Check for updates

OPEN LETTER

The Genetic Sphygmomanometer: an argument for routine genome-wide genotyping in the population and a new view on its use to inform clinical practice [version 1; referees: 3 approved, 1 approved with reservations]

Nicholas John Timpson [ID] [1], Frank Dudbridge[2]

[1]MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
[2]Department of Health Sciences, University of Leicester, Leicester, UK

Abstract

Initial genomewide association studies were exceptional owing to an ability to yield novel and reliable evidence for heritable contributions to complex disease and phenotype. However the top results alone were certainly not responsible for a wave of new predictive tools. Despite this, even studies small by contemporary standards were able to provide estimates of the relative contribution of all recorded genetic variants to outcome. Sparking efforts to quantify heritability, these results also provided the material for genomewide prediction. A fantastic growth in the performance of human genetic studies has only served to improve the potential of these complex, but potentially informative predictors. Prompted by these conditions and recent work, this letter explores the likely utility of these predictors, considers how clinical practice might be altered through their use, how to measure the efficacy of this and some of the potential ethical issues involved. Ultimately we suggest that for common genetic variation at least, the future should contain an acceptance of complexity in genetic architecture and the possibility of useful prediction even if only to shift the way we interact with clinical service providers.

Keywords
Genetic association studies, GWAS, prediction

This article is included in the Avon Longitudinal Study of Parents and Children (ALSPAC) gateway.

Open Peer Review

Referee Status: ✔ ? ✔ ✔

	Invited Referees			
	1	**2**	**3**	**4**
version 1 published 31 Oct 2018	✔ report	? report	✔ report	✔ report

1 **Nicholas G. Martin**, University of Queensland, Australia

2 **David J. Porteous** [ID], University of Edinburgh, UK

3 **Zoltan Kutalik**, Swiss Institute of Bioinformatics, Switzerland

4 **Dorret I. Boomsma** [ID], Vrije Universiteit Amsterdam, The Netherlands

Brandon Johnson, Avera McKennan Hospital & University Center, USA

Any reports and responses or comments on the article can be found at the end of the article.

Corresponding author: Nicholas John Timpson (N.J.Timpson@bristol.ac.uk)

Author roles: Timpson NJ: Conceptualization, Writing – Original Draft Preparation; **Dudbridge F**: Conceptualization, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

How to cite this article: Timpson NJ and Dudbridge F. **The Genetic Sphygmomanometer: an argument for routine genome-wide genotyping in the population and a new view on its use to inform clinical practice [version 1; referees: 3 approved, 1 approved with reservations]** Wellcome Open Research 2018, **3**:138 (https://doi.org/10.12688/wellcomeopenres.14870.1)

First published: 31 Oct 2018, **3**:138 (https://doi.org/10.12688/wellcomeopenres.14870.1)

Currently most genetic association studies are undertaken using common variants which tend to have small relative risks for common diseases. This prompts an approach to characterising the genetic contribution to disease through use of aggregate scores of risk alleles which can explain non-trivial proportions of variance in disease risk. Indeed the majority of common complex diseases and phenotypes are likely to be influenced by hundreds, if not thousands of common variants of small effect scattered across the genome[1]. This means that there is information in the mass of common variants that individually are not considered to be statistically significant in a genome-wide association study. Allelic scores constructed from thousands of common variants (usually single-nucleotide polymorphisms, SNPs) across the genome can be used even without explicit knowledge of the mode of contribution of each variant, as the aggregate score uses all common genotypic contributions to explain the variance in the trait concerned[2–4]. Taken together, aggregate genetic risk can be one of the strongest single risk factors for a common disease, given the substantial heritability of most traits[5].

We should also bear in mind that most genetic association studies for complex traits are underperforming. Nearly all genetic associations are discovered through the analysis of data collected in cross-section. In most instances, prevalent cases or spot measurements are used in studies systematically hunting for correlations between phenotypes and SNPs. However, whilst it is known that genetic data is temporally stable (for the most part), the same cannot be said for phenotypic assessment. Consequently, the failure to measure outcomes adequately (i.e. in true longitudinal, or incident fashion) will at the very least compromise analytical power and cloud the inference from genetic association. Furthermore, it is increasingly clear that when longitudinal data are assessed, quite different patterns of association and variance explained can be revealed[6,7]. At certain points in the life course, then, the genetic risk of an individual may well be much more predictive than initially suggested by available association studies.

So if we are racing towards a world of improved genotyping scale and better studies of phenotypic association, what should we do with this information? The direct use of human genetic data in directing pharmaceutical development[8] and in causal inference analyses[9] has had a marked impact. However, if we regard genetic data as another tool of measurement in the hands of clinicians attempting to describe or predict the biological condition of an individual – the "genetic sphygmomanometer" – a natural question is: what is there to be gained from adding this particular information to the history taking process? It is well known that predicting the health trajectory of an individual is made enormously complicated by the idiosyncrasies of patients and the events of chance[10]; so can genetic data be of any use? The answer must ultimately be yes as a more precise measurement of family history, but this does not have to be considered wholly via conventionally defined causal pathways – the detection of which suffers a natural, genetic architecture versus analytical power limited, aggregate predictive ability.

Taking the example of cancer, it is possible to generate a genome-wide aggregate score, from birth, which explains some of the

risk of presenting with a malignant neoplasm later in life[11–13]. This is, of course, a combination of both direct pathway effects, but also relationships derived from the heritable contributions to any predicting factor – causal or otherwise. Done for each individual in a population of common ancestry to discovery studies educating the location of genetic risk, this would allow any one person to be located in a distribution of scores summarising the common genetic contribution to cancer risk. Position in this distribution is of course not a reason for immediate intervention with, say, a regime of growth factor inhibitors from the moment of full maturity, but it may offer information that could warrant something as simple as a more regular check-up regime through life. Such a non-invasive subject specific augmentation of basic clinical consultation could be the only intervention, but one which over thousands of individuals would yield life extending marginal returns at a rate greater than chance. The proposal here is that the routine incorporation of genetic data into clinicians' hands will provide an informative addition to the standard tool kit.

Such a mechanism would be possible given two core components. The first is a centralised repository of the results of the largest, most well conducted genome-wide association studies. This would include as many disease outcomes as possible, though would only include studies of suitable rigor and with contributory information. Early versions of such repositories have already been published[9,14–16]. Genetic epidemiologists have a growing collection of formidable meta-analyses that summarise effects by variant across the genome for hundreds of thousands of individuals. Made central, they could be easily accessed to provide disease specific genetic scores for any individual, which could then be used to assign that person to an augmented regime of non-invasive, lifecourse, follow-up.

Secondly, to achieve population scale benefits this approach requires all members of the population to have a comprehensive screen of common genetic variants which could be matched to the reference panel "du jour" in order to provide the material for generating individual aggregate risk scores. The actual use of these, from a consent point of view, would be the sole decision of the genotyped party (or executive) and these data could be linked to electronic health records for further clinical investigation if so required later in life or as the value of these data changes with technological advance.

A key element of this use of genetic data at population scale is the likely cost implication of personalised courses of check-ups (clinical assessment) versus the usual lifecourse clinical experience. This is a calculation which must involve a consideration of the effectiveness of altering the frequency of assessments for a person in the tails of a disease risk profile versus the cost of regular service. Factors that affect these calculations include not only the cost of assessment but also that of implementing more personalised programmes and of the resultant rates of over- and under-diagnosis of multiple conditions[17]. However the cost of the genome-wide screening itself can be regarded as negligible once it is conducted routinely, given that the benefit is distributed across a spectrum of health outcomes, in contrast to measurement of risk markers such as lipoproteins that are specific to certain classes of disease and must be repeatedly measured over time.

Furthermore, there is the societal or ethical cost to those located in the lower part of the risk profile. This may generate knee-jerk feelings of discrimination or injustice, however in light of our knowledge of common genetic effects of this nature (especially their interplay with the true non-genetic environment) is only as threatening as conventional diagnoses which themselves sit on a continuum of determinism. Indeed, as patients, we seem relaxed with the idea of having our own unique balance of phenotypic abnormality measured by a combination of crude instruments and clinician eye, but we remain strangely disturbed by the notion of having an additional source of information (which may at least be measured more precisely) used in the same examination.

Similarly, concerns that individuals may prefer not to know of their increased genetic risk seem to overlook how most of us live more or less easily with the knowledge of genuinely non-modifiable risk factors, age and sex. These concerns may originate from perceptions of genetic determinism occurring in rare genetic diseases, and the burden may be on genetic epidemiologists to effectively convey the probabilistic nature of genetic risk in complex disease. That said, few parents would endorse the removal of any part of a physician's toolkit when having a child assessed or diagnosed. We should therefore in principle not feel any different about genetics when delivered appropriately. Even if it is perceived as undesirable or unethical to know genetic risk in advance of disease, once disease is established we may wish to rationalise the hand of fate by asking whether there was a genetic origin.

Can we test this approach? What guarantee is there that the formulation of differing check-up regimes will, in time, yield favourable outcomes for those subjected to them versus normal access to clinical services? The obvious step is to undertake a randomized controlled trial in a design comparable to some already ongoing in the UK and elsewhere, for example the ProtecT trial[18] concerning prostate cancer outcomes following randomly allocated screening. Under these conditions it would be practical to examine the impact of differential rates of clinical assessment through the lifecourse in strata assigned by aggregate genotypic scores for disease outcomes of choice. A study of this nature would not only give evidence of the value of this type of approach to realise, at population scale, the utility of the small effects of common genetic variants when combined *en masse*. It would also allow for health economic, ethical and participant experience analyses to be undertaken and further studies to be designed to bring forward genetic data as yet another part of the clinical tool kit.

This comment was first drafted in 2012 and whilst no longer potentially novel[19,20], is worth blowing the dust off. In the last five years there has been an explosive development of genetic association studies, and increasing emphasis on paradigms of precision medicine. Indeed most recent initiatives in the world of genome-wide association studies have shown potential utility in prediction which may be the opening confirmation of this thesis[21–23]. Coupled with this is the observation that with the human genetics community continuing to pursue studies of ever increasing magnitude and power, then the ability to detect the heritable contribution to the most distal and complex of human phenotypes (including behaviour which in itself opens the door for altered environmental exposure) is increasing. Consequently the capacity for genetic studies to predict complex outcomes increases (de facto) and we are left in a position where (i) genetic prediction may not only be a flag of rare syndromes, but a reasonable guide by which one might tailor lifecourse clinical interaction, a form of "precision" medicine, (ii) we seek to identify the truly environmental factors contributing to disease as unlike measuring genetic variation, this does not have an information threshold beyond which further measurement is redundant (i.e. the challenge of tracking true non-shared environment at the individual level) and (iii) we need to challenge the inference from genetic association studies perceived to add clarity through biological/molecular insight - sampling frame, power, genetic architecture and the type of genetic study all contribute to complexities in inference[24]. For common genetic variation and complex traits, the future should not only be causation and dissection of the wealth of new signals, but also the acceptance of complexity and the prediction of low-cost, lifecourse shifts in the way we interact with clinical service providers.

Disclaimer
The views expressed in this article are those of the authors. Publication in Wellcome Open Research does not imply endorsement by Wellcome.

Data availability
No data is associated with this article.

References

1. Visscher PM, Yang J, Goddard ME: **A commentary on 'common SNPs explain a large proportion of the heritability for human height' by Yang *et al*. (2010).** *Twin Res Hum Genet.* 2010; **13**(6): 517–24.
    PubMed Abstract | Publisher Full Text

2. International Schizophrenia Consortium, Purcell SM, Wray NR, *et al*.: **Common polygenic variation contributes to risk of schizophrenia and bipolar disorder.** *Nature.* 2009; **460**(7256): 748–52.
    PubMed Abstract | Publisher Full Text | Free Full Text

3. Dudbridge F: **Polygenic Epidemiology.** *Genet Epidemiol.* 2016; **40**(4): 268–72.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

4. Wray NR, Lee SH, Mehta D, *et al.*: **Research review: Polygenic methods and their application to psychiatric traits.** *J Child Psychol Psychiatry.* 2014; **55**(10): 1068–87.
    PubMed Abstract | **Publisher Full Text**

5. Visscher PM, Brown MA, McCarthy MI, *et al.*: **Five years of GWAS discovery.** *Am J Hum Genet.* 2012; **90**(1): 7–24.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

6. Sovio U, Mook-Kanamori DO, Warrington NM, *et al.*: **Association between common variation at the *FTO* locus and changes in body mass index from infancy to late childhood: the complex nature of genetic association through growth and development.** *PLoS Genet.* 2011; **7**(2): e1001307.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

7. Renström F, Shungin D, Johansson I, *et al.*: **Genetic predisposition to long-term nondiabetic deteriorations in glucose homeostasis: Ten-year follow-up of the GLACIER study.** *Diabetes.* 2011; **60**(1): 345–54.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

8. Nelson MR, Tipney H, Painter JL, *et al.*: **The support of human genetic evidence for approved drug indications.** *Nat Genet.* 2015; **47**(8): 856–60.
    PubMed Abstract | **Publisher Full Text**

9. Hemani G, Zheng J, Elsworth B, *et al.*: **The MR-Base platform supports systematic causal inference across the human phenome.** *eLife.* 2018; **7**: pii: e34408.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

10. Smith GD: **Epidemiology, epigenetics and the 'Gloomy Prospect': embracing randomness in population health research and practice.** *Int J Epidemiol.* 2011; **40**(3): 537–62.
    PubMed Abstract | **Publisher Full Text**

11. Garcia-Closas M, Gunsoy NB, Chatterjee N: **Combined associations of genetic and environmental risk factors: implications for prevention of breast cancer.** *J Natl Cancer Inst.* 2014; **106**(11): pii: dju305.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

12. Szulkin R, Whitington T, Eklund M, *et al.*: **Prediction of individual genetic risk to prostate cancer using a polygenic score.** *Prostate.* 2015; **75**(13): 1467–74.
    PubMed Abstract | **Publisher Full Text**

13. Frampton MJ, Law P, Litchfield K, *et al.*: **Implications of polygenic risk for personalised colorectal cancer screening.** *Ann Oncol.* 2016; **27**(3): 429–34.
    PubMed Abstract | **Publisher Full Text**

14. Staley JR, Blackshaw J, Kamat MA, *et al.*: **PhenoScanner: a database of human genotype-phenotype associations.** *Bioinformatics.* 2016; **32**(20): 3207–3209.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

15. Neale B: **An updated GWAS of the UK Biobank.**
    Reference Source

16. Zheng J, Erzurumluoglu AM, Elsworth BL, *et al.*: **LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis.** *Bioinformatics.* 2017; **33**(2): 272–9.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

17. Pashayan N, Duffy SW, Neal DE, *et al.*: **Implications of polygenic risk-stratified screening for prostate cancer on overdiagnosis.** *Genet Med.* 2015; **17**(10): 789–95.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Martin RM, Donovan JL, Turner EL, *et al.*: **Effect of a Low-Intensity PSA-Based Screening Intervention on Prostate Cancer Mortality: The CAP Randomized Clinical Trial.** *JAMA.* 2018; **319**(9): 883–95.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

19. Lewis CM, Vassos E: **Prospects for using risk scores in polygenic medicine.** *Genome Med.* 2017; **9**(1): 96.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

20. Torkamani A, Wineinger NE, Topol EJ: **The personal and clinical utility of polygenic risk scores.** *Nat Rev Genet.* 2018; **19**(9): 581–90.
    PubMed Abstract | **Publisher Full Text**

21. Khera AV, Chaffin M, Aragam KG, *et al.*: **Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations.** *Nat Genet.* 2018; **50**(9): 1219–24.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

22. Maas P, Barrdahl M, Joshi AD, *et al.*: **Breast Cancer Risk From Modifiable and Nonmodifiable Risk Factors Among White Women in the United States.** *JAMA Oncol.* 2016; **2**(10): 1295–302.
    PubMed Abstract | **Publisher Full Text** | **Free Full Text**

23. Inouye M, Abraham G, Nelson CP, *et al.*: **Genomic risk prediction of coronary artery disease in nearly 500,000 adults: implications for early screening and primary prevention.** *bioRxiv.* 2018.
    Publisher Full Text

24. Haworth S, Mitchell R, Corbin L, *et al.*: **Common genetic variants and health outcomes appear geographically structured in the UK Biobank sample: Old concerns returning and their implications.** *bioRxiv.* 2018.
    Publisher Full Text

Open Peer Review

Current Referee Status: ✓ ? ✓ ✓

Version 1

Referee Report 25 February 2019

✓   **Dorret I. Boomsma** (iD) [1], **Brandon Johnson** [2]

[1] Department of Biological Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands
[2] Avera Institute for Human Genetics, Avera McKennan Hospital & University Center, Sioux Falls, SD, USA

Genomic prediction is likely to be of increasing relevance in clinical applications, prevention and to further scientific discovery. In this review, we would like to focus on three points: what might be expected when a polygenic risk score (PRS) is obtained say at birth for later outcomes in life? To answer this question we turn to twin research, in particular research with monozygotic (MZ) twins. Secondly, we want to add to the discussion on the ideal centralized repository and which members of the population should be included. Thirdly, we mention that some groups will benefit greatly from the developments in PRS research and applications as this may be their only way to learn about their genetic risks.

More specifically to our first point of what might be expected from a PRS to tell individuals about their risk over the course of life? There probably is no better estimate for this than the concordance rates in monozygotic twin pairs: the concordance rates for disease and disorders in MZ twins gives the best limit for how predictive a PRS might be. MZ twins share (nearly) all their DNA sequence, many of the same prenatal exposures, e.g. a smoking mother, and are usually brought up in the same family. Yet, even for highly heritable traits their concordance rarely exceeds 50%. For example, for type 2 diabetes, a recent meta-analysis estimated the heritability at 72% but the concordance in MZ pairs at 2.1% (1). The heritability of schizophrenia was estimated at 79% based on data from two nationwide registries in Denmark, while the MZ twin pair concordance was 33% (2). A more comprehensive summary of MZ twin concordances is presented in reference (3) and reference (4) gives the mathematical reasoning of why this is so: If liability to disease is continuous, and disease status is diagnosed after a threshold has been passed, the probability of discordance in MZ twins depends on the heritability of the underlying liability and on the threshold (4). Especially for rare disorders, which have a high threshold, many affected MZ twins are discordant even if the heritability is high. The reasoning that applies to an MZ twin with an affected cotwin of also becoming affected themselves, is the same reasoning that applies to an individual with a high genetic risk.

For our second comment we would like to introduce into the discussion sample size and power versus representativeness of the proposed centralized repository. To draw accurate conclusions about a population, we do not need to include the entire population, but need to make certain that within the discovery group all individuals that reside in the population have the same chance to be part of the discovery group, e.g. through random sampling. The first issue of power receives ample attention in

GWAS, the second one of representativeness currently much less.

Lastly, will PRS replace family history? There are groups in the population for which PRS have very clear benefits. Think of individuals who were adopted, possibly from abroad, who became orphans at a young age or for other reasons do not know one or both of their biological parents. They cannot obtain their family history of disease but might learn about genetic risks through PRS.

References

1. Willemsen G, Ward KJ, Bell CG, Christensen K, Bowden J, Dalgård C, Harris JR, Kaprio J, Lyle R, Magnusson PK, Mather KA, Ordoňana JR, Perez-Riquelme F, Pedersen NL, Pietiläinen KH, Sachdev PS, Boomsma DI, Spector T: The Concordance and Heritability of Type 2 Diabetes in 34,166 Twin Pairs From International Twin Registers: The Discordant Twin (DISCOTWIN) Consortium.*Twin Res Hum Genet*. 2015; **18** (6): 762-71 PubMed Abstract | Publisher Full Text
2. Hilker R, Helenius D, Fagerlund B, Skytthe A, Christensen K, Werge TM, Nordentoft M, Glenthøj B: Heritability of Schizophrenia and Schizophrenia Spectrum Based on the Nationwide Danish Twin Register.*Biol Psychiatry*. 2018; **83** (6): 492-498 PubMed Abstract | Publisher Full Text
3. van Dongen J, Slagboom PE, Draisma HH, Martin NG, Boomsma DI: The continuing value of twin studies in the omics era.*Nat Rev Genet*. 2012; **13** (9): 640-53 PubMed Abstract | Publisher Full Text
4. Smith C: Heritability of liability and concordance in monozygous twins.*Ann Hum Genet*. 1970; **34** (1): 85-91 PubMed Abstract

Is the rationale for the Open Letter provided in sufficient detail?
Yes

Does the article adequately reference differing views and opinions?
Yes

Are all factual statements correct, and are statements and arguments made adequately supported by citations?
Yes

Is the Open Letter written in accessible language?
Yes

Where applicable, are recommendations and next steps explained clearly for others to follow?
Partly

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Behavior genetics, twin research, quantitative genetics, GWAS

We have read this submission. We believe that we have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Referee Report 19 February 2019

Zoltan Kutalik

Swiss Institute of Bioinformatics, Lausanne, Switzerland

The paper pinpoints the potential benefit of using genetic risk scores in clinical settings. While the outlined thoughts are individually not necessarily original, I like that these arguments are collected in a coherent piece of work. I hope that more health-economic studies will follow to assess direct health benefits of routine collection of genome-wide genetic data and providing up-to-date genetic risk to inform clinicians. Below are some suggestions that could improve the paper.

1. Some relevant work (supporting the author's claims) could be cited:
Iribarren et al.(2016)[1]
Sanyoura et al.(2018)[2]
McCarthy and Mahajan (2018)[3]

2. Some caution would be necessary with application of such risk scores in clinical practice, some papers drawing attention to potential problems with transferring GRS from one population to another should be mentioned: e.g.

Martin et al.(2018)[4].
Kim et al.(2018)[5]
Martin et al.(2017)[6]

3. In addition, it has to be admitted that even current studies of very large sample size tend to overestimate the importance of GRS due to uncorrected population structure, or simply GRS predicting recent migration patterns/geographic origin:

Abdellaoui et al.(2018)[7]
Haworth et al.(2019)[8] (by the author of this work)
Sohail et al.(2018)[9]
Belsky et al.(2018)[10]

4. The paper focused on genotyping, but could you add a few words on the benefit of sequencing?

5. While I fully agree that routine collection of genome-wide genetic data should become the norm, improved health records on family history, health outcomes, environmental exposures, lifestyle, etc. should be equally important. For example, a part of the genetic risk score to convert to T2D is simply predicting BMI (because of the BMI->T2D causal link), hence with available BMI data, the T2D risk score should be updated accordingly.

References
1. Iribarren C, Lu M, Jorgenson E, Martínez M, Lluis-Ganella C, Subirana I, Salas E, Elosua R: Clinical Utility of Multimarker Genetic Risk Scores for Prediction of Incident Coronary Heart Disease: A Cohort Study Among Over 51 Thousand Individuals of European Ancestry.*Circ Cardiovasc Genet*. 2016; **9** (6): 531-540 PubMed Abstract | Publisher Full Text
2. SANYOURA M, LETOURNEAU L, NAYLOR R, PHILIPSON L, WEEDON M, ORAM R, GREELEY S: Clinical Utility of the T1D Genetic Risk Score—Examples from the U.S. Monogenic Diabetes Registry. *Diabetes*. 2018; **67** (Supplement 1): 264-OR Publisher Full Text
3. McCarthy M, Mahajan A: The value of genetic risk scores in precision medicine for diabetes. *Expert*

Review of Precision Medicine and Drug Development. 2018; **3** (5): 279-281 Publisher Full Text

4. Martin A, Kanai M, Kamatani Y, Okada Y, Neale B, Daly M: Current clinical use of polygenic scores will risk exacerbating health disparities. *bioRxiv*. 2018. Publisher Full Text

5. Kim M, Patel K, Teng A, Berens A, Lachance J: Genetic disease risks can be misestimated across global populations. *Genome Biology*. 2018; **19** (1). Publisher Full Text

6. Martin AR, Gignoux CR, Walters RK, Wojcik GL, Neale BM, Gravel S, Daly MJ, Bustamante CD, Kenny EE: Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations.*Am J Hum Genet*. 2017; **100** (4): 635-649 PubMed Abstract | Publisher Full Text

7. Abdellaoui A, Hugh-Jones D, Kemper K, Holtz Y, Nivard M, Veul L, Yengo L, Zietsch B, Frayling T, Wray N, Yang J, Verweij K, Visscher P: Genetic Consequences of Social Stratification in Great Britain. *bioRxiv*. 2018. Publisher Full Text

8. Haworth S, Mitchell R, Corbin L, Wade K, Dudding T, Budu-Aggrey A, Carslake D, Hemani G, Paternoster L, Smith G, Davies N, Lawson D, J. Timpson N: Apparent latent structure within the UK Biobank sample has implications for epidemiological analysis. *Nature Communications*. 2019; **10** (1). Publisher Full Text

9. Sohail M, Maier R, Ganna A, Bloemendal A, Martin A, Turchin M, Chiang C, Hirschhorn J, Daly M, Patterson N, Neale B, Mathieson I, Reich D, Sunyaev S: Signals of polygenic adaptation on height have been overestimated due to uncorrected population structure in genome-wide association studies. *bioRxiv*. 2018. Publisher Full Text

10. Belsky D, Domingue B, Wedow R, Arseneault L, Boardman J, Caspi A, Conley D, Fletcher J, Freese J, Herd P, Moffitt T, Poulton R, Sicinski K, Wertz J, Harris K: Genetic analysis of social-class mobility in five longitudinal studies. *Proceedings of the National Academy of Sciences*. 2018; **115** (31): E7275-E7284 Publisher Full Text

Is the rationale for the Open Letter provided in sufficient detail?
Yes

Does the article adequately reference differing views and opinions?
Partly

Are all factual statements correct, and are statements and arguments made adequately supported by citations?
Yes

Is the Open Letter written in accessible language?
Yes

Where applicable, are recommendations and next steps explained clearly for others to follow?
Yes

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: statistical genetics

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Referee Report 08 February 2019

?　**David J. Porteous** (iD)
Centre for Genomic and Experimental Medicine (CGEM), Institute of GenMolecular Medicine (IGMM), Western General Hospital, University of Edinburgh, Edinburgh, UK

I agree that genomic prediction has enormous potential, but whereas the many GWAS associations and polygenic risks scores now reported are statistically significant, the proportion of variance explained remains modest for most conditions. Moreover, there is still work to be done to answer the general question as to how cohort-based predictions translate to whole populations, or more relevantly (and not really considered in this piece) *individuals* in that population. If nothing else, the letter makes an argument for family history to be given much greater prominence in user-health service settings. We may not yet know which set of genetic variants will provide maximum predictive value, but we can certainly say that simply asking about medical histories and clinically relevant measures (BP etc) of close relatives is valuable.
I appreciate that this is a 'dusting off' of a comment first drafted in 2012, but I think it remains pertinent and I hope that my comments will be taken broadly as supportive of the proposition.

It would be useful however to take more account of what has happened since 2012. Two obvious examples come to mind.

The first is the release and preliminary GWAS analysis of the UK Biobank cohort (Oriol Canela-Xandri, Konrad Rawlik & Albert Tenesa (2018)[1]) (GeneATLAS, http://geneatlas.roslin.ed.ac.uk)
At 500,000 strong UKB is not even 1% of the UK population, but nevertheless it should be large enough to test these predictions at a quasi-population level, for example by asking with what success do any or all polygenic risk scores predict disease incidence post-recruitment?

The second is to acknowledge the parallel acceleration of sign-up by subscription to private sector, GWAS-based companies such as 23andMe (www.23andme.com). The predictive validity of these over the counter tests are controversial. That, we presume, will change. Where will that lead regarding the patient-doctor relationship? Who is the guardian and who the expert of medical knowledge and wisdom.

If GWAS is to become part of the GP tool kit then it would have to do at least three things – pass national and international standards of scrutiny for predictive validity, cross the cost/benefit boundary of clinical utility, and not add to GP consultation time. This is not to challenge the precept, but I think it would provide a useful end point to the discussion.

Other reviewer comments
I endorse the comments made by fellow reviewer Nick Martin and note the author's response.

Minor points:
I have a bit of a problem with the title. Anyone coming to this letter from outside scientific circles (perhaps an ALSPAC participant?) may well be put off by 'The Genetic Sphygmomanometer'. They are more likely to read further if it were 'The genetic blood pressure monitor'.

Some of the text is a bit convoluted.
'the detection of which suffers a natural, genetic architecture versus analytical power limited, aggregate predictive ability.' Is convoluted and opaque. Please rephrase [sometimes two sentences are better than one]

'Done for each individual in a population of common ancestry to discovery studies educating the location of genetic risk,' is convoluted. Please rephrase. My suggestion: 'If done for each individual in a population of common ancestry to define the genetic risk loci,'

'for those subjected to them versus normal access' sounds like a punishment. Perhaps 'for those offered them versus normal access'?

'an explosive development of' 'an explosion of' and add current number with date.

References

1. Canela-Xandri O, Rawlik K, Tenesa A: An atlas of genetic associations in UK Biobank.*Nat Genet*. 2018; **50** (11): 1593-1599 PubMed Abstract | Publisher Full Text

Is the rationale for the Open Letter provided in sufficient detail?
Yes

Does the article adequately reference differing views and opinions?
Yes

Are all factual statements correct, and are statements and arguments made adequately supported by citations?
Yes

Is the Open Letter written in accessible language?
Partly

Where applicable, are recommendations and next steps explained clearly for others to follow?
Not applicable

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Psychiatric genetics Cohort studies

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Referee Report 05 November 2018

https://doi.org/10.21956/wellcomeopenres.16206.r34148

Nicholas G. Martin
Queensland Brain Institute, University of Queensland, Brisbane, Qld, Australia

3rd para, [beg So if we..] last sentence is so convoluted as to be incomprehensible; please rephrase.
6th para [beg Secondly..] Why does it need to be compulsory ["requires all members"] ? Why cant it be voluntary and incremental? Of course it would be best if 100% did it, but there's utility in less than perfect, and much more acceptable to the populace.
7th para [beg A key..]; authors might consider the influence of neurotic/conscientious vs. lackadaisical/thick in uptake and use; these are biases that can be quantified from known PRS.

8th para [beg Furthermore..] Not clear which is lower and upper tail; is lower = higher risk ? If lower = lower risk then these people will feel relief, surely.

9th para [beg Similarly]: "concerns that indivs.." Why should there be concern? It's an individual choice and people can pick and chose what they want to know. For example, they may wish to know their PRS for glaucoma [actionable] but not for Alzheimer's [not].

10th para [beg Can we test] what do they speculate on is already happening - lots at ASHG this year[1]. Last paragraph, last sentence; couldn't agree more!

References

1. Brown SN, Jouni H, Marroush TS, Kullo IJ: Effect of Disclosing Genetic Risk for Coronary Heart Disease on Information Seeking and Sharing: The MI-GENES Study (Myocardial Infarction Genes).*Circ Cardiovasc Genet*. 2017; **10** (4). PubMed Abstract | Publisher Full Text

Is the rationale for the Open Letter provided in sufficient detail?
Yes

Does the article adequately reference differing views and opinions?
Yes

Are all factual statements correct, and are statements and arguments made adequately supported by citations?
Yes

Is the Open Letter written in accessible language?
Yes

Where applicable, are recommendations and next steps explained clearly for others to follow?
Yes

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Author Response 10 Dec 2018
Nic Timpson, University of Bristol, UK

Response to initial review (NM)

3rd para, [beg So if we..] last sentence is so convoluted as to be incomprehensible; please rephrase.

- This has been attended to and will be included in the next posted version.

6th para [beg Secondly..] Why does it need to be compulsory ["requires all members"] ? Why cant it be voluntary and incremental? Of course it would be best if 100% did it, but there's utility in less than perfect, and much more acceptable to the populace.

- This is absolutely right of course. Our phrasing:

"this approach requires all members of the population to have a comprehensive screen of common genetic variants"

… was really to assert that the approach here would be hypothetically a population based effort which would provide best utility for understanding the complex contributions to complex health outcomes in all. Indeed this is far from deterministic and one would have to operate efficient and population scale changes in health care interaction to see on average benefits. We agree with the point that there is an ideal scenario that would have everyone participating, but in reality an optional scheme is more acceptable and would still be useful.

7th para [beg A key..]; authors might consider the influence of neurotic/conscientious vs. lackadaisical/thick in uptake and use; these are biases that can be quantified from known PRS.

- This is a sensible and real detail which would have to be accounted for in any real-world development of this type of approach to incorporating genetic information into health care practice in the proposed way. Key to this, however, is the step-change notion that genetics (even of complex traits) is not to be treated in a manner different to other forms of partial diagnosis and medical examination. These features also suffer the same issues, though are not abandoned as a result.

8th para [beg Furthermore..] Not clear which is lower and upper tail; is lower = higher risk ? If lower = lower risk then these people will feel relief, surely.

- Either way there is of course the interesting position that those with low predicted risk will (a) feel potential relief, but (b) will have to understand their relative risk position and that this is a predicted average. For clarification, being "located in the lower part of the risk profile" does mean at lower predicted risk and could potentially elicit feelings of discrimination in that other people are getting more attention when they themselves remain at risk. This is a central tension in that the probabilistic assessment of risk will not satisfy individual feelings of need for health care – nor should it preclude access. There is a need to discuss the potential impact of even small changes of behaviour with relation to interaction with healthcare in light of this – though it seems less likely that those at (on average) lower risk would chose to alter their use of healthcare provision.

9th para [beg Similarly]: "concerns that indivs.." Why should there be concern? It's an individual choice and people can pick and choose what they want to know. For example, they may wish to know their PRS for glaucoma [actionable] but not for Alzheimer's [not].

- This paragraph was not as clear as it could have been and will be updated in the next posted version. The concerns mentioned refer to both those of the clinician and the patient/health member of the public and you are right – there should be no concern that some chose not to know of risk.Individuals clearly should have the choice whether to be informed of their genetic risk, but that choice should itself be informed not only by the potential impact to health individually, but also the potential benefit to the population more broadly.

10th para [beg Can we test] what do they speculate on is already happening - lots at ASHG this year[1].
Last paragraph, last sentence; couldn't agree more!

- It is not clear that there are explicit trials of this type of activity currently, but this could be undertaken. Further to this – we are delighted that the reviewer agrees with our last point.

** of note – all manuscript changes will be made/updated on the receipt of 3 reviewer contributions.

Competing Interests: No competing interests were disclosed.