Your Covid-19 Risk: Reflections on the Development of the Tool

KEBEDE BEYENE  
AMY HAI YAN CHAN  
JAMES A. GREEN  
SANDER HERMSSEN  

All authors share joint first authorship  
*Author affiliations can be found in the back matter of this article

CORRESPONDING AUTHOR:  
Kebede Beyene  
School of Pharmacy, The University of Auckland, NZ  
k.beyene@auckland.ac.nz

KEYWORDS:  
COVID-19 risk tool; health risk communication; behavior change; risk behaviour; agile development

TO CITE THIS ARTICLE:  
In the absence of a vaccine, our only defence against the spread and transmission of the coronavirus depends on the behaviours of individuals. In mid-March of 2020, a small group of researchers/collaborators/friends known to each other through the European Health Psychology Society ‘social’ channel recognised the need to develop a tool to support the behaviour changes required to adhere to the new public health recommendations. The group started working on a project led by a core team from the Netherlands (Gjalt-Jorn Peters, Sylvia Roozen, Gill ten Hoor, Rik Crutzen). This project became the “Your Covid-19 Risk” tool (https://your-covid-19-risk.com), which has been used by over 50,000 people globally. In this editorial, we briefly describe the project and the development of the tool, and then offer three viewpoints on important lessons learned from this project and their implications for future work: ‘Agile’ development, risk communication, and cross-cultural development and translation. Together, these three independent sections offer insights into problems that are relevant for many intervention development processes in public health and elsewhere, and may contribute to the debate on how to solve these problems.

1. THE DEVELOPMENT OF THE “YOUR COVID-19 RISK” TOOL

The project started small, but from mid-March, the project team expanded rapidly, recruiting people with needed skills ranging from infectious diseases experts, virologists, epidemiologists, and behavioural scientists to designers and programmers. Ultimately our project team ended up with over 140 people involved (94 represented on the volunteers list), representing remarkable geographic, cultural and linguistic diversity. Over 30 countries were represented, with the tool available in 22 languages. The primary method of communication was Slack — a communication platform that enables different ‘channels’ for chatting on specific topics, as well as direct messaging and notice-board type functionality. This enabled groups of people to work on specific topics, ranging from risk communication, server capacity, translations into specific languages, to design and media considerations. Rather than ‘death by email’, it served as a central hub for the project. The core team also held regular meetings to discuss overall strategy, and there were also some social videocalls to enable people to meet each other and to chat. Occurring as it did, during the middle of the coronavirus, this provided a welcome break for some! Sander Hermsen reflects on this more agile design process in Section 2: Agile Development. By early May, the tool development reached a point where the team felt it was ready to be released.

The tool development process involved: determining target behaviours, developing and testing questionnaire/intervention content, and designing and programming the online environment. An overview of the development process and the theory used to inform the intervention part of the tool can be found here: https://osf.io/cbmd8/.

The tool had three parts. The first part was a questionnaire that asked about people’s behaviours related to the transmission of COVID-19 such as social distancing, hand-washing, self-isolation; the second part was an optional questionnaire with questions on the determinants of those behaviours, and the third part presented the tool users with their risk estimate and a set of tailored messages to support preventive action on their part. The questionnaire in the first part had three goals. First, for use in the intervention; the individual results informed the risk communication visualisation and messages in the third part. Second, the data on perception, intention and adherence to these behaviours gave us the opportunity to learn what we needed to know to further develop the intervention; and third, the results from the questionnaire data offer insights for public health and policy institutions and governments to inform their risk communication efforts. The data from the second questionnaire further helped us develop the intervention, and could also provide insights to public health professionals and governance.

Open Science principles were to the fore. Aside from Slack (which is free (as in beer) but proprietary software¹), the infrastructure underpinning the tool was all built on free/libre open source software (FLOSS). The tool itself was constructed using a bespoke R package (written by Gjalt-Jorn Peters) that authored a LimeSurvey questionnaire in each language, with all data and resources stored on GitLab and moved with Git. Data was processed in R, and then published on GitLab. We were lucky to have key hosting resources donated by Slack, Netlify, and Limesurvey GmbH. The anonymous and semi-“Born Open’ data is freely available from a public repository affiliated to the project website.

A key reason for the bespoke R package (“limonaid” – at https://r-packages.gitlab.io/limonaid/) underpinning the survey creation was that the tool was translated into 22 languages, giving us coverage of over 70% of the world’s population. This meant that any change to, for example, the risk estimate could be automatically updated through all the different language versions of the tool. As the tool was intended to be global, there were considerations that needed to be taken into account in terms of languages, translations and meaning. The considerations for developing a tool across diverse cultures and translating it into multiple languages is discussed further in Section 4: Cross-cultural Development and Translation, by Kebede Beyene.
2. AGILE DEVELOPMENT — SANDER HERMSEN

When developing digital interventions for healthy behaviour change, there is a disconnect between the relatively slow, non-parallel, pre-defined methods that behavioural scientists use, and the higher-paced, more flexible and pragmatic methods used by designers and software developers. Integrating insights from the behavioural sciences into modern ‘Agile’ development approaches (such as Scrum, Kanban, Google Design Sprint, and many others) is often difficult (Hermsen et al., 2016; Hermsen et al., 2020).

The “Your Covid-19 Risk” project has been a fine example of scientists trying out development methods that are normally used by ICT developers and designers. At no time, has the “Your Covid-19 Risk” project pretended to be rigorous enough to be scientific research, or even “Science lite”, but the approaches used in the project constitute an interesting and strong first step towards a more agile science, as called for by many scholars who recognise the shortcomings of the current gold standards in intervention development and evaluation (Hekler et al., 2016).

The strengths and weaknesses in the carrying out of this project neatly align with findings among designers and health professionals. For instance:

We had some really good, effective, and time-saving methods to collect epidemiological insights to inform our tool: we performed a rapid literature review where dozens of researchers integrated findings from hundreds of papers within days, and we performed a mini-Delphi study among 11 virologists and epidemiologists from all over the world. Similarly, our methods to validate the tool were fast and lean, and touched upon methods from design thinking. For instance, we performed a vignette study to test the risk estimate model in which experts rated user stories based on ‘personas’: fictitious archetypes of users, each reflecting a distinct pattern in goals, attitudes and behaviours, based on empirical research among potential users (Cooper, 1999). This shows that there is ample room for more creativity in laying the foundations for evidence-driven intervention development processes than current practice.

Furthermore, the ‘parallel’ development which is generally seen as one of Agile’s strongest points really worked out. People took up the jobs that most fitted their expertise and capacity, and teams formed (in new Slack channels) to deal with new tasks that came up. When a team encountered issues that needed dealing with at a higher level, those issues were fed back to the general Slack channel, where they could be discussed and dealt with by the whole group or referred to team members who had the skills to handle the task.

On the other hand, there were some shortcomings. Agile processes prefer speed and development over rigour and reflection. As a consequence, it was at times very hard to keep track of what is being done, what decisions were being made, and what rationales informed those decisions. This is the case in most Agile developments (Hermsen et al., 2020) so it did not come at a surprise. We tried to tackle this by using reflective journals (day-to-day notes of what was done and what decisions were made (Thorpe, 2004)), and the core team provided an almost daily overview of the work currently being performed, posted in the main Slack channel. Even so, it remained difficult to be aware of everything that was being done. The fact that everybody on the project was a volunteer who did this in their spare time, and maybe did not have the extra time needed for reflective journaling, certainly did not help here.

In the beginning, we were a large group of very enthusiastic people, who could devote a lot of time at short notice to the project. The topical nature of the pandemic and mutual need to contribute our skills to an evolving and growing global pandemic were key driving forces behind the project. This energy, however, could not last forever. At the end of the project, the group that actually kept things going had dwindled to quite a small crew. Now that the first versions of the app have been launched, developments have halted, even though new insights and policies (for instance on face mask wearing) would definitely warrant further updates. This is a common pitfall in every science-driven digital technology project. Most apps and wearables are abandoned as soon as the funding for the RCT runs out (the infamous “plague of pilots” (Huang, Blaschke, & Lucas, 2017)). This means that there is a strong need for viable business model development at an early stage of the project, so there is an opportunity and a driving force to keep it going after the volunteer energy runs out (Korpershoek et al., 2020; Van Limburg et al., 2011) show examples of a development process with attention to business modelling taken into account).

All in all, this project shows that evidence-based intervention development can certainly benefit from modern methods such as Agile. The high pace, responsiveness, interaction, and parallel development all contribute to a positive result. For this method to be really useful in scientific contexts, however, we also need to be aware of the shortcomings, especially in keeping track of the development process, and in ensuring sustained development when the energy or funding from grants or sponsors run out.
3. COMMUNICATING RISK – MORE A SCIENCE THAN AN ART? AMY H Y CHAN

The onset of the COVID-19 pandemic has brought about significant levels of anxiety and uncertainty worldwide, as societies struggle to grapple with new knowledge about the coronavirus, city- and country-wide lockdowns, and new behaviours in line with recommended and mandatory public health measures (Sibley et al., 2020; United Nations, 2020). A key focus has been on the characterising and quantifying the risk of getting COVID-19 and the effect of public health measures on reducing this risk. The “Your COVID-19 Risk” tool was developed to provide individuals an idea of these risks, and what they can do to reduce this. As with any risk prediction model, the most challenging part after the risk model was developed was working out how to illustrate this risk in a way that promotes positive behaviours (Glik, 2007). We know that simply providing information is not effective in changing behaviour (Kelly & Barker, 2016). The key to effectively communicate information – in this case risk information – to achieve sustained behaviour change is to do so in a way that encourages and motivates behaviour change. Risk perception is part of why people act the way they do, but this is not the key aspect of behaviour change – the intervention should also consider and target the determinants of the behaviour (e.g. social norms, habits etc.). Any behaviour is influenced by hundreds of determinants, for example an individual’s motivation, ability and environment (Horne, Cooper, Wileman, & Chan, 2019; Michie, Stralen, & West, 2011). For example, some individuals may not have the financial resources to allow them to follow self-isolation/stay-at-home measures due to pressures to work and earn an income. In contrast, some may not be motivated to follow public health measures due to a lack of perceived need to do so. Identifying an individual’s specific set of determinants that shape their behaviour is a pivotal step towards achieving effective behaviour change.

The science behind risk communication comes from the social and behavioural sciences, and considers target population, messaging and purpose of the information communication. A suggested analysis tool by the World Health Organization (WHO) is the HIC-DARM (Hear, Inform, Convince, Decide, Act, Reconfirm, Maintain) analysis framework, which allows priorities for risk communication to be identified and messages/actions to be targeted to the relevant audience (World Health Organization, 2017, 2020). The framework focuses on ensuring the people who need to know about the information do so in a way that encourages action – they Hear about the behaviour; get Informed about it; are Convinced that it is worthwhile; and then Decide to do something; Act on the new behaviour; Reinforce action by feeling satisfied about participating; and Maintain the behaviour. Global health agencies such as the WHO have also produced guidelines for emergency risk communication (World Health Organization, 2017). These guidelines focus on the practice of risk communication and strategies to adapt to communicate risk effectively in public health emergencies (such as building trust and ensuring consistency of messaging), and provide high-level guidance on what principles effective risk communication should consider (e.g. tailoring information and communication systems to user needs, and ensuring that messages only promote actions that people can take to protect their health). The development team considered these principles and guidelines when designing the risk tool; however, the challenges with these high-level guidelines are that they do not provide specific recommendations on exactly how to illustrate and communicate risk in the confines of a digital, on-screen tool.

For this, we looked to evidence from medical decision-making, patient engagement and medication adherence, such as work related to communicating cardiovascular risk to patients in a way that promotes medication uptake and adherence (Goldman et al., 2006; Timmermans, Ockhuysen-Vermey, & Henneman, 2008). The key principle behind this health risk communication is ensuring an exchange of information about risk in a way that leads to better understanding of the risk in question, thus promoting better decisions about management. In our COVID-19 risk tool, what we wanted to achieve was a persuasive depiction of the risk of getting coronavirus, and the behaviours that would ‘get rid of’ or ‘remove’ the coronavirus (thus reducing the risk of COVID-19). As there were two separate behaviours that needed to be considered, we developed two separate risk estimates showing the two behaviours. The next challenge however was to illustrate different risk levels. Allocating a risk ‘score’ seemed the logical step to take, as many medical risks are communicated in numbers. However, increasing evidence suggests that allocating a number to risk may be a barrier to effective action in some situations (Gigerenzer & Edwards, 2003). Numbers suggest a scale or measurable element of risk, which our COVID-19 risk algorithm did not have — at the time, it would not have been possible (and to date it still isn’t) to distinguish or define the difference between a score of 7/10 risk versus 6/10 risk — what would each scale point of ‘1’ mean? There was not enough information at the time to provide individual-level risk estimates with accuracy, and even now, most of the data that exists are only estimates from between-person, rather than within-person data, leading to large uncertainties in any measurements. There were also concerns that numbers would falsely suggest that the risk of getting and getting rid of coronavirus was quantifiable in a stepped manner — which again is not true. Numbers also allow direct comparisons, which we felt may lead some individuals to be too focused on getting a ‘perfect’ score or a ‘pass’
score, or fuel stigma by suggesting that some scores were ‘better’ than others. We also recognised that numeracy is not a universal attribute — many patients, clinicians, journalists, politicians may have a lack of basic competencies to understand health statistics.

With this in mind, we moved away from the use of a numerical scale to using a ‘heat map’ on a gradient scale (with no numbers) to show a shaded change in risk. In line with avoiding numbered scales, we defined the gradient scale only at the two extreme points (similar to a semantic scale instead of a Likert type scale). We also opted to calculate the score in such a way that we never produced either minimum or maximum values on the scale — no-one was ever at no risk whatsoever (to avoid a false sense of security) and no-one was ever at maximum possible risk (to avoid undue anxiety). Studies have shown that providing fewer data points — even just two points — helps to facilitate information processing compared to using more data points (Lipkus & Peters, 2009; Timmermans et al., 2008). The question then came about what colours to use to show the risk.

From research on colours and communication, and the psychology of colour, we chose to use red arrows to show getting coronavirus onto the body, and the use of green arrows to show getting it off. The meaning of colours differs greatly between cultures, and within cultures between different contexts. As such we went with red as it tapped into the universal metaphor of red indicating stop and green meaning go, and there are no negative cultural connotations against using these two colours.

The ‘stop’ and ‘go’ traffic light analogy aligned with our intended action of encouraging actions to remove or ‘clean’ off coronavirus after exposure. Alongside this, we used a graded map of yellow to dark orange to illustrate the change in risk; use of orange and yellow was chosen as the ‘neutral’ or ‘in-between’ colour between red and green based on the traffic light metaphor. However, the risk graphic is still imperfect as there are still issues of comparisons with ‘pass’ and ‘fail’ messages with traffic light systems. We also considered how colours may be perceived in different cultures, and avoided whites and blacks which are related to death in some cultures.

In line with behaviour change principles, with every risk that was shown, we also ensured that actionable advice was given. This was tailored according to the responses that the individual gave at the risk determination stage through the online questionnaire. The resulting risk estimate is shown in Figure 1. Overall, whilst individual beliefs and backgrounds will always influence the perception of the final risk shown to the individual, we believe we found a way to communicate a complex risk to individuals so that people from most backgrounds will be able to understand and act on the risk. Initial pilot testing showed good consumer acceptability of the risk estimate, though potential areas of confusion relating to understanding the meaning of the arrows and the ‘level’ of coronavirus on the avatar remain, which are considerations for later versions (see the last section by James Green who discusses the practical issues of creating and updating a tool in a rapidly evolving environment).

---

**Figures 1–3** Screenshots of the home page, explanation sections, and example risk estimate from the ‘Your COVID-19 risk’ tool.
4. CROSS-CULTURAL DEVELOPMENT AND TRANSLATION — KEBEDE BEYENE

Culture influences many aspects of our life; as such measurement of risk behaviour is also likely to be influenced by cultural characteristics of the target populations. Thus, if risk behaviour assessment tools are to be used across cultures or languages, rigorous standards are required to ensure equivalence between the original and translated versions of the tools. In the absence of this, the risk assessment tool accuracy and effectiveness for the intended purpose cannot be guaranteed (Eremenco, Cella, & Arnold, 2005).

Translating the COVID-19 risk assessment tool from English into other languages was challenging. I was involved in translating the tool to one of the South Semitic languages known as Amharic. Amharic is widely spoken in Ethiopia, and it is the official national language of the country (Meyer, 2006). Some of the problems we had during translation included not having the same term in Amharic and having to use several words to sustain the original English meaning. Differences in sentence or grammar structure, verbal nuances, and tense are some of the other issues we encountered. Amharic language follows subject-object-verb (SOV) grammatical pattern, similar to Asian languages such as Korean and Japanese, as opposed to English language which has SVO sequence of words. This has created some issues in translating questions written as short incomplete phrases (e.g. I prefer ... “Being much more relaxed”/“Being much more excited”). The use of long sentences for short English expressions was another challenge. This is particularly problematic as COVID-19 is a new disease and several terminologies associated with the disease did not have equivalent Amharic words at the time of translation. Another difficulty was with the translation of the response scales, for example, in English “very probable” and “very likely” have subtle differences in meaning. However, both response scales have the same meaning in Amharic. Translating words, phrases, and concepts that do not have clear equivalence in Amharic was also a challenge. “Your risk of getting coronavirus particles on you”, “Your risk of keeping the coronavirus particles on you”, and “Your COVID-19 Risk” are few of such examples. Moreover, there was hot debate among team members as to whether to include or exclude some culturally insensitive risk behaviour assessment items.

Most of the team members involved in forward and backward translations of the COVID-19 risk tool were highly educated and at least bilingual. For example, all the members of Team Ethiopia are bilingual and trained in Western countries. As a result, all of us adopted some of the values and attitudes of the Western society, and we may not represent monolingual Amharic-speaking Ethiopians. Translations that are produced by highly educated health professionals could sometimes be complicated and may not be easily understood by less educated or disadvantaged members of the society. This has a potential to introduce inequity. The WHO recommends at least the back translation to be performed by an independent translator whose mother tongue is English and who has no knowledge of the original risk assessment tool (World Health Organization, 2009). However, given the urgency of the work and the agile development nature of the tool (see section 2), this was not a viable option for several teams involved in the development of the COVID-19 risk assessment tool, including the Ethiopian team. Ensuring the quality of risk assessment tool translation is very important to collect comparable data across nations. Poor translations could lead to measuring concepts that were not intended to be measured, and the data collected by poorly translated tools often reflect systematic errors rather than meaningful behavioural differences between target populations. Overall, “Your COVID-19 Risk” is a novel tool and timely, but subsequent development of the tool should pay attention to the validity of translation and cross-cultural validity of the tool. The tool is well designed; however, it still requires refining to make it more accessible by people with physical or intellectual disabilities and populations vulnerable to COVID-19 (e.g. indigenous people and ethnic minorities). This could maximise response rate and usability of the tool.

5. REFLECTIONS FOR THE FUTURE — JAMES GREEN

As Sander Hermsen noted above, we had originally hoped to begin work on a version two and even three of the tool. This would have incorporated the feedback we had from participants, our own self-reflections, and new behaviours associated with the preventing the transmission of COVID-19 such as wearing masks. For those of us in academic roles, preparing for the delivery of our teaching in radically different circumstances, as well as other tasks became more pressing. The declining case numbers through July and August will have also made it seem ‘less urgent’, and our initial enthusiasm waned.

Despite our agile — for academics — development of the tool, initial timeline estimates of around three weeks blew out to something like two months. Although we were agile, perhaps we were not agile enough. A key issue is that while much work was completed in parallel, some technical parts of the project pipeline, such as the development of the Limonaid package, were dependent on only one person/legend. By the time we launched, the first peak was receding in many countries. Delays were also in part due to scope creep (like more languages!), things inevitably taking longer than planned, but also some real disagreements that needed to be resolved.
A key example was around the determinant mapping questions, and particularly whether the wording of some questions was culturally acceptable across all of the nations involved. Kebede mentions the internal discussions within Team Ethiopia above. This discussion primarily played out on Slack, but there was also some outside discussion, though that is harder to quantify. Ultimately a mutually acceptable solution was found, but it caused around two weeks delay.

Related to that, I had some concerns about relational asymmetries. A number of us know each other really well and have established relationships. But for others, who came into the project perhaps knowing one person, having been brought in for specific technical or language expertise, they might feel less able to speak up. There are also inherent hierarchies within academia (we ranged from full Professors to PhD students to health professionals). There are also potential cultural differences (some cultures are more direct, others more conflict avoidant) that may have impacted on effective working within the team. The social space enabled us to get to know some of the new participants, but this by no means resolves all issues described here.

An amazing element of the project for me was the breadth of skills in the team, and the depth which elements could be considered. For example, if I'm doing a project, I have some colours that I like and some typefaces that I tend to use. Here, however, we had an entire Slack channel discussing multiple proposals for colour palettes and fine detail of lettering, culminating in something truly-professional looking (see e.g. Figure 1 – the risk estimate picture earlier). Similarly, if I design some questions for a project, I might have input from 2–3 people, rather than a large number of highly skilled people honing wording for clarity, accessibility, intent and so forth.

Overall, it was a remarkable project. I cannot immediately think of anything comparable where such a large diverse team came together to pull together a quasi-academic project in such short timeframes, in response to an evolving and pressing global issue. Included in the team also were people with design and technical skills, which would not normally be intimately involved with such a project. This was a fine example of the successes but also pitfalls of such a process — importantly though we made lifelong connections (well we hope they are!) that will set us in good stead for future collaborations and have templated an approach that might be adopted for other future collaborations/global responses. As researchers we all want to measure outcomes and from our work here — the real measure of success may not just be the tool that we produced but the journey that we took to get there.

NOTES

1 Free (as in beer) is an old joke to differentiate between things that you are able to use for free (here Slack, but also products like Gmail) from things that are free and open to adapt and use as you would like. The difference between being given some beer or a beer recipe, perhaps. With the recipe, you can change and adapt it, whereas with free beer, you are reliant on someone else continuing to give it to you, and you get the beer that they give you.

2 'Born Open' data refers to data that is anonymous on capture, and immediately publicly archived. The data from this project never captured referrers, IP Addresses or any of the usual trackers that many websites have. Updated data was made available online at regular intervals rather than immediately (hence only semi-‘Born Open’).

COMPETING INTERESTS

The authors have no competing interests to declare.

AUTHOR CONTRIBUTIONS

All authors contributed equally to this manuscript.

AUTHOR AFFILIATIONS

Kebede Beyene  
orcid.org/0000-0002-1057-2593  
School of Pharmacy, The University of Auckland, NZ

Amy Hai Yan Chan  
orcid.org/0000-0002-1291-3902  
School of Pharmacy, The University of Auckland, NZ

James A. Green  
orcid.org/0000-0002-7309-0751  
School of Allied Health and Physical Activity for Health (PAfH), Health Research Institute, University of Limerick, IE

Sander Hermens  
orcid.org/0000-0001-8781-5445  
OnePlanet Research Center, Wageningen, NL

REFERENCES

Cooper, A. (1999). The inmates are running the asylum. Indianapolis, IA: SAMS, Macmillan. DOI: https://doi.org/10.1007/978-3-322-99786-9_1


