Franks, P. W., & Timpson, N. (2018). Genotype-based recall studies in complex cardiometabolic traits. *Circulation: Genomic and Precision Medicine, 11*(8), [e001947]. https://doi.org/10.1161/CIRCGEN.118.001947

Peer reviewed version

Link to published version (if available): 10.1161/CIRCGEN.118.001947

Link to publication record in Explore Bristol Research

PDF-document

This is the final published version of the article (version of record). It first appeared online via AHA at https://www.ahajournals.org/doi/10.1161/CIRCGEN.118.001947. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

**General rights**

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/user-guides/explore-bristol-research/ebr-terms/
Title: Genotype-based recall studies in complex cardiometabolic traits

Running title: Genotype-based recall

Paul W Franks PhD MPhil MS BS(Hons)1,2,3,4, Nicholas J Timpson PhD MS BS5

1. Department of Clinical Sciences, Genetic and Molecular Epidemiology Unit, Lund University Diabetes Centre, Skåne University Hospital, SE-21741, Malmö, Sweden.
2. Oxford Centre for Diabetes, Endocrinology and Metabolism, University of Oxford, Oxford,
3. Department of Nutrition, Harvard TH Chan School of Public Health, Boston, MA, USA
4. Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden
5. MRC Integrative Epidemiology Unit, Avon Longitudinal Study of Parents and Children, Population Health Science, Bristol Medical School, Oakfield House, Oakfield Grove, Bristol, BS8 2BN, UK

Correspondence:

Prof. Paul W. Franks
Department of Clinical Sciences
Genetic and Molecular Epidemiology Unit
CRC, Building 91, Level 10
Jan Waldenströms gata 35
Skåne University Hospital Malmö
SE-205 02, Malmö, Sweden
Tel: +46703045755
Paul.Franks@med.lu.se

Prof. Nicholas J Timpson
MRC Integrative Epidemiology Unit
Avon Longitudinal Study of Parents and Children
Population Health Science, Bristol Medical School
Oakfield House, Oakfield Grove, Bristol, BS8 2BN, UK
N.J.Timpson@bristol.ac.uk

Word count: 194 (abstract); 5744 (main body)
Figs: 2
Refs: 49
Abstract

In genotype-based recall (GBR) studies, people (or their biological samples) who carry genotypes of special interest for a given hypothesis test are recalled from a larger cohort (or biobank) for more detailed investigations. There are several GBR study designs that offer a range of powerful options to elucidate i) genotype-phenotype associations (by increasing the efficiency of genetic association studies, thereby allowing bespoke phenotyping in relatively small cohorts), ii) the effects of environmental exposures (within the Mendelian randomization framework), and iii) gene-treatment interactions (within the setting of GBR interventional trials). In this review we overview the literature on GBR studies, as applied to cardiometabolic health outcomes. We also review the GBR approaches used to date and outline new methods and study designs that might enhance the utility of GBR-focused studies. Specifically, we highlight how GBR methods have the potential to augment randomized controlled trials, providing an alternative application for the now increasingly accepted MR methods usually applied to large-scale population-based datasets. Further to this, we consider how functional and basic science approaches alongside GBR designs offer intellectually intriguing and potentially powerful ways to explore the implications of alterations to specific (and potentially druggable) biological pathways.
Introduction

In the context of cardiometabolic disease, as with all complex disease, genetic variation is neither the sole nor absolute determinant of morbidity. Nevertheless, DNA sequence variation plays a profound role in determining phenotypic diversity, largely because the nuclear genome is the first step in the process of protein (amino acid) synthesis, thus setting the stage for all subsequent biological processes. A person’s genetic characteristics also affect his or her sensitivity to environmental exposures (diet, exercise, smoking and alcohol, for example), such that predisposition to disease is conditional on the combination of genetic and non-genetic risk factors. Thus, genomic variants that affect disease predisposition do so by reducing the potential health-enhancing features of environmental exposures (e.g., by interfering with the extraction or synthesis of micronutrients from food), or by raising susceptibility to risk exposures (e.g., by restricting cellular macronutrient transportation or by inhibiting autophagy). Pharmacotherapies are in essence environmental exposures, albeit man-made, and their effects are also prone to modification by genetic variations, such that some variants may raise or lower sensitivity to drugs and others may increase the risk of side-effects.

Variation in a person’s DNA sequence is essentially randomized to conventional confounding factors as a result of the mechanics of intergenerational inheritance, and remains the same throughout the life-course, features that distinguish inherited DNA variation from all other known forms of biological information. Moreover, because of the static nature of germline DNA, genotype-phenotype associations are unidirectional and therefore less prone to reverse causality. These features provide a unique opportunity to assess causal relationships between risk factors of interest (including modifiable environmental exposures) with which DNA variants are associated (e.g., macronutrient preference, tobacco use, alcohol consumption) and disease outcomes. In analyses such as this, carriage of gene variants that affect the extent to which a person is exposed to environmental factors can provide an easily measured proxy for the factor itself. The approach, termed ‘Mendelian randomization’ (MR) benefits from the independence of the instrumental variable (the gene variant(s)) from most factors that might otherwise confound the relationship between the exposure and outcome.

Genetics can also provide powerful insights into how specific perturbations in molecular pathways, which might be mimicked by artificial or natural chemical compounds, affect the development of disease. Reciprocally, the molecular perturbations caused by gene variants may parallel the effects of specific drugs (as ‘phenocopies’), providing insights into the long-term effects that treatment with the drug is likely to convey. In this regard, genetics has already helped accelerate the pace of drug development, halving the rate of failure and substantially reducing costs. Genetics has also proven extremely valuable in drug repurposing by i) guiding the use of drugs for hitherto unknown conditions and ii) improving safety by predicting adverse drug reactions. While far less well understood, genetics has the potential to play a parallel role in the development of nutraceuticals.
Studying participants with and without a specific phenotype (case-control studies), particularly for a disease whose clinical presentation is striking, such as congenital obesity, familial dyslipidemia, or permanent neonatal diabetes, has long been used to maximize power to detect rare genetic causes of disease. However, specifically enrolling participants with contrasting genetic characteristics into a study as a feature of its design (genotype-based recall: GBR) in order to elucidate the mechanisms linking genotypes to disease, is less common. The GBR approach is particularly powerful when the ratio of the two genotype groups of interest is heavily imbalanced in the general population (thus allowing substantial ‘rebalancing’ by selectively over-recruiting carriers of the less frequent genotype) and new assays and/or interventions will be implemented, which would be too expensive or burdensome to deploy in large cohorts.

Here, we define GBR studies as those within which individuals are purposefully recruited (or biological specimens for these individuals are sampled) from a larger background cohort (or biobank) on the basis of the individual’s genotype. Typically, two groups are sampled with contrasting genetic characteristics (e.g., the two homozygous genotypes at a given SNP or from the upper and lower tails of a genetic risk score), thereby contrasting high and low degrees of genetic burden. GBR can be applied to a number of conventional study designs, including cross-sectional studies, clinical trials and functional experiments. GBR studies are attractive as they afford improvements in statistical power and the possibility to deploy more precise phenotyping methods than in conventional trials or sampling strategies. Applying GBR approaches to conventional trial designs can also strengthen causal inference.

Most common genetic variants explain a very small proportion of the overall variance in a given complex trait such as obesity, CHD or diabetes, and most variants that convey larger effects in these traits are rare. These characteristics have led researchers to undertake enormous studies in order to acquire adequate statistical power to detect genetic effects or, in the case of rare variants, focus on family pedigrees. Undertaking detailed phenotyping or complex interventions in large populations is usually untenable and focusing on family pedigrees limits the generalizability of the findings owing to the many other characteristics (genetic, environmental and phenotypic) shared by family members. In these settings, GBR studies in which participants are recruited from population-based cohorts provide an attractive alternative.

Overview of published GBR studies

The vast majority of published GBR studies have focused on elucidating how genetic variants manifest phenotypically. Other forms of GBR have looked to assess the causal impact of specific risk factors and also to enhance trial design through informed participant selection. More than a dozen GBR studies focusing on cardiometabolic traits have been reported in the past two decades with contrasting genetic characteristics, with the PPARG Pro12Ala variant (rs1801282) being the single most extensively studied locus.
PPARG is a nuclear receptor that binds long-chain polyunsaturated fatty acids (PUFA), among other ligands\(^{21}\). Epidemiological\(^{22}\) and animal\(^{23}\) studies suggest that Ala12 allele carriers may benefit more from the insulin-sensitizing effects of dietary PUFA. The first published example of a GBR study involved the comparison of Pro12Pro and Ala12X (n=10 vs. 10) genotype carriers\(^{14}\). The study examined whether the Pro12Ala genotype was associated with insulin sensitivity and insulin secretion basally and after 5-hrs of Intralipid (a 10% intravenous fat emulsion) plus heparin infusion. The authors found nominal evidence that insulin secretion in response to arginine increased in Pro12Pro genotype carriers, and decreased in Ala12X carriers, but few other comparisons yielded detectable differences. A more detailed investigation of Pro12Pro and Ala12X (n= 12 vs. 12) male adults from the UK\(^{15}\) found that, following a mixed-meal containing \([1,1,1-(13)C]\)tripalmitin, Ala12 carriers had lower plasma NEFA concentrations, higher adipose tissue and muscle blood flow, and greater insulin-mediated postprandial hormone-sensitive lipase suppression along with greater insulin sensitivity, than Pro12 homozygotes. A recent observational GBR study of 12 Pro12Pro, 15 Pro12Ala, and 13 Ala12Ala variant carriers, all older Swedish men, found no statistical differences in anthropometric or serological traits by genotype.

To assess whether the Pro12Ala genotype interacts with dietary fats over prolonged durations, Rubin et al undertook a randomized, placebo- controlled, double-blind, crossover trial in 16 male Pro12Pro and 23 male BMI-matched Ala12Ala carriers\(^{13}\). Four dietary fatty acid interventions lasting 28-days each were delivered: i) c-9, t-11CLA or t-10; ii) c-12 CLA; iii) a commercial available 1:1 mix of both isomers (Tonalin®); and iv) placebo (linoleic acid from safflower oil). Four enrolled participants (all carriers of the Ala12Ala genotype) were exclude from the data analysis and comparisons did not account for changes observed in the placebo group, which may have biased the results; moreover, the approach employed to control for the large number of statistical comparisons conducted may not have adequately controlled alpha-inflation. Nevertheless, the study's data suggest that serum insulin concentrations fell more across all fatty acid interventions in Ala12Ala carriers, whereas plasma leptin concentrations fell more in Pro12Pro carriers. No differences in the many other traits were reported. In a second dietary fat intervention trial, Pihlajamäaki and colleagues\(^{12}\) recalled 17 Pro12Pro and 14 Ala12Ala male genotype carriers from the METSIM cohort, and randomized them to isocaloric PUFA- and saturated fatty acid-enriched diets in an 8-week cross-over trial. The investigators failed in their goal of matching the two genotype groups by age and BMI, as the Ala12Ala participants were considerably older (mean 61.2 yrs vs. 56.7 yrs) and heavier (mean 83.5 kg vs. 82.1 kg) than the Pro12Pro genotype group, and analyses were not performed on the basis of intention-to-treat, factors that may bias or confound the reported results. Nevertheless, the authors found that the baseline metabolic profiles of the Ala12Ala carriers were generally more favorable, despite their older age and heavier weight, and that following dietary intervention, the Pro12Pro carriers’ triglyceride concentrations improved more with the PUFA diet than did the Ala12Ala carriers’ triglycerides.

The majority of GBR studies have been descriptive, focused on elucidating the mechanisms underlying an established association between a gene variant and trait through the use of bespoke assessments. However, some studies have employed GBR trials as adjuncts to larger experiments, sometimes with
the objective of providing proof of mechanistic concept using drugs or other chemical compounds to perturb the pathways encoded by specific genes. For example, Tang et al undertook a, double-blind, variable dose (0 (placebo), 10, or 20 mg), GBR randomized trial of oral yohimbine (a compound extracted from pausinystalia johimbe tree bark). Treatment was administered once every 2-weeks. The authors had previously determined that the rs553668 variant at ADRA2A (the gene encoding the α2AAR receptor) raises type 2 diabetes risk because of impaired insulin secretion owing to α2AAR receptor overexpression24. Participants were recalled by rs553668 genotype (21 were non-risk homozygotes, 21 were heterozygous, and 7 were homozygous risk allele carriers). The primary outcome was 30-min plasma insulin concentrations during a 75g OGTT (comparing 20mg oral yohimbine vs. placebo). The study showed that ADRA2A encoded defect in insulin secretion can be overcome by blocking the expression of the α2AAR receptor. GBR trials are most powerful when risk and non-risk genotypes are balanced, thereby elevating the minor allele frequency (MAF) from its naturally occurring level (about 0.15 to 0.35 in the case of rs553668 [according to 1000 Genomes and NCBI]), to 0.50. Like all other published GBR trials, no clear sample-size calculations were reported in Tang et al’s paper, and the frequency of major and minor allele homozygotes was not balanced (MAF=0.36), in part because heterozygotes were enrolled. Thus, the study was not optimally powered given the sample-size. Nevertheless, the study provides a rare example of a randomized controlled trial (RCT) using GBR that showed that a specific gene defect can be successfully treated with a chemical compound.

Elsewhere, Tuomi et al undertook a GBR intervention study in 23 GG (risk) and 22 CC (non-risk) homozygous genotype carriers at the MTNR1B rs10830963 variant; the authors showed that the inhibitory effect of melatonin treatment on endogenous insulin secretion was greatest in the GG genotype group17.

A single GBR study of monogenic obesity has been reported in the literature. Here, van der Klaauw et al enrolled lean controls (n=20), obese controls (n=20) and heterozygous loss-of-function MC4R mutation carriers (n=14) whose ad libitum food intake was assessed in response to a high fat meal 19, 25; MC4R deficient participants consumed 95% more of the meal than lean participants, and 65% more than obese participants. In a second study reported in the same paper, 20 obese controls, 20 lean controls and 10 MC4R deficient volunteers underwent a sucrose preference test; the MC4R deficient volunteers liked and consumed substantially less of the high sucrose meal than lean and obese controls.

Thus far, almost all published GBR studies of cardiometabolic traits have focused on carriers vs. non-carriers of specific genotypes within a particularly penetrant or otherwise well-characterized gene. However, in a recent GBR study26, we recalled participants using a GRS comprised of multiple BMI-associated gene variants derived from a large published GWAS meta-analysis27. The study sought to assess the causal role of body mass index (BMI) in clinically relevant cardiovascular phenotypes within young adults of the Avon Longitudinal Study of Parents and Children (ALSPAC). Participants (N=4,602) with high or low genetic burden for obesity (defined as the lower or upper ~30% of a genome-wide GRS distribution) were invited to engage in the study, and 419 non-pregnant and diabetes-free participants
(average age 21 years) were successfully enrolled. The study results indicated that higher BMI causes higher blood pressure (systolic and diastolic blood pressure, pulse pressure and mean arterial pressure) and higher left ventricular mass index, the latter suggesting an adverse effect on cardiac structure, even in young adults. The GBR analyses also suggested that heavier BMI increased cardiac output, which appeared to be solely driven by increased stroke volume, as neither MR nor GBR analyses indicated a causal effect of BMI on heart rate. This study illustrates the potential for improving phenotypic resolution using a GBR study design, whilst preserving analytical power. The study also provides strong causal evidence that preventing obesity from a young age may help improve long-term cardiovascular health.

**Design considerations for GBR trials**

Existing Phase III RCTs, genotyped after the trials have concluded, have been used extensively to assess gene-treatment interactions, albeit with minimal success. Conventionally, participants in clinical trials are not recruited on the basis of specific genetic characteristics, and thus the genotype frequencies within large trials broadly conform to those observed in the population at large (although, it should be noted that this may not hold true in clinical trial populations enriched for highly heritable traits). Thus, given the relatively small size of most trials (and the corresponding paucity of rare variant carriers) and the small magnitude of most gene-treatment interaction effects, existing trials are likely underpowered to detect gene-treatment interactions in cardiometabolic disease.

There are few current solutions to this problem, as the default approach to overcoming issues of statistical power in interaction analyses using epidemiological data (i.e., ramping-up sample size by meta-analyzing the interaction test statistics derived from multiple cohorts) is unlikely to work in the clinical trial setting; this is because the study populations, designs, and outcomes in clinical trials are often very diverse, factors that can have a substantial detrimental impact on statistical power when the focus is on detecting gene-treatment interactions, and clinical trials are usually much smaller than epidemiological cohorts. This point is illustrated in a meta-analysis of eight weight-loss trials (n=9563) that sought to test FTO gene-lifestyle interactions in weight change: despite compelling evidence from epidemiological studies that FTO variants interact with lifestyle in obesity predisposition, the clinical trial result was null. Whilst these findings may appear contradictory, it is likely that the clinical trial analysis, despite being the largest of its kind to date, simply lacked power to detect the interaction. By contrast, even fairly small-scale appropriately designed GBR trials may be adequately powered to detect gene-treatment interaction effects.

To place the statistical advantages of GBR trials in context with conventional trial analyses, we examined power and sample-size considerations across diverse scenarios, including comparisons with published examples of gene-treatment interactions in conventional clinical trials. We selected two published examples taken from the Diabetes Prevention Program, a randomized controlled trial of intensive lifestyle intervention, metformin treatment and placebo control. The first example was for the metformin-enhancing effects of a variant (rs8065082) at SLC47A1 in diabetes prevention, whereas
the second example related to a gene-lifestyle interaction in dyslipidemia\textsuperscript{36}. With power set at 80%, using parameters such as marginal and interaction effect sizes and incidence rates derived from the original publication and assuming recall of only the homozygous genotypes in equal proportions, we found that a conventional RCT designed to detect the interaction of metformin and rs8065082 (reported in \textsuperscript{35}) would require roughly 3,000 participants, whereas, with all else held equal, a GBR trial would require approximately 1,700 participants. We selected this example of gene-drug interaction, as it is biologically plausible and has been reported elsewhere \textsuperscript{37}. However, the relatively high minor allele frequency (0.44) limits the extent to which GBR improves power; the statistical power-gains expected from GBR trials focusing on carriers of rarer genotypes are generally far greater. We provide proof of this in our paper for single nucleotide variants and genetic scores \textsuperscript{3}. The latter focused on a GRS comprised of multiple dyslipidemia-associated variants and its interaction with the DPP intensive lifestyle intervention; the outcome was 1-yr change in lipoprotein sub-fraction concentrations (small LDL particles) \textsuperscript{36}. We showed for this example that the sample-size requirements to detect the previously reported gene-lifestyle interaction using a conventional RCT would be around 1,900 compared with only \textasciitilde400 in a GBR trial, where all else is held equal.

The two key advantages afforded by GBR trials compared with conventional trials are the smaller sample-sizes and/or shorter durations of follow-up required to achieve comparable event rates and thus also adequate statistical power, features that maximize cost-effectiveness for targeted hypothesis tests \textsuperscript{38}. In some instances, the costs saved using GBR could be reinvested to improve the precision and accuracy of the measurements undertaken in the trail. In settings where very detailed measurements are essential to test a given hypotheses, the GBR approach is likely to be highly appealing. It is also worth considering that improving the precision of assessments is also likely to improve statistical power when modeling gene-treatment interactions in clinical trials \textsuperscript{39}. In our paper\textsuperscript{3}, we provide an example illustrating the extent to which reductions in measurement error influence sample size requirements to detect gene-treatment interactions (assuming 80% power) in trials using conventional recruitment strategies compared with GBR: building on the interaction scenarios for quantitative outcomes discussed above, we showed that where the outcome is assessed with low, moderate or high measurement error, the required sample sizes for trials using random population sampling (conventional trials) would be \textasciitilde200, \textasciitilde600 and \textasciitilde1,800 respectively, compared with \textasciitilde20, \textasciitilde70 and \textasciitilde300 participants respectively for a GBR trial, with all else held equal.

The GBR design is strengthened by the biological insights derived when integrating well-characterized genotype(s) into trials. In this context, this is not a recall framework for understanding a genotype effect, nor is it an MR-style recall study looking to mimic an exposure (both of which occur in observational studies \textsuperscript{2}). Moreover, featuring genotypes into the trial design is not the main route to causality, as this comes from the randomized nature of a randomized controlled trial. The value added is the delivery of a specific type of biological perturbation, and perhaps also within a characterized pathway, that has the potential to augment intervention efficacy or to elucidate mechanisms of action. This does assume knowledge of genotype function, but where this exists the genotype can act as a targeted modification

\textsuperscript{8}
of the trial main effect. In essence, designs of this nature serve to optimize the interrogation of interactions within trials where genotypic data can help improve the efficacy of the main intervention.

In some instances, the index variant(s) used in GBR studies may influence behavioral traits that interfere with recruitment and/or adherence to interventions. If these effects were known prior to undertaking the study, such variants should not be used for recall. Nevertheless, the absence of evidence is not evidence of absence. Thus, given that differential rates of adherence by genotype would likely cause bias, monitoring adherence during the trial and testing whether this differed by genotype would be prudent.

GBR studies might be prone to confounding if the index genotype conveys pleiotropic effects. In the single variant scenario, numerous methods for the detection of pleiotropy exist and using such methods to determine which variants act pleiotropically might preclude their use in GBR studies. When recall strata are composed of many loci, it is almost inevitable that pleiotropic effects will be present and excluding these loci may not be pragmatic. Indeed, providing the combination of pleiotropic effects are balanced across the trial’s arms, which is likely when recalling participants based on polygenic scores, confounding is unlikely to present a major problem.

The heterogeneous nature of polygenic GRSs makes it difficult or impossible to deconvolute the biological basis of the genetic association. By contrast, GRSs comprised of multiple variants that predispose to perturbations within the same molecular pathway may be useful in studies that seek to elucidate biological mechanisms. Although interpreting the biological meaning of polygenic GRSs can be challenging, they remain useful when the primary objective is to maximize the variance explained in a given trait (e.g., where prediction is the focus) or where a robust proxy for an environmental exposure is sought (e.g., in MR experiments). The biological heterogeneity of polygenic GRSs can also help reduce genetic confounding attributable to unbalanced pleiotropy, as highlighted above.

Population stratification is an additional factor that might cause confounding in GBR studies. Population stratification occurs when allele frequencies coincidentally correspond with an outcome of interest. Typically, this occurs in admixed populations that include people with different ancestral linages (and thus genetic characteristics) who also differ in their predisposition to disease, possibly owing to social disparities or other risk factors. In genetic association studies, this type of “hidden structure” is usually resolved by adjusting for genomic markers of these subgroups. In a GBR study, population stratification could also cause confounding. Suppose for example that the index genotype or GRS is in fact an excellent marker of ethnic diversity; recalling people using this genetic marker would then result in an uneven distribution of ethnicity between the recalled groups. If ethnicity impacts the effects of the intervention, one might than falsely conclude that the genotype interacts with treatment, whereas the true causal interaction would be between treatment and ethnicity. To overcome this problem requires that the index genotype (or GRS) is not a marker of hidden structure, which should be checked when designing the study by testing the correlation of the GBR marker with explicit markers of structure (e.g., genomic principle components).
It may be appropriate to feed-back genetic information to study participants, but this will depend the objective of the study and the information provided to participants during the consent process. However, it is very important that if genetic information is given to the participants, this is done after the study has concluded, to avoid introducing bias. Indeed, the genetic characteristics should remain masked from the participants, field staff and data analysts at least until the results of the study are finalized. We provide a range of general recommendations relating to these issues elsewhere.\textsuperscript{2}

Notwithstanding the appeal of GBR trials when studying gene-treatment interactions, selectively recruiting genetically at-risk participants can also enhance power to assess treatment efficacy by virtue of higher rates of the trial’s outcomes and/or greater treatment sensitivity in people with specific gene variants\textsuperscript{42-44}. In contrast to the examples outlined above, where interactions are the central feature of the experiment, GBR trials that enroll participants purely to enhance event rates or treatment effects do so with the expectation that the index genotypes will not interact with the treatment; indeed, a failure of this expectation may scupper the trial.

**GBR and Mendelian randomization (MR)**

MR studies use genetic variants to provide proximal measures of biological perturbations or environmental exposures in order to determine with high-certainty whether these factors are likely to causally affect one or more outcomes of interest. MR relies on specific properties of DNA variants (random assignment at meiosis and saliency across the life-course) to permit a degree of causal inference that can rarely be achieved using conventional epidemiological methods. Moreover, MR studies performed in cross-sectional settings can provide evidence of longitudinal effects (because the same genotypes are carried throughout life and are thus proxies for long-term environmental exposures) and are not susceptible to reverse causality (because the development of disease does not affect a person’s genotype), which are two major limitations of conventional cross-sectional studies.

As we recently described the integration of MR into GBR studies\textsuperscript{2}. By exploiting the key properties of genetic variants that arise from the random allocation of alleles at conception, GBR studies enhance the ability to draw causal inferences in population-based studies and minimize (or completely obviate) confounding, reverse causality and bias. Hence, GBR studies are similar to classical case-control studies, but without the key caveats of the latter. Case-control studies involve the comparison of people with or without a given phenotype (usually a disease diagnosis) and are prone to various sources of confounding and bias, particularly when the accuracy with which a participant recalls an exposure is affected by diagnostic labelling or treatment. For example, a CVD diagnosis is likely to affect responses to questions about lifestyle habits, because many people recognize that CVD is caused by unhealthful lifestyles and acknowledging that one’s poor lifestyle choices likely caused a major health event may be a source of embarrassment that affects how questions about lifestyle habits are answered. In contrast, because genotypes are unaffected by diagnostic labelling, patient education or other exposures, and
can be measured with high accuracy, GBR studies are generally immune to the key limitations of case-control studies.

In GBR studies, it is possible to define the genetic instrument as a single variant (or set of variants) within a specific biological pathway (GBRsv) or conversely the instrument might be defined using multiple variants from several loci that each affect the same phenotype, but through different pathways (GBRmv) (see Fig. 1). The former of these is particularly useful when seeking to understand how perturbations within a given biological pathway impact disease; recalling major and minor allele homozygotes (excluding heterozygotes) is likely to be the most efficient study design strategy when the study is focused on a single variant. By contrast, in GBRmv designs, recalling participants who carry few or many effect alleles across the set of loci (excluding those with an intermediate number of effect alleles) will provide a comprehensive instrument for the index environmental exposure that is often more powerful than a single variant instrument. However, because GBRmv instruments generally characterise heterogeneous biological processes, and the specific set of alleles each participant carries will vary from one participant to the next, these instruments do not characterize specific biological perturbations and should not typically be used for this purpose. In GBRsv studies, particularly where the focus is on rare variants, ensuring these are functional loci, rather than tag-SNPs, is important, as causal variant misclassification (i.e., recalling participants who carry a tag-SNP, but not the causal variant) can undermine statistical power and inference. Nevertheless, genetic variants that have uncharacterized or predicted effects (i.e., loss-of-function variants, cis-regulatory variants or intronic variants that alter DNA-protein binding at potential drug targets) may also be of interest. GBRmv instruments are likely to be more robust to causal variant misclassification, owing to their polygenic nature and that the focus of GBRmv studies is on environmental exposures rather than specific biological perturbations.

To aid the design and initiation of new GBR experiments, we previously developed two complimentary open-access GBR power calculators. The first focuses mainly on the design of GBR interventional trials with single variant and multi-variant approaches and quantitative and time-to-event models (https://gbr-power.crc.med.lu.se), whereas the second calculator emphasizes the design of GBR experiments focused on incorporating MR for causal inference (http://www.bristol.ac.uk/integrativeepidemiology/facilitiesresources/software/ (under ‘RbG Study Planner’)).

**Functional GBR studies**

The process of purposefully restructuring the genetic characteristics of a study population using GBR is in some senses analogous to certain genome engineering methods, such as CRISPR-Cas9. In the case of the latter, one or more alleles within the nuclear DNA’s sequence can be altered; specifically, a double-stranded break is made at a selected site in the genome. Subsequently, the cell’s natural DNA repair mechanisms are leveraged to introduce a specially designed DNA repair template, which, at its
ends, matches the “broken” section of DNA, but also includes an engineered piece of DNA\textsuperscript{45}. By consequence, a population of cells can be created within which a specified ratio of index genotypes exists, such that half of the cell population might carry the minor allele homozygous genotype at a given locus, with the other half carrying the common homozygous genotype. More complex genotype configurations can also be achieved by introducing multi-locus gene edits. Although GBR studies do not seek to engineer DNA edits into an individual’s genome, the genetic characteristics of the study population are engineered to have specific characteristics through the process of recall. These parallel characteristics of GBR studies and CRISPR-Cas9 experiments present a powerful opportunity to study the effects of genes \textit{in vivo} using GBR, and then to validate these effects \textit{in vitro} using CRISPR-Cas9 (see Fig. 2).

In a study focused on elucidating the mechanisms underlying the obesogenic effects of the \textit{FTO} gene, Claussnitzer and colleagues recruited 100 healthy (BMI 20-24 kg/m\textsuperscript{2}) young adults of European ancestry\textsuperscript{7}. Of these, 52 participants carried both copies of the risk allele at the rs9930506 genotype (previously associated with obesity in GWAS meta-analyses) as well as at two other variants tagged by rs9930506 (rs1421085 and rs1558902); collectively, these variants form an obesity risk haplotype comprised of 89 variants in introns 1 and 2 of \textit{FTO}. The remaining 48 participants were considered low risk on the basis that they did not carry any of the risk alleles at the three selected \textit{FTO} variants (rs9930506, rs1421085 and rs1558902). From the subcutaneous adipose tissue of these participants, primary human adipose–derived progenitor cell cultures were obtained for bidirectional CRISPR–Cas9 editing, which was used to demonstrate cell-autonomous causal effects of the target variant. Initially though, the authors showed that \textit{FTO} influences thermogenesis and energy metabolism through the long-range activation of \textit{IRX3} and \textit{IRX5}; they also showed that the expression of mitochondrial, browning, and respiration genes was lower and expression of lipid-storage genes was higher in risk allele carriers compared with non-risk allele carriers. \textit{FTO} risk allele carriers also had larger adipocytes, lower mitochondrial DNA content and basal oxygen consumption, and impaired UCP1 response to cold and \textbeta-adrenergic stimulus. The authors then used CRISPR-Cas9 to reverse the \textit{FTO} genotypes (from high-risk to low-risk, and vice versa) in the preadipocytes; knocking-down \textit{IRX3} and \textit{IRX5} by reversing the high-risk \textit{FTO} genotypes restored oxygen consumption and thermogenesis response to non-risk levels. By contrast, \textit{IRX3} and \textit{IRX5} over-expression by reverse-engineering \textit{FTO} in preadipocytes from non-risk-allele carriers reduced basal oxygen uptake and thermogenesis (including UCP1 expression) to the levels observed in risk-allele carriers.

The study by Claussnitzer \textit{et al}\textsuperscript{7} was the first example of how GBR and \textit{in vitro} genetic engineering might be combined to provide an extremely eloquent proof-of-concept. In the future, this approach might be extended by coupling GBR intervention trials with interventional experiments in CRISPR-Cas9 edited cells that mimic the conditions of the GBR trials.

\textbf{Summary & conclusion}
GBR studies offer a range of powerful options to elucidate the causal effects of i) genotypes (by increasing the efficiency of genetic association studies, thereby allowing bespoke phenotyping in relatively small cohorts), ii) environmental exposures (within the MR framework), and gene-treatment interactions (within the setting of GBR interventional trials). As our narrative review shows, the literature on GBR studies as applied to cardiometabolic health outcomes is expanding but remains small. Most existing GBR studies are descriptive in nature\textsuperscript{4, 8-11, 13-15, 20}, having used recall to minimize the sample size required to deploy bespoke phenotyping methods; a handful of additional GBR studies have examined gene-drug or gene-diet interactions\textsuperscript{12, 17-19, 24, 25}; and only one GBR study has incorporated MR to enhance causal inference\textsuperscript{26}, whilst one further study combined GBR and gene-editing technologies to study mechanism of action\textsuperscript{7}.

Throughout this paper, we have attempted to illustrate the strengths of GBR studies and how this general approach can leverage maximum value out of existing and new datasets. However, GBR experiments are not without their caveats. Perhaps the most obvious disadvantage of GBR studies is that, by design, the distribution of the index genotypes (and correlated phenotypes) within the enrolled cohort will be distinct from the background population. Thus, whilst conventional observational studies and clinical trials may be suitable for a wide-spectrum of secondary hypothesis tests, the very specific design of GBR studies means that the data may not be appropriate for secondary analyses, a limitation that is not considered in previous cost-effectiveness analyses of GBR trials\textsuperscript{38}. Moreover, GBR studies that focus on index variants that are neither functional nor in complete LD with the functional variant will be less powerful than those where the opposite is true, as these synthetic relationships will lead to misclassification of participants across the GBR groups.

It is also likely that in some settings GBR studies may face ethical barriers. For example, many older cohorts did not obtain explicated informed consent to recall participants by genotype. For such cohorts to serve as sampling-frames for GBR studies thus requires that participants are either re-consented (a costly and time-consuming task for very large cohorts, which even when done may see many participants excluded who might not truly object to the use of their samples/data) or, more controversially, legislation is changed to accommodate the principle of “presumed” consent, where the exclusion of a person’s data and/or samples from research requires that they actively opt-out \textsuperscript{46}. Presumed consent places the responsibility firmly in the hands of the donor for electing NOT to allow their samples and/or data to be included in biomedical research, which can be viewed as empowering individuals and facilitating society’s desire to promote biomedical research, or alternatively as an encroachment on human rights. The Icelandic genomics company deCode used presumed consent when establishing their genomics database of all of Iceland’s citizens, successfully arguing that public support for their research was strong and data would be de-identified and securely stored – hence explicit informed consent was not necessary \textsuperscript{47}. Recognizing the challenges of using extant cohorts for GBR sampling-frames, numerous large \textit{de novo} bioresources are being accrued, where participant consent for GBR sub-studies is obtained at enrolment; these include the UK Biobank (N\textapprox500,000)\textsuperscript{48},

\@13
INTERVAL (N~50,000)\(^4^9\), and the NIHR BioResource (N~20,000–100,000) (https://bioresource.nihr.ac.uk/).

Coupling GBR studies with cellular genome engineering experiments is particularly exciting, as this combination not only provides a powerful opportunity to establish causal relationships, but also stands to provide insights into the molecular mechanisms underlying these effects that would otherwise remain concealed. As technologies advance, and ever more insightful experiments can be performed at lower costs, the progressive hybridization of GBR and functional studies seems likely.

Ultimately, GBR trials are only one of many approaches that can be used to incorporate genetics into biomedical research, and they are no panacea. However, they have clear advantages where statistical power is a rate limiter, such as in studies of gene-environment and gene-treatment interactions, where the genotypes of interest are exceptionally rare, or where the cost of deploying the desired interventions or phenotyping technologies is prohibitive.
Figure legends:

Figure 1. Genotype-based recall randomized controlled trials using single variant (GBRsv) (panel A) and multivariant (GBRmv) (panel B) approaches.

Figure 2. Combining in vivo genotype-based recall (GBR) interventional trials with ex vivo or in vitro genome engineering experiments in human cells provides a potentially powerful type of hybrid research that can be used to demonstrate the presence of gene-treatment interactions, as well as determine the functional basis to these effects. The example shown here focuses on lifestyle interventions, but the principle also applies to drug interventions.

Sources of funding:

PWF is supported by grants from the Swedish Heart-Lung Foundation, the Novo Nordisk Foundation, the Swedish Research Council, the Innovative Medicine Initiative (DIRECT: grant agreement #115317; RHAPSODY: grant agreement #115881), the Swedish Foundation for Strategic Research (Industrial Research Center in Precision Medicine) and the European Research Council (CoG-2015_681742_NASCENT). Some of the concepts described herein emanate from the CoG-2015_681742_NASCENT research program (written by PWF).

NJT is a Wellcome Trust Investigator (202802/Z/16/Z), is the PI of the Avon Longitudinal Study of Parents and Children (MRC & WT 102215/2/13/2), is supported by the University of Bristol NIHR Biomedical Research Centre (BRC), the MRC Integrative Epidemiology Unit (MC_UU_12013/3) and works within the CRUK Integrative Cancer Epidemiology Programme (C18281/A19169).

Disclosures:

PWF is a consultant for and has stock options in Zoe Global and has received research funding from Boehringer Ingelheim, Eli Lilly, Janssen, Novo Nordisk, Sanofi Aventis, and Servier.

References:

