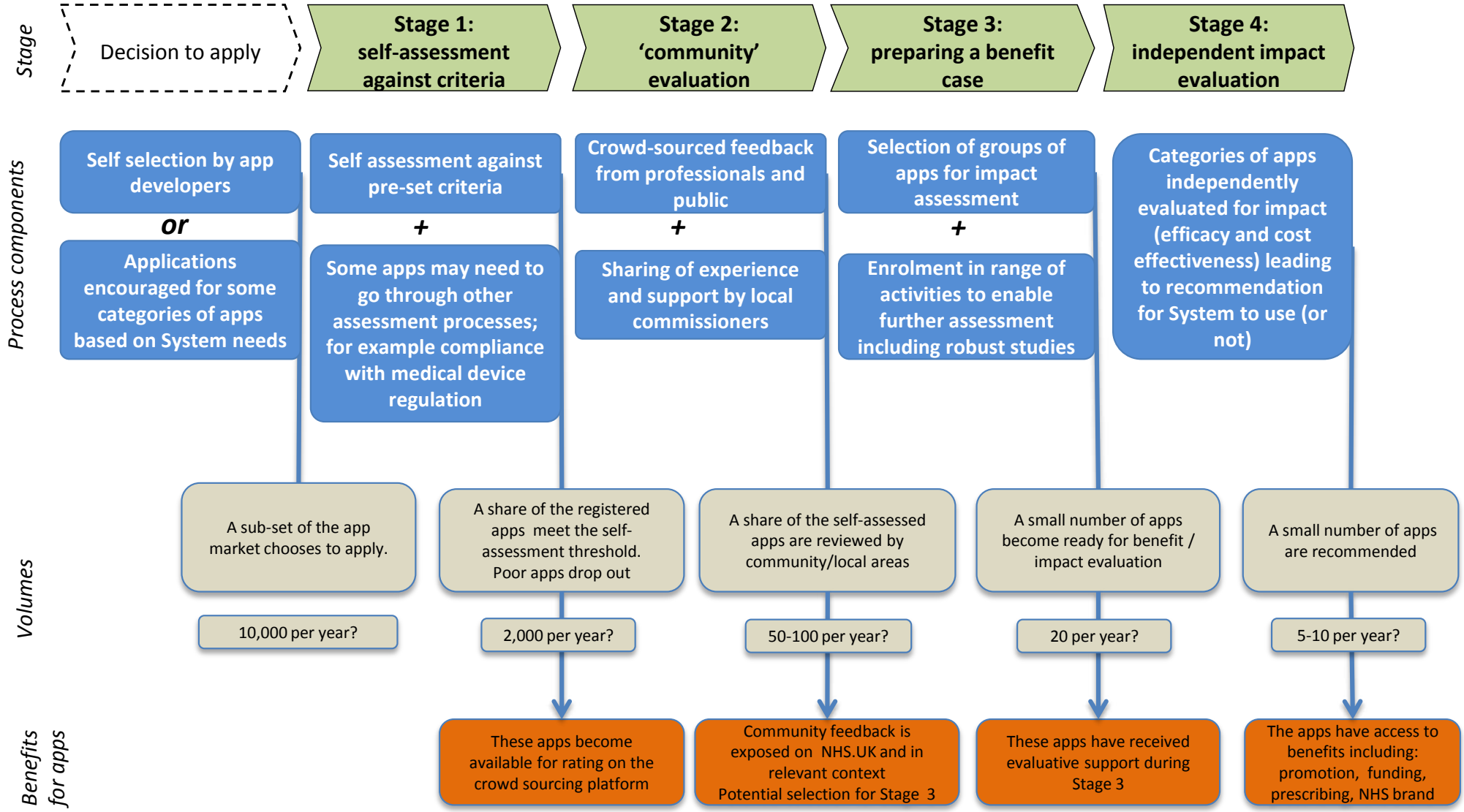
Apps emerging from Stage 4 will become recommended interventions and as such it is expected they will benefit from a range of adoption support mechanisms that could include branding, commissioning support, or reimbursement.

For illustrative purposes, as few as 10 apps per year may be formally recommended by the NHS. Recommendations may increase over time as more apps evaluated through Stages 1 and 2 undertake robust studies of impact.

3.5 Visual overview of the assessment framework

The 4 stages of the proposed framework are summarised in figure 2 below, focusing on key process steps and the benefits that apps receive after each stage. Illustrative numbers of apps entering and exiting each stage of the assessment are presented.

Figure 2: a high level outline of the proposed assessment framework



3.6 Next steps

Key strands of activities expected from June 2015 through to early 2016 include:

- **Continued engagement with stakeholders and app developers** – engagement with a range of stakeholders to gather feedback on the proposed framework will continue following the publication of the NIB roadmaps in June 2015. This, together with the feedback from the user research (see below), will be used to evolve the model where necessary. The adjustments required may be more or less substantial depending on the nature feedback received after June 2015.

- **User needs research** - We will test that the proposed model serves actual user needs. This will include separate research with commissioners, health and care professionals, patients and citizens. Nurses and GPs will be a key user group to engage as we expect they will be users of the system as well as core contributors to the crowd-sourced component.

  A key challenge of the proposed solution is one of terminology: how the different stages of assessment will be referred to and whether the different stages of the solution can be (or need to be) correctly understood by the different groups of users.
  - Should an app emerging from Stage1 (self-assessment) carry any reference to 'quality'?
  - Can Stage 2 (crowd-sourced evaluation) carry any reference to 'trust' or 'endorsement'?
  - At what stage can a reference to an `NHS recommendation' be applied to the app?
  - Can users understand the difference between an app being assessed by an avid community of users or GPs and an app being formally evaluated and recommended by the NHS based on evidence of impact?

  Linked to these labels are the issues of risk and liability that the scheme will carry.

  The user research will also help understand the touch points and channels where the exposure or promotion of apps to potential users will be most effective. This will help define the platforms we need to use and shape truly valuable promotional 'benefits' to app developers.

Assuming the key components of the framework are retained, further work to develop the various stages of the framework will be required as follows:

- **Further development of the Stage 1 self-assessment questions and algorithms** - An initial draft of the self-assessment questions will be available at the end of June as part of the workstream Discovery phase deliverables. These will need to be iterated and refined with application developers but also with organisations that have expertise in the various components of the assessment. A scoring algorithm including a threshold for 'good enough', which qualifies apps for Stage 2, will need to be agreed. Another remaining challenge will be to produce a meaningful 'categorisation' of apps. Whist the model will be open; a good categorisation will be required to allow an effective use of the apps emerging positively from Stage 1 of the process and beyond. Stage 1 must also include a more in-depth assessment of higher risk apps. How this is achieved will need to be defined as well. Finally, there is

also a question as to whether self-assessment should include random audits to ensure greater quality of responses

- **Development of the digital platform for Stage 1** – an initial prototype for Stage 1 of the assessment will be delivered as part of the workstream Discovery phase. This was created with a view to continue to test the concept of self-assessment with pilot app developers. This prototype will be iterated through the piloting process and developed into a beta site. Where the digital services are hosted and how they will be accessible to app developers needs to be decided as part of the development plans.

- **Defining the principles for Stage 2, the crowd-sourcing of feedback** – More work is needed to establish who should contribute, how contribution can be made, how to promote and incentivise contributions, and the technical implications of operationalising crowd-sourced evaluations.

- **Further developing Stages 3 and 4 evaluative processes -** the latter stages of the model require further engagement with experts in the field of impact and effectiveness assessment. For Stage 3, advice needs to be developed on how to conduct valuable research for digital technologies including advice on selecting outcome and comparators. This may differ for different types of application as identified under section 3.3. The advice should recognise that digital technologies may engage new types of outcome such as patient empowerment. The advice should also recognise that the technologies may be dependent on new models of care being adopted and new capabilities being available across the system such as the ability to handle larger amounts of patient-generated data. For Stage 4, agreement on how to conduct a formal assessment and whether this assessment is carried on individual applications or categories of applications has to be established.

- Across the entire model, **the feasibility of providing sufficient benefits** for app developers, clinicians and end users must be investigated and proposals should be published. Other NIB workstreams will be critical as the some of the benefits will come from these workstreams, for example work on NHS Choices (workstream 1.1), supporting professional access to digital technologies (workstream 2.1), platform for apps accessing national data (workstream 8), professional education (workstream 6), commercial model and funding to support adoption including prescribing (workstream 5).

- **Operating model and business case** – an investigation into how the model will be run after the initial pilot phase needs to be undertaken as part of the development of a first business case for implementing the overall assessment framework.

3.7 Deliverable timing including quick wins

The workstream deliverables at the end of June 2015 will be:

- Proposals for an application assessment framework – as articulated in this Roadmap document
- Digital (early) prototype of the self-assessment stage (Stage 1)

- Structured reviews (as listed under section 2.1.1) posted on the NIB pages of Gov.UK
- Following the NIB commitment to support the development, diffusion, and adoption of low-cost high-efficacy apps with a particular priority on mental health services, NHS England, Nesta and the Behavioural Insights Team will announce a digital mental health development programme in June 2015. The programme would enable the development of digital tools that follow the assessment process and offer an opportunity to test the proposed model in practice.

Key activities will take place in the summer leading to the end of November 2015. These will include:

- User research to test audience needs
- Piloting of the stage 1 prototype
- Engagement with clinicians: in Autumn 2015, we will be directly engaging with a cohort of GPs, and other clinicians to consider the feasibility and interest in the crowd-based component of the model.

These activities will achieve better stakeholder engagement if they can be focused on selected areas of care that will be presented as 'experiments' to the market – experiments to test, learn from and continue to develop the right model.

By early 2016, we expect that a 'beta' version of the Stage 1 digital service can be launched and Stage 2 will have been piloted. Decisions on whether and how to progress Stages 3 and 4 will also have been reached by this time and plan for operationalising the broader framework will be available as part of a broader business case for implementation.

4 Benefits

4.1 Key end benefits

A number of benefits can be assumed from putting in place an assessment framework for apps. It should reiterated that these benefits are based on the assumptions that an assessment framework can truly identify the better apps, influence decisions and change the adoption levels of apps and lead to the increased use of higher quality apps. If these assumptions hold, then the key end benefits of an assessment framework for digital applications could include:

- **Reduced costs to the health and care system:** Digital interactions have been shown to be cheaper than other interactions between professionals and the public. There is some good evidence to show that digital interactions cost about 1% of face to face costs.
- **Improved outcomes:** Although there is some evidence to suggest that digital interactions are more cost effective that face to face equivalents in certain circumstances, this evidence is relatively poor. It is expected that better outcomes will be delivered by channelling scarce funds to deliver face to face services where they are most needed, using digital as an integral part of the suite of treatments and interventions available to the health and care service in therapeutic areas where a good evidence base exists for the efficacy of digital interventions;

- **Improved system:** other industries are 10-20 years ahead of the health and care system when it comes to using technology. This model will help to improve the reputation of the English health and care system. The assessment model brand, and the organisations delivering it may come to be synonymous with good quality, trusted digital therapies.

4.2 Intermediary Benefits

Some intermediary benefits can also be identified and include:

- Developers of good quality, safe and effective apps will be able to market their products much more effectively to the health and care system and, to a lesser extent, the public here and the wider world. More good applications will be developed by application developers and their visibility will improve.
- End users will have a more informed choice of apps they choose for themselves or in conjunction with a care professional.
- Clinicians will be able to access apps assessed as meeting some minimum quality criteria and validated by their peers, enabling them to promote or use apps with much less fear of them being ineffective or even causing harm.
- Commissioners will be able to use information produced through the different stages of the assessment process to inform a cost/benefit assessment of adoption of an app for a service or across a catchment area.
- Gaps and areas of needs in the market could be identified and the market influenced to respond to these needs.
- Over time, stronger evidence around the impact of apps and digital services will emerge, which should lead to higher levels engagement with and support for digital services.

4.3 Value case for Stages 3 and 4

The value case for assessment cannot be quantified without reference to individual app categories. We expect that for a category of apps to move into Stages 3 and 4 of the assessment, which are likely to carry a greater central cost, that a business case should be submitted which elaborates on the likely the value of pushing the adoption of the set of digital tools. The benefit case would need to include:

- Potential saving of replacing non-digital services with digital equivalent
- Relative efficacy of the digital services in terms of cost per positive outcome
- Access to new populations and/or increase in overall treatment rates across digital and non-digital interventions

Apps or categories of apps with promising business cases would be put forward for Stages 3 and 4 so that the NHS / broader system can independently validate the claims articulated in the business case and formally recommend, or not, the use of the type of applications by the system.

5 Costs

It is too early to provide a complete cost profile for the app assessment framework, in terms of both the set-up and on-going operating costs. This is because the detailed operating model for many components of the proposal is not yet defined. Also, some elements of the framework may yet change significantly based on stakeholder and user feedback expected over the coming months.

An outline of the type of costs and cost profile to be expected in the set up phase and once the framework is in operation have however been put forward and tentatively estimated where it was possible to do so.

At this point in time, greater focus is on estimating the cost involved in continuing to manage and deliver workstream 1.2 over the next 6 months. This will require continued investment to provide the level of leadership, range of skills and focus required to progress to implementation at the end of 2015 as outlined in the deliverable section.

5.1 Summary of indicative costing:

Cost area	Estimated Costs
Project design and development costs to December 2015	up to 700k
Annual running costs	Fixed costs: up to 390k Variable costs: unknown

5.2 Project design and development costs to December 2015

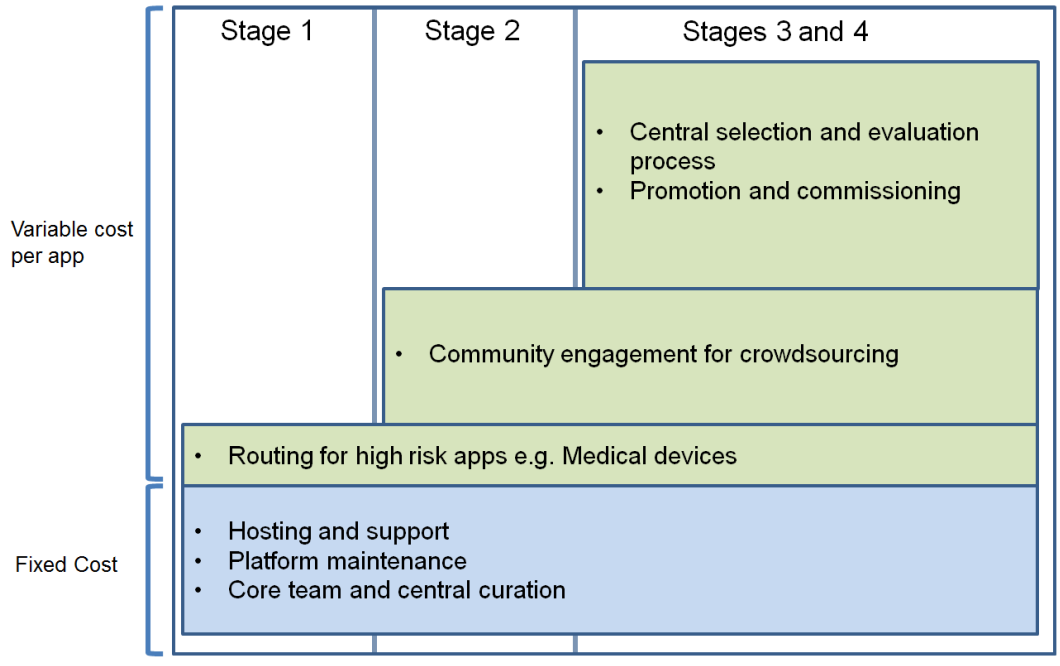
The cost of delivering the Stage 1 of the assessment framework is broken down into internal and external costs in the table below:

Description	Estimated Costs	Key assumptions
Core leadership and programme team to support engagement and coordinate delivery	Up to 150k	Staffing – for 6 months • Dedicated senior project lead • Project manager
Conceptual development and iteration of the framework (Stages 1 to 4) with stakeholders and using pilots	Up to 250k	Staffing – for 6 months: A multidisciplinary team will be needed including the following capabilities: • One or more user researchers • One or more people with design and development experience • Subject matter expertise • Buyer perspective Tasks will include: • iteration and building the algorithm

		for Stages 1 • prototyping to support crowdsourcing approach in Stage 2 • research and consensus building on suitable evaluation methodologies for Stages 3 and 4.
Delivery of digital tool and platform to support Stage 1	Up to 200k	• Based on an indicative quote from an agency • Includes requirements gathering, development and testing. • Further options and associated costs for delivery will be evaluated in June/July • Includes technical review
User research	Up to 100k	Separate research with citizens, patients, clinicians and health care professionals, done at multiple stages of development.
Maximum Total	**Up to 700k**	

5.3 Annual running costs

Operational costs can be broken down into fixed and variable cost as illustrated in the diagram below. Variable costs are dependent on the number of apps going through the various stages of the model. It is not currently possible to estimate variable costs.



Fixed costs	**Estimated Cost**	**Assumptions**
Hosting and support	up to 30k	• Externally provided – options to be evaluated
Platform maintenance	up to 60k	• Externally provided – options to be evaluated
Core team and central curation	up to 300k	• Programme manager • Subject matter expertise • Development support
Total	**Up to 390k**	