

'Your DNA, Your Say': global survey gathering attitudes toward genomics: design, delivery and methods

Anna Middleton^{*,1,2}, Emilia Niemiec³, Barbara Prainsack⁵, Jason Bobe⁶, Lauren Farley¹, Claire Steed⁷, James Smith⁷, Paul Bevan⁷, Natasha Bonhomme⁸, Erika Kleiderman⁹, Adrian Thorogood⁹, Christoph Schickhardt¹⁰, Chiara Garattini¹¹, Danya Vears¹², Katherine Littler¹³, Natalie Banner¹³, Erick Scott⁶, Nadezda V Kovalevskaya¹⁴, Elissa Levin¹⁵, Katherine I Morley^{1,16,17} & Heidi C Howard^{1,3}

¹Society & Ethics Research, Connecting Science, Wellcome Genome Campus, Cambridge, UK

²Faculty of Education, University of Cambridge, Cambridge, UK

³Centre for Research Ethics & Bioethics, Uppsala University, Uppsala, Sweden

⁴Department of Political Science, University of Vienna, Vienna, Austria

⁵Department of Global Health & Social Medicine, King's College London, London, UK

⁶Department of Genetics & Genomic Sciences, Icahn Institute, Icahn School of Medicine at Mount Sinai, NY 10029, USA

⁷Web Team, Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, UK

⁸Expecting Health, Genetic Alliance, Washington, DC 20008-2369, USA

⁹Centre of Genomics & Policy, McGill University, Montreal, Quebec, Canada

¹⁰Department of Medical Oncology, National Center for Tumor Diseases, Heidelberg University Hospital, Heidelberg, Germany

¹¹Health & Life Sciences, Intel Corporation, Swindon, UK

¹²Department of Public Health & Primary Care, Centre for Biomedical Ethics & Law, KU Leuven, Leuven, Belgium

¹³Policy Team, Wellcome, London, UK

¹⁴Repositive Ltd, 104 Hills Road, Cambridge, UK

¹⁵Policy & Clinical Affairs, Helix OpCo LLC, San Francisco, CA 94158, USA

¹⁶National Addiction Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, UK

¹⁷Centre for Epidemiology & Biostatistics, Melbourne School of Global & Population Health, The University of Melbourne, Melbourne, Australia

*Author for correspondence: am33@sanger.ac.uk

Our international study, 'Your DNA, Your Say', uses film and an online cross-sectional survey to gather public attitudes toward the donation, access and sharing of DNA information. We describe the methodological approach used to create an engaging and bespoke survey, suitable for translation into many different languages. We address some of the particular challenges in designing a survey on the subject of genomics. In order to understand the significance of a genomic result, researchers and clinicians alike use external databases containing DNA and medical information from thousands of people. We ask how publics would like their 'anonymous' data to be used (or not to be used) and whether they are concerned by the potential risks of reidentification; the results will be used to inform policy.

First draft submitted: 28 March 2018; Accepted for publication: 25 April 2018; Published online: 1 June 2018

Keywords: attitudes • data sharing • film • genetics • genomics • metaphor • public • public engagement • science communication • survey

The background, context and subject matter of this study has already been published [1,2]. This Special Report contains methodological details pertaining to the development and delivery of a cross-sectional online survey, called 'Your DNA, Your Say' (YDYS) to gather public attitudes toward genomic data donation, access and subsequent sharing.

Why use an online survey?

Online or web-based surveys are increasingly being used in epidemiological and social science research [3,4]. This format permits greater response time and scope than traditional paper-based surveys, and can be particularly effective

for accessing populations often missing from 'traditional' survey methods (e.g., younger sections of population who are less likely to have fixed phone lines and more likely to change home address) [5–7]. The cost of web-based surveys may be less than 10% of paper-based methods (Ebert *et al.* 2018). It can also increase data quality as the absence of an interviewer means this approach is less likely to be affected by social desirability bias, and direct entry of responses by participants along with integrated real-time validation and consistency checks reduces the likelihood of data entry errors [5–7].

Survey design

The Participant Values Task Team (PVTT) is an active group within the Regulatory and Ethics Work Stream of the Global Alliance for Genomics and Health (www.ga4gh.org/aboutus/). The PVTT comprises an international group of researchers (both nonprofit and for-profit), policy makers, clinicians and patient representatives all with expertise in scientific, legal or social aspects of clinical and/or research genomics. The PVTT had a teleconference call in 2015 where a collective decision was made to conduct an empirical study on global public attitudes toward genomic data sharing, using a cross-sectional, online survey as the tool. This methodological approach was guided by practicality in terms of gathering large-scale attitudes across languages, and between countries and cultures. The survey items were designed primarily by A Middleton, HC Howard and E Niemiec with iterative input from the rest of the PVTT.

Figure 1 provides an overview of the design process to create the survey and films. A literature review was conducted within PubMed, Scopus and Web of Science in 2015 to explore existing empirical work on genomic data sharing. With a focus on discussion pieces as well as public attitude studies, we were also particularly interested in the richness of qualitative data gathered on the subject, to inform our survey items. We found no published work on our exact area of interest that used validated items. Thus, we used an iterative design method to create our own exploratory survey. Consensus was reached on the survey topics via two teleconference calls with the PPVT. Three members of the PVTT (A Middleton, HC Howard and E Niemiec) then met in person to spend 1-month full time work in February 2016 writing the survey questions. These three members have expertise in survey design, genomic data sharing, application of genomic technology in clinical and research settings and bioethics. HC Howard and E Niemiec are also fluent in five languages between them. Creation of the novel, bespoke survey items followed a democratic process, using five basic principles: are the questions congruent with the topics of interest?; Do the questions make sense?; Do the questions appear to gather what we want to study?; Do the questions meet readability scores indicating they are easily understood by participants with a reading age of up to 15 years?; And will the questions be easily translatable into other languages?

Drawing on previous experience in delivery of public surveys about genomics [8], we chose to use a film–survey combination for this study. Film was used as the mechanism to explain what participants needed to know about the context of genomic data sharing, in a manner that would be engaging and interesting. What is novel about this project is that the films are used as a method to build a bridge between the research participant and the subject matter. Empirical research tells us that there are significant gaps in public knowledge and understanding of genetics and genomics [9] and due to this the subject of genomics is often seen as 'not for me' because the topic is perceived as a topic only understood by 'experts' [10]. We thus created the films deliberately to address this, demystifying the subject matter, using metaphors [11] and lay language.

We conducted three sets of face validity testing with experts in genomic data sharing from the Global Alliance for Genomics and Health to ensure that the survey items were deemed appropriate to the subject matter and able to collect data on the issues of importance. We conducted readability testing at two time points using readability software and adapted the survey items so that the reading level was <15 years [12]. We also conducted pilot testing at three time points with 45 participants with expertise in public engagement, genomic data sharing and also 30 nonexperts (broader public). The survey underwent iterative improvement after each intervention. The body of the survey contains a glossary: each time a technical or scientific term is used (e.g., DNA information, medical information, database etc.) this term was underlined and a glossary description would appear on the screen if the user hovered their mouse over the term. Creating the descriptions for the glossary was achieved through an iterative process between the survey design team. Where disagreements occurred, the language was discussed and fine-tuned until consensus was reached.

The survey contains nine sections where participants can enter data (Figure 2). Details of the sections of the survey have been published previously [2] and the actual survey can be accessed at: (www.YourDNAYourSay.org). Participants complete the survey anonymously and all socio-demographic questions were deliberately designed so

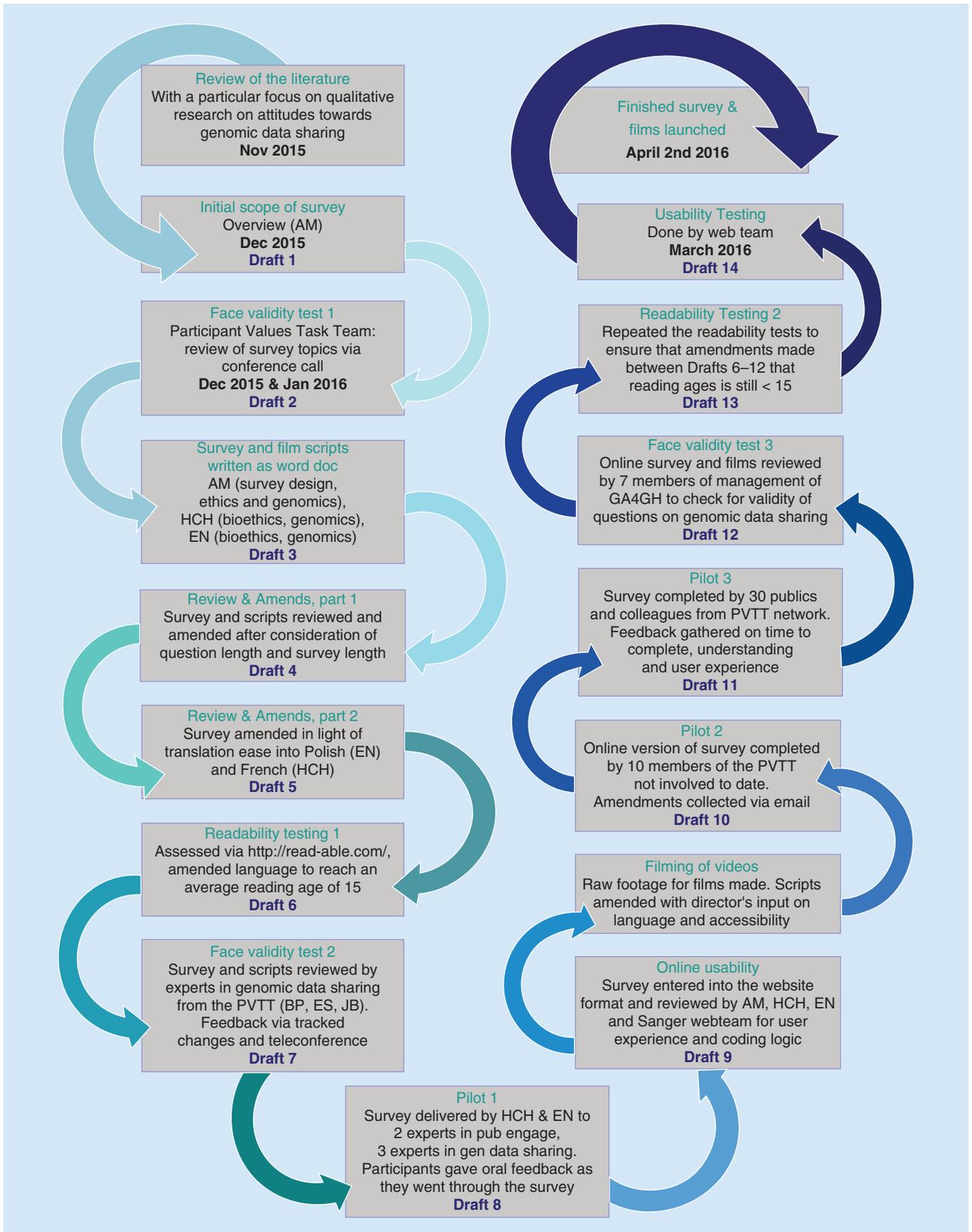


Figure 1. Summary of survey design process.

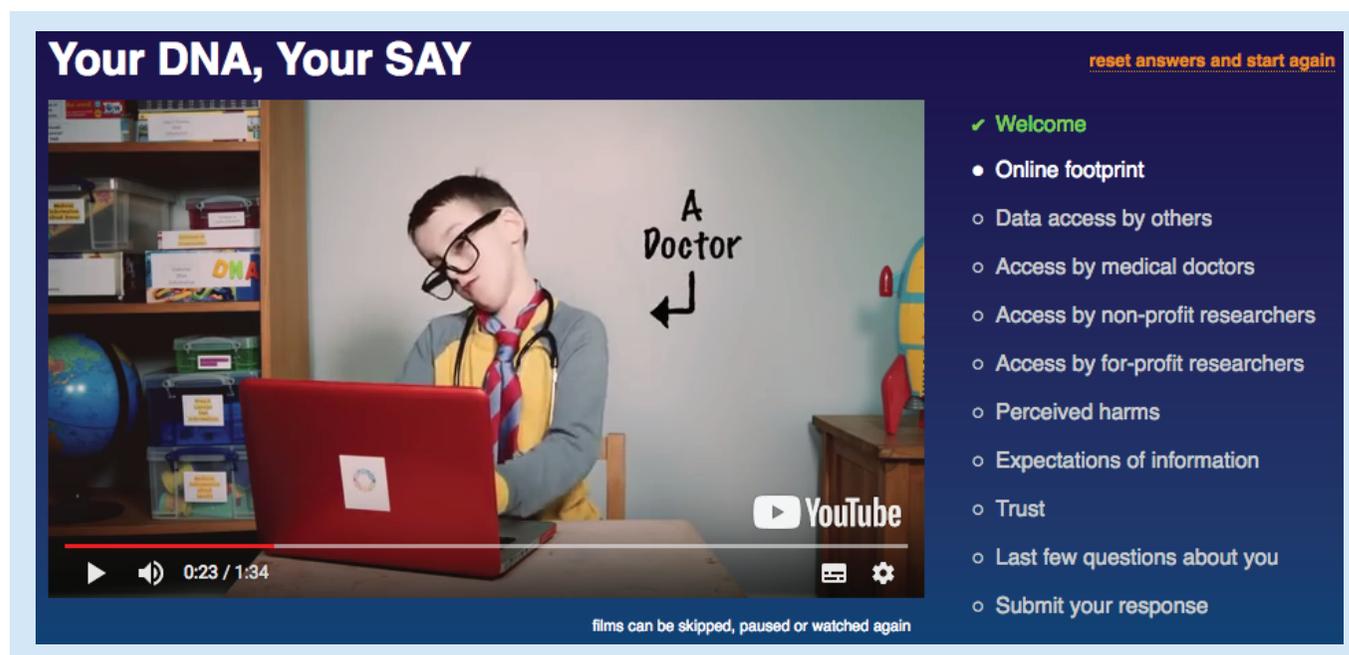


Figure 2. Survey.

that individuals could not be identified. IP addresses were collected to allow participants to partially complete the survey and return to it at a later date (IP is Internet Protocol and identifies the geographical location of the computer). These were deleted from the dataset used in analysis so it was not possible to identify the computer that participants used. Thus, the survey is considered anonymous.

As per feedback from the pilot process, we designed the majority of the questions to be closed, with fixed response choice options. The fixed options were based on the expertise of the authors, their own research findings from related fields and the literature scoping exercise. The finished survey had 29 questions and took on average of 15–20 min to complete.

Film design

Nine films were scripted, using visuals and voice over, to explain what genomic data sharing is. See Supplementary Material for an example of one of the explanatory films.

Our previous empirical research has shown that the vast majority of British publics have not heard of the word ‘genomics’ [13], and this is likely to be indicative of knowledge and understanding in other countries as well – as we know that public understanding of genomics is low [9,10]. There is a real need to ‘socialize’ genomics for publics [14] so that they can make informed choices about how and whether to engage with genomic technologies. Thus, in the films and survey we use the term ‘DNA information’ instead of ‘genomics’, as we were aware that this was a term more widely understood.

The films were designed to have an open, nonthreatening and positive feel, [15] using a child as the main character. Under the guise of a medical doctor, a nonprofit scientist and a for-profit scientist, he describes what genomic data sharing entails. His playroom sets the scene as a metaphor for the internet – shelves of toys representing databases containing DNA data that are accessed by researchers. The use of metaphor has shown to be a valuable and consistent method of communication about genomics [16]. The films provide the background and contextual information that is needed to answer the questions. Using narrative to explain science is well established: ‘research suggests that narrative communication is encoded using a unique cognitive pathway’ [17] and narrative communication has been shown to improve comprehension and generate interest and engagement [18]. Scholars have also shown that film itself has the capacity to transport viewers into a narrative [19] and we utilized the films as an engagement tool to generate and maintain interest in the survey [15]. Participants could click on different sections of the survey and see a new film in each section. Attention was given, within the script development, to ensure the information in the films was delivered in as neutral a manner as possible, so as not to bias participants’ answers. As a testament to their

broad public appeal, the films have been awarded official selection status at the following film festivals: Raw Science Film Festival 2016 (Los Angeles, October 2016); Viten Film Festival 2016 (Bergen, Norway, November 2016); New Filmmakers New York 2017 (New York City, January 2017); and International Freethought Film Festival 2017 (Orlando, Florida, May 2017).

We assumed that participants knew nothing about the subject matter on survey commencement. We also assumed (and will measure in future studies) that the films helped participants to develop a connection with the subject matter. With this in mind, the film–survey combination was designed in such a way so that the survey respondent could evolve their thinking as they processed through the survey. It is thus a public engagement tool as much as a social sciences method to aid attitude capture.

Survey translated into other languages

In order to reach larger and more diverse audiences of non-English speakers, the survey and accompanying films were translated into a number of languages. Translations were provided by native speakers who were experts in genomics and/or related fields known to the PVTT. Throughout the process, translations were checked with other native speaker(s) of the language, where possible, and problematic issues were discussed and resolved. While the translations were aimed to be highly consistent with the English version, the wording used was intended to be understandable to general audiences for each language. The final versions of translations were back translated to English by a person not yet familiar with the original English version to provide a validation step and ensure that the translation was adequate. Back translations were then compared with the original English version by the main translator. Any discrepancies between the versions that introduced changes in meaning were reviewed and translations were amended. The translated survey and films were made available online and further rounds of reviewing and amendments were made through user testing among the collaborating team, before being available for participant recruitment. At the time of article submission, the survey has been translated into, and is recruiting participants in Russian, Polish, French, German and Portuguese, with translations at various stages of completion for the following: Spanish, Italian, Japanese, Chinese (Mandarin), Urdu, Hindi, Zulu, Arabic, Icelandic and Swedish.

Recruitment

This survey aimed to be as inclusive as possible and thus participants could be of any age, gender, socio-demographic background, ethnic group and nationality worldwide. The only limiting factor is access to an internet connection and ability to read and understand the questions. In order to understand more of the profile of people completing the survey, one of the questions identifies if people had any previous knowledge or experience of genomics, either through being a patient, working in the industry, a customer of a direct-to-consumer genetic test, a research participant or having a family history of an inherited condition.

Recruitment for the English-language version started in 2016 and non-English versions in 2017 (likely to continue through 2018–2019). As the project was being managed by multiple collaborators around the world, each local team decided upon the most appropriate methods for recruitment of their own population. Where funds allowed, we used the market research company, ResearchNow (via the London office, with global reach to publics worldwide) to recruit a representative sample of participants from the UK, Germany, USA, Canada, Australia, France, Switzerland, Belgium, Portugal, Poland and Russia. Using this service, participants are offered participation in return for a small financial reward. Careful consideration is given to ensuring participants are not over sampled and become professional survey takers (thus potentially biasing their participation), see ResearchNow for details of sampling frameworks, representativeness and data quality. The majority of surveys were recruited using this sampling technique. However, for some translations, where it was not possible to use a market research company, a snowballing method of recruitment was used, which was pertinent to the local market. For example, we added relevant social media buttons to the survey for individual languages, in other words, social media in China is different from social media in the UK and thus the delivery of these were country-dependent. This created an opportunity for collaborators to advertise the survey to their local populations and make use of social media reach.

As of May 2018, 20,000+ completed surveys from across the world have been received from 25+ different countries. The rationale for our methodology was to reach as many publics from across the world as we could, given the practical and logistical limitations of an online survey and the practical limitations of the recruitment network we could access. We aim to have recruited some publics from every continent across the world and while no study can ever claim to have 'representative global views' our objective is to explore a snapshot of public views at a single time period, from a breadth of different countries. Collaboration on the project was achieved through

a snowballing mechanism of networking among the PVTT, GA4GH, all encouraging respected and established researchers around the world to translate and deliver the survey within their own countries. The management of the project and collaborations was overseen by the Society and Ethics Research Group, Wellcome Genome Campus in Cambridge, UK.

Survey data storage, protection & privacy

Each of the collaborators on the project will determine their own recruitment strategy and also framework for analysis of data. In the English-speaking survey (involving representative publics from UK, USA, Canada and Australia) we have completed surveys from <10,000 participants and have only used surveys in the analysis from participants recruited via ResearchNow (thus removing ~300 participants who found the survey online, separately from our recruitment strategy). By using only participants from ResearchNow we were able to keep the sample as cleanly 'representative' of a public sample, but it also meant that we had fully completed surveys in our final dataset and did not have to remove surveys that were partially completed.

The YDYS survey is hosted at the Wellcome Sanger Institute, Wellcome Genome Campus, Cambridge, UK. The data within it are protected, encrypted and stored as per the Sanger Institute's data protection policies. Consent to participate in the survey was deemed implicit if participants started answering the questions. Some of the collaborators (e.g., in Australia and South Africa) needed to obtain Human Research Ethics Committee approval before becoming involved in the study, and so for them, an additional consent process was added onto the landing page of the survey, where participants could click on a button that would take them to a consent form to complete before starting the survey. For social sciences researchers at the Wellcome Genome Campus gathering anonymous attitude data from publics, no Human Research Ethics Committee approval was required.

Limitations

At the time of survey design, no previous studies had been published on the exact topic of interest (i.e., surveys designed to gather attitudes from a nonexpert, global public audience). This meant that no previously validated survey items existed for use in the survey. Given the absence of prior research in this area, this study is hypothesis generating as opposed to experimental, and our results should only be considered exploratory. Consequently, no power calculations were undertaken prior to participant recruitment, and all results would benefit from confirmation via further studies designed to test specific hypotheses generated by the study. Likewise, the YDYS survey was never designed to be longitudinal. It captures a flavor of attitudes at a single time period and our results should always be considered in the context of other social sciences work, using different methodological approaches on the same subject.

Despite the advantages of using online surveys to gather extensive data, like all research methods, they also have some important limitations. They have previously been criticized for having low response rates, or unknown response rates where the sampling frame is undefined, and recruiting nonrepresentative samples due to disparities in internet access [5–7,20]. This may be counterbalanced by the capacity to recruit much larger samples than conventional approaches or, as is the case with YDYS, using market research strategies to recruit a representative sample of a defined size. The reliability and validity of surveys administered via the web has also been questioned, but research to date suggests that potential problems in this area relate mainly to clinical measures, such as diagnostic questionnaires, rather than direct reporting of characteristics or opinions [5].

The films we created were designed as a public engagement tool, replacing what could have been delivered via text, had this been a paper-based survey. Given the high cost of film development it was not possible to change the films after they were made. Therefore, we chose not to evaluate the finished films in any way (as we would not have been able to change them). The input into design of the films happened before they were made, the scripts were piloted along with the survey items to ensure they were as unbiased and informative as they could possibly be. Future studies would benefit from an evaluation of the films, measuring whether they did indeed foster engagement and provide the necessary background information of use to participants.

Conclusion

This methodological paper describes how our cross-sectional online survey, 'Your DNA, Your Say', was designed. The survey uses a novel film-survey approach where innovative films offer public engagement around genomic data sharing, providing enough background information so that participants can answer the questions. Genomics is often seen as a 'difficult' subject for non-experts and so we paid specific attention to 'socializing' the subject

matter via film, so that anyone, irrespective of their knowledge or understanding, could participate in the research. Global publics are invited to offer their attitudes towards genomic data donation, access and subsequent sharing. The survey has already been translated into 14 other languages, with more being planned. To date, this project is on track to be the largest study ever on global public attitudes towards genomic data sharing.

Future perspective

Given the rate and increase of public exposure to multimedia, survey research needs to keep pace with the modes of engagement to which publics are becoming accustomed. We anticipate that social sciences research will evolve to incorporate more multimedia within its design and delivery. An ideal advancement would also be to have the ability to capture and code video responses from participants as they engage with a survey. The survey methods (and indeed the results from the survey) will inform how we as a society engage in discourse about genomics.

Executive summary

- This methodological paper describes the design process of the 'Your DNA Your Say' survey which gathers attitudes of global publics toward genomic data donation, access and subsequent sharing.
- The survey uses a novel film–survey approach where innovative films offer public engagement around genomic data sharing and provide enough context to the subject matter so that participants can answer the questions.
- We have addressed the challenges surrounding genomics as a topic for engagement using film and a novel metaphor.
- The survey has been translated into more than 14 different languages, with a view to gathering international attitudes from global publics.

Open access

This work is licensed under the Attribution-NonCommercial-NoDerivatives 4.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Financial & competing interests disclosure

The authors have no relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript. This includes employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties.

No writing assistance was utilized in the production of this manuscript.

Supplementary data

To view the supplementary data that accompany this paper please visit the journal website at: www.futuremedicine.com/doi/full/10.2217/pme-2018-0032

References

Papers of special note have been highlighted as: • of interest; •• of considerable interest

1. Middleton A. Society and personal genome data. *Hum. Mol. Genet.* 27(R1), R8–R13 (2018).
 2. Middleton A. Your DNA, Your Say. *New Bioeth.* 23(1), 74–80 (2017).
 3. Van Gelder M, Vorstenbosch S, Derks L, Te Winkel B, Van Puijtenbroek EP, Roeleveld N. Web-based questionnaires to assess perinatal outcome proved to be valid. *J. Clin. Epidemiol.* 90, 136–143 (2017).
 4. Debin M, Turbelin C, Blanchon T *et al.* Evaluating the feasibility and participants' representativeness of an online nationwide surveillance system for influenza in France. *PLoS ONE* 8(9), e73675 (2013).
 5. Van Gelder MM, Bretveld RW, Roeleveld N. Web-based questionnaires: the future in epidemiology? *Am. J. Epidemiol.* 172(11), 1292–1298 (2010).
 6. Barratt MJ, Ferris JA, Lenton S. Hidden populations, online purposive sampling, and external validity: taking off the blindfold. *Field Methods* 27(1), 3–21 (2014).
 7. Fripiat D, Marquis N. Web surveys in the social sciences: an overview. *Population* 65(2), 285–311 (2010).
 8. Middleton A, Bragin E, Morley KI, Parker M. Online questionnaire development: using film to engage participants and then gather attitudes towards the sharing of genomic data. *Soc. Sci. Res.* 44C, 211–223 (2014).
- **Our previous work on survey design about public attitudes toward genomics.**

9. Lea DH, Kaphingst KA, Bowen D, Lipkus I, Hadley DW. Communicating genetic and genomic information: health literacy and numeracy considerations. *Public Health Genomics* 14(4-5), 279–289 (2011).
 - **Helpful overview of genomic literacy among the general public.**
10. Boersma R, Gremmen B. Genomics? That is probably GM! The impact a name can have on the interpretation of a technology. *Life Sci. Soc. Policy*. 14(1), 8 (2018).
 - **Interesting overview of the way the genomic terminology can be interpreted by the public.**
11. Avise JC. Evolving Genomic metaphors: a new look at the language of DNA. *Science* 294(5540), 86–87 (2001).
 - **Offers a very useful overview of the way metaphors can help with communication about genomics.**
12. Cheng C, Dunn M. Health literacy and the Internet: a study on the readability of Australian online health information. *Aust. N. Z. J. Public Health* 39(4), 309–314 (2015).
13. Middleton A. Socialising the genome: making genomics resonate [version 1; not peer reviewed]. *F1000Research* 7, 149 slides doi:110.7490/f1000research.1115249.1115241 (2018) (Epub ahead of print).
 - **An overview of the societal and ethical issues raised by genomic data sharing, offering a review of the subject matter behind our methodological paper.**
14. Parry V, Middleton A. Socialising the genome. *Lancet* 389(10079), 1603–1604 (2017).
15. Jonah Berger, Milkman KL. What makes online content viral? *J. Mark. Res.* 49(2), 192–205 (2012).
16. Stelmach A, Nerlich B. Metaphors in search of a target: the curious case of epigenetics. *New Genet. Soc.* 34(2), 196–218 (2015).
 - **Offers a very useful overview of the way metaphors can help with communication about genomics.**
17. Dahlstrom MF, Ho SS. Ethical considerations of using narrative to communicate science. *Sci. Commun.* 34(5), 592–617 (2012).
18. Dahlstrom MF. Using narratives and storytelling to communicate science with nonexpert audiences. *Proc. Natl Acad. Sci. USA* 111(Suppl 4), 13614–13620 (2014).
19. Nomura R, Hino K, Shimazu M, Liang Y, Okada T. Emotionally excited eyeblink-rate variability predicts an experience of transportation into the narrative world. *Front. Psychol.* 6, 447 (2015).
20. Richiardi L, Pizzi C, Pearce N. Commentary: representativeness is usually not necessary and often should be avoided. *Int. J. Epidemiol.* 42(4), 1018–1022 (2013).