

Published in final edited form as:

Obesity (Silver Spring). 2016 August ; 24(8): 1630–1638. doi:10.1002/oby.21554.

Body mass index: has epidemiology started to break down causal contributions to health and disease?

Laura J. Corbin and Nicholas J. Timpson*

MRC Integrative Epidemiology Unit (IEU) at the University of Bristol

Abstract

Objective—To review progress in understanding the methods and results concerning the causal contribution of body mass index to health and disease.

Method—In the context of conventional evidence focused on the relationship between BMI and health, we consider current literature on the the common, population based, genetic contribution to body mass index BMI and how this has fed into the developing field of applied epidemiology.

Results—Technological and analytical developments have driven considerable success in identifying genetic variants relevant to BMI. This has enabled the implementation of Mendelian randomization to address questions of causality. The product of this work has been the implication of BMI as a causal agent in a host of health outcomes. Further breakdown of causal pathways by integration with other omics technologies promises to deliver additional benefit.

Conclusion—Considerable progress has been made, though gaps remain in our understanding of BMI as a risk factor for health and disease. Whilst promising, applied genetic epidemiology should be considered alongside alternative methods for assessing the impact of BMI on health and in light of potential limitations which relate to inappropriate or nonspecific measures of obesity and the improper use of genetic instruments.

Keywords

body mass index; causal analyses; epidemiology; Mendelian randomization; obesity; Adiposity; BMI

License (if available): Unspecified

*corresponding author: CONTACT INFO: MRC Integrative Epidemiology Unit, Oakfield House, Oakfield Grove, Bristol, BS8 2BN. laura.corbin@bristol.ac.uk. N.J.Timpson@bristol.ac.uk.

Disclosure: The authors declared no conflict of interest.

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via Wiley at onlinelibrary.wiley.com/doi/10.1002/oby.21554/abstract. 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/about/ebri-terms>

Introduction

The pandemic status of high BMI (obesity) has been attributed to the rise of an “obesogenic” environment which tips the balance between energy intake and energy expenditure, driving individuals towards increased adiposity along environmentally determined lines (1, 2). Despite this it is important to realise that, within the same environment, not all individuals become overweight or obese and those that do have differential disease risk. In reality, a complex interplay of both genetic and environmental factors must be considered in order to better understand BMI and why it appears to have such a great health impact. Focusing on BMI specifically, whilst it is absolutely clear that there are strong and replicable associations between this risk factor and health, the interpretation of existing associations is not straightforward. In reality, a complex interplay of both genetic and environmental factors must be considered in order to better understand BMI as a phenotypic proxy for adiposity, why it appears to have such a great health impact and how this impact might be mitigated in future both at the individual and the population level.

The underlying aetiology of relationships between BMI and health outcomes is clearly complex and likely to be heterogeneous across differing populations, apparently healthy individuals and those with disease. It is perhaps not surprising that efforts to counter the impact of ~2.3billion overweight, >700million obese and ~\$100billion per annum care bill (including targeted dietary intervention (3), weight loss programmes (4, 5) and pharmaceutical interventions (6, 7)) have failed to deliver lasting reductions in BMI (>2yr) at least at the level of the population. Currently, the only effective intervention for weight reduction is bariatric surgery, which is costly, not favourable for the treatment of moderate obesity (8).

Questions therefore remain as to why we continued to focus on BMI when we struggle to understand it as a measurement and fail to control or augment it at a policy or population level? More so, if we are content that the ease of BMI measurement is a justification for continued use, how might we gain insight into how and why BMI appears to be causally related to disease? In the context of conventional evidence focused on the relationship between BMI and health, this review aims to consider the current literature around the common, population based, genetic contribution to BMI/adiposity and how this has fed into the developing field of applied epidemiology in an effort to assess if and how the metric kg/m^2 exerts a causal effect on health. In doing this, we will discuss the complications of measurement, complex genetic aetiology and idiosyncrasy of human phenotyping (and its effect on analysis and inference) before attempting to suggest likely future moves for BMI research.

Conventional approaches to the analysis of BMI and health

More than 1.9 billion adults, 18 years and older, were overweight with over 600 million obese in 2014. This represents the worldwide prevalence of obesity more than doubling between 1980 and 2014 and the consequences of this are put into morbid focus when one is reminded that raised BMI is a substantial risk factor for disease cardiovascular disease (which were the world leading cause of death in 2012), diabetes, musculoskeletal disorders

and some cancers (endometrial, breast, and colon) (World Health Organisation, 2015). The evidence for these relationships comes from a variety of sources, but importantly the relative simplicity of height and weight measurement has allowed for the formation and analysis of substantial BMI related data sets focused on these relationships.

Examples of this include the Prospective Studies Consortium (able to assess observational relationships between baseline BMI and mortality in a collection of 57 studies delivering 894,576 participants mostly from Western Europe and North America(9) and an equally well sized initiative in Asian population based samples (including more than 1.14 million people recruited in 20 cohorts in Asia(10)) which have been able to give estimates as to the likely contribution of BMI variation to the risk of death and specific disease outcomes. Away from population specific differences hinting at the potential importance of body composition in BMI related effects, the relationship between BMI and mortality (with a marked cardiovascular component) is broadly consistent. Whilst not proven to be fully causal, these studies present a compelling illustration of these relationships.

This type of work has not been limited to the collection of semi-focused, large-scale investigations of mortality and common disease outcome, but also has been undertaken in a manner targeting specific disease outcomes. For example, a detailed examination of UK Clinical Practice Research Datalink (CPRD, www.cprd.com) was able to characterize the observational associations between BMI and cancer risk for the 22 most common cancer sites seen in UK medical record data(11). In this work, more than 5 million individual records reporting over 160000 cancers were investigated yielding evidence of association between BMI and 17 of 22 disease sub-types. Outside of likely confounding events driven by smoking, compelling association between BMI and disease risk are evident for cancers of the uterus, kidney, thyroid and leukaemia with more complex association signatures seen for liver, colon, ovarian breast cancers and together add to the growing range of non-specific disease risk alterations that appear linked to population based fluctuations in BMI.

Outside the realm of observational epidemiology, interventional studies in the form of randomized control trials (RCTs) have of course been applied. The most commonly evaluated interventions for BMI involve modifications to diet and/or physical activity levels as implemented in both children (12, 13) and adults (5, 14, 15, 16, 17). There are also RCT that have tested the efficacy of pharmacological interventions, most often alongside behavioural changes with the most commonly tested agent being Orlistat (4, 18, 19). However, the potential for pharmacological intervention is somewhat limited due to a lack of suitable drugs with favourable properties (20). Whilst most, although not all (12, 13), behavioural and/or pharmacological interventions result in a reduction in adiposity (as assessed by BMI or body weight), a major limitation of these studies with respect to inferring causality between BMI and health, is the lack of long-term follow-up, with a 12-month endpoint being typical. Therefore, the conversion of this reduction in BMI to a reduced incidence of disease later in life is not well evidenced. Indeed, apparently beneficial changes in cardiovascular risk factors, such as lipid profile and blood pressure, have been used to bolster conclusions regarding health benefits despite results from at least one longer-term trial suggesting that the assumed link between these intermediates and cardiovascular mortality may not be valid (14).

In contrast to most behavioural intervention studies, RCTs of surgical intervention have had longer followup periods allowing a more direct assessment of the impact of weight reduction on mortality. The long-term health impact results of RCTs for surgical intervention have been mixed and whilst there is evidence of a reduction in mortality following surgery (21, 22, 23), concerns have been raised around the potential for differences between surgical cases and untreated controls to complicate analyses (22). There is also no assurance that the effects seen after these procedures is directly related to BMI/weight reduction, with short term impact of surgery being marked and arguably independent of weight(24). Furthermore, the cost-effectiveness of surgery depends on the patient's level of obesity on admittance and the relative improvement in quality of life and health achieved subsequently (25).

Taken together, whilst there is a deep literature focused on the examination of associations between BMI and common health outcomes within both observational and intervention designs, these approaches remain limited in their ability to assess the causal contribution of BMI to disease. Observational studies have been undertaken at scale, but retain the conventional limitations to inference in confounding, bias and reverse causation and although trials of BMI intervention are conceptually more inceptive, limitations to the interventions themselves and the ability to alter BMI hamper the interpretation of long term health implications.

Genetic contributions to BMI

Common form obesity, assessed simply by BMI level and which does not segregate in families, has a multifactorial basis. Individuals may carry any number of common genetic variants which contribute to variation in BMI at the level of the population, but most of these exert only small effects on adiposity. Genomewide association studies (GWAS) employ a hypothesis-free approach to identify variants consistently associated with complex traits (26) and use genotyping chips with the ability to score hundreds of thousand to millions of single nucleotide polymorphisms (SNPs) positioned across the entire genome. This approach has revolutionised the search for genetic associations and the interrogation of the common disease/common variant hypothesis specifically (27) and in the case of BMI, the first real progress in the application of GWAS approaches came with a study of just under 40,000 participants and from an initial search for type 2 diabetes loci (28). This work identified a locus with common variants reliably associated with BMI where carriers of two copies of the minor allele at *FTOrs9939609* were on average 3kg heavier than the major allele carriers (29). Immediately after this first wave of GWAS analyses, it was acknowledged that substantially larger sample sizes, greater genomic coverage through advanced reference panel use and imputation (30) and more rigorous discovery and replication phases through extensive consortia derived meta-analysis were needed to fully explore the common genetic contribution to complex traits like BMI (31, 32). The most recent of these involves 125 independent cohort studies and totalling nearly 340,000 participants and has brought the list of confirmed associated genetic variants to 97 (Figure 1) (33).

Despite the recent success in identifying and verifying almost 100 loci with confirmed BMI association, together they only explain in the region of 2.7% of the phenotypic variance in BMI (33). Even with the addition of these new associated genetic variants, it is evident that

although each new (bigger) GWMA offers new biological insight through novel gene discovery, the newly discovered associations are the product of larger studies and not larger effects (Figure 2). Saving the scaling up of population based sequencing initiatives with the capacity to score rare variants(34), the next steps are therefore to make use of the variants we have to try to understand the effects of BMI. Importantly even small genetic effects are potentially useful for this in the correct conditions and the development of MR has given utility to the “so what” gene variant associations GWAS is efficient at capturing.

Applied genetic epidemiology and Mendelian randomisation

Developments in the genetics of obesity have opened up a new avenue of investigation to researchers interested in dissecting the relationship between BMI and health – Mendelian randomisation (MR). In contrast to direct measurement, germline genotypes reliably associated with risk factors can act as proxy measurements for risk factors offering several advantages: Genotypes are relatively easy to measure, are stable through time, are largely immutable and are not correlated with confounding factors as a result of the mechanisms of Mendelian inheritance(35, 36). An alternative approach to the analysis of BMI is therefore to use genetic predictors to act as proxies of the feature (or exposure) one is concerned with in order to help investigate causality(37, 38, 39)(Figure 3A&B). In MR, genetic variation fulfils the role of an instrumental variable (40) where the presence of variance in BMI explained by genotype is orthogonal to confounding factors and where genotype is assumed only to exert effect on health outcome through BMI. Whilst these assumptions are clearly open to challenge (through pleiotropy and other phenomena discussed at the end of this review), this approach provides an important contribution to the weight of evidence that may exist around a given epidemiological association. It is important to note that applied genetic epidemiology and MR is just one approach to the assessment of causality outside of RCTs(41).

The first application of MR to BMI followed rapidly on from the discovery of *FTO*(rs9939609) and examined 10 metabolic traits. Authors of this study concluded that the *FTO* genotype was associated with metabolic traits to an extent entirely consistent with its effect on BMI although power limitations meant causal relationships could only be confirmed for fasting insulin, glucose, triglycerides and lower high-density lipoprotein cholesterol (HDL-C) (42). By exploiting the ability of genetic variants to model lifetime exposure, researchers have also been able to explore the potential long-term effects of increased adiposity on health. To date, MR studies using BMI-associated variants have provided evidence of a causal effect of greater adiposity on a number of indicators of reduced cardiovascular health, including increased blood pressure (43), increased fasting glucose and insulin (44), decreased HDL-C (44) and increased systemic inflammation (45). Causal inference with respect to complex diseases is challenging however, a causal role for increased adiposity has been evidenced for type 2 diabetes T2D) (44) and ischaemic heart disease (IHD) (46). There are also examples of MR being applied to outcomes and traits beyond the classical cardiometabolic outcomes, including mental health (47, 48), childhood asthma (49), bone mass in childhood (50), uric acid (51), cancer (52) and trans-generational effects such as foetal over nutrition (40).

As well as these simple investigations of the causal impact of BMI on health-related factors, bidirectional assessments (39) have also been undertaken (Figure 3C). For example, exploring the relationship between the acute phase reactant C-reactive protein (CRP) and BMI, bidirectional MR has been used to exploit variation at independent BMI and CRP associated variants to evaluate whether BMI had a causal effect on CRP and simultaneously variation at the *CRP* loci to assess whether CRP had a causal effect on BMI (53). This work has provided evidence implicating BMI as a causal agent in inflammation and asserting directionality in an otherwise unclear network of complex phenotypes (45, 54). Other cases where causal association is likely to run in both directions or suitable instruments are unavailable, results may be less clearly interpreted. Work implicating BMI in the aetiology of activity patterns in young participants recently illustrated this point - whilst the impact of BMI on activity was marked and likely to be real, the reciprocal relationship (which is likely to be present) was not possible to either exclude or describe precisely (55).

One of the key parts of the MR process is the referral of evidence for causal relationships generated through the use of genetic data back to the existing observational estimates. A good example of this and the potential impact of MR analyses can be seen in that of C-reactive protein and the marked differences in effect estimates that have been generated from observational and MR analyses with respect to effects on cardiovascular health outcomes (56, 57, 58, 59, 60). In this case, it is the contrast between MR derived and observational estimates which provides information given a lack of support for strong observational effects. In contrast to this, one of the paradoxically dissatisfying observations from MR analyses of BMI as a risk factor is the breadth of apparently causal associations (bar perhaps dental caries and depression and foetal overnutrition (61, 62, 63, 64).

Blood pressure and cardiovascular health outcomes serve to illustrate this well with both within study (Figure 4) and between studies (9, 43), however this is the case elsewhere. Bone health (50), cancer (52), asthma (49), T2D (27, 44), osteoarthritis (65) have all shown some level of agreement in the associations delivered by the best available observational studies and MR. Indeed a novel approach to the examination of BMI as a causal risk factor (66) looking across a large number of possible intermediate phenotypes was able to chart a broad spread of BMI related effects. This highlights one of the key problems in the analysis of BMI in that, unlike others so far, there seems to be an underlying causal contribution to a large number of health related outcomes. This therefore leaves questions as to what are the underlying pathways and mechanisms responsible for these apparently causal relationships flagged by the broad exposure "BMI" and also as to the validity of these MR tests.

BMI effects: breaking down pathways

Despite the demonstration of the likely causal relationships between BMI and IHD, it is possible that weight itself is not the causal agent in disease, rather that there are a suite of intermediate phenotypes between BMI and outcome that deliver risk. Work not dissimilar to that originally published exploring the downstream impact of genetic variation at the *FTO* locus on cardiometabolic traits (42) used multiple intermediate phenotypes and also the health outcomes T2D and CHD (44) to try and unpick the pathways of BMI effect. In a relatively large collection of European participants (4,407 T2D, 6,073 CHD, and 3,813

stroke cases) the causal effect of a change in BMI of $1\text{kg}/\text{m}^2$ on fasting glucose, fasting insulin, interleukin-6, systolic blood pressure, reduced HDL-C and low-density lipoprotein cholesterol (LDL-C) was estimated alongside the change in odds for disease outcome given the same exposure. This work was able to identify a host of intermediate risk factor associations and related this to the strongest health outcome effects (T2D), although additional power was needed to obtain precise estimates of BMI effect on CHD.

Coming from a similar methodological approach, MR can be employed in network analyses to identify causal risk factors in efforts to locate refined targets for therapeutic intervention (67). Any network of observational associations can be explored by the use genetic variants that independently predict the nodes of that network. A simple exposition of this network MR approach has been applied to break down the association between BMI and IHD (believed to be causal from previous MR analysis (46)) through the use of available intermediate risk factors measured in a population and novel instruments for each of them that have come from new GWAS studies. In an example of this, the Copenhagen General Population Study (N=71,407) and the Copenhagen City Heart Study (N=10,314) and a case-control study, the Copenhagen Ischemic Heart Disease Study (N=5,262), have been used to attempt to further dissect the BMI/IHD association using MR and causal mediation analyses (68) (Figure 5). This work suggests that it is likely that BMI driven elevations of non-fasting remnant cholesterol and LDL-C, elevated blood pressure and possibly elevated non-fasting glucose levels may contribute causally to risk of IHD.

Intermediate phenotype and the “omics” revolution

Mendelian randomization studies have generally focused on a limited number of intermediate phenotypes, but recent applications of omic technologies into large scale population-based studies present new opportunities for identifying predictive biomarkers and causal links between established phenotypes and disease outcomes (69, 70, 71, 72, 73, 74). This has particular gravity for the types of network analysis being proposed above where, the combination of large-scale genetics data (and successful associations) are able to provide genetic proxy measures for an equally large collection of pathway specific intermediate risk factors. There is of course no guarantee that use of multi-omic phenotype data will avoid any of the problems encountered in observational epidemiology, but in combination with MR approaches, there is an opportunity to undertake network MR at scale. Omic technologies are now generating phenotypic data at a staggering rate and the use of these data in large scale population-based studies is presenting new opportunity for identifying novel predictive biomarkers and causal links between established phenotypes and disease outcomes (75, 76, 77, 78, 79, 80).

As an example of this, it is known that metabolite profiles are useful in the prediction of cardiometabolic disease (81, 82), but that their role as modifiable targets for intervention or causal mediators of disease risk remains unclear. It is also known that many metabolites have substantial heritability and that it is possible to find robust associations between genetic variation and the same metabolite features (83, 84). Together, it is then possible to examine the causal role of risk factors (in this case BMI) in the formation of metabolomics profiles(85) and then in a second stage (currently not applied to BMI systematically) to

consider the causal role of those BMI driven metabolites in disease outcomes. This is a process termed two-step MR(69) and when applied across multiple collections with measurements of BMI, intermediate phenotypes (such as metabolites) and health outcomes, has the potential to informatively reduce the omics measure data space (to a set of anchoring genetic variants(39)) and to breakdown causal pathways to disease.

Challenges and limitations in the causal analysis of BMI

(i) Measurement – the idiosyncrasy of human phenotyping

The options available to both clinicians and researchers for measuring adiposity are many and varied. They range from simple indices of body weight for stature (e.g. BMI as focused on here) to detailed imaging protocols using magnetic resonance imaging (MRI), computed tomography (CT), ultrasound, DXA and electron microscopy. As well as permitting the differentiation of body fat compartments relevant to health, such as subcutaneous versus visceral fat, such advanced technologies as combined positron emission tomography (PET)-CT, allow even finer resolution such that quantities of brown fat (containing mitochondria-rich brown adipocytes) can be measured.

The statistical construct that is BMI ($\text{weight}(\text{kg})/\text{height}(\text{m})^2$) was first proposed as an index of relative bodyweight by Adolphe Quetelet in 1842 (86). Promoted more formally in 1972, BMI was suggested to be the optimum derivation for weight given stature based on the fact that, in a population of healthy men aged 18 to 60 years this index had lowest correlation with height and the highest correlation with measures of body fatness(87). Perhaps surprisingly, considering the technological advances of recent years, a recent commentary on the evolution of BMI came to much the same conclusion as that of Keys over forty years ago – that BMI is “*a robust, useful and surprisingly accurate measure of fatness in ‘healthy’ adults*” (88).

However, there remain two serious limitations of BMI as a measure of adiposity that are likely to hinder causal analyses. Firstly, there is the apparent inability of BMI to adequately describe body composition and related to the specific characteristics of different subgroups of the population. Comparisons of BMI both with alternative indices of body weight such as waist circumference, and with MRI and DXA derived measures of body composition have shown that BMI fails to discriminate well between major contributors to body composition. For example, studies have shown that short and tall individuals and those from different ethnicities have similar but not identical body compositions (89, 90). This type of limitation has a bearing on the generalisability of BMI as a measure and has driven the development of both alternative measures of adiposity, such as a body shape index (ABSI)(91) and modifications to BMI itself, for example, by optimising the power term for height to minimise its influence (92).

Secondly, BMI is clearly not specific. Whilst the correlation of BMI with health outcomes is undeniable, the biological interpretation of these relationships is complex. This problem was rather eloquently described by Wells who stated “*Paradoxically, it seems that the various limitations of BMI as a specific index of adiposity may also be its strengths as a composite index of cardio-metabolic risk*” (88). But this concept of BMI as a “*composite index of risk*”

is what makes its use in causal analyses so challenging. It is of course not impossible to consider the utility of specific genetic variants or collections thereof to help dissect more specific components of BMI, though this has not been systematically undertaken to date.

(ii) Undertaking MR - don't trust your genetic proxies

Statistical power, correlation between genetic variants (linkage disequilibrium (LD)), the non-specificity of genetic effects (pleiotropy), developmental plasticity (or “canalization” (93)) and population stratification have all been recognised as potential limitations to the MR approach(38). However, it has become possible to assess and overcome issues of statistical power, LD and population stratification through the combination of large data sets which are based in homogeneous population based collections and that use independent genetic variants for analyses. Canalisation and pleiotropy remain potentially serious limitations. The former of these has yet to really escape the bounds of theory and if present, may only act to nullify genetic associations before they are found. On the other hand, pleiotropy and more generally the blind use of biologically complex genetic variation (and potentially large collections of complex genetic proxies from GWAS studies) remains one of the real challenges to these applied approaches.

It is becoming increasingly clear that there are important potential complications in the formation of genetic proxy measures for MR studies. Through either the analysis of complex or derived phenotypic outcomes which can generate genetic associations which are driven by artefactual biases(94) or just the presence of complex biological underpinnings, the chosen genetic variation for MR studies may bring just as many complications as they appear to avoid. In circumstances where well-characterised, candidate driven and biologically understood genetic variants as proxy measures (relied upon in previous MR studies derived from smaller scale genetic association studies) are unavailable, but where extremely large GWAS consortia yield apparently reliable association signals, it is tempting to use exhaustive lists of genetic risk factor in a genetic risk scores to undertake MR analyses (95).

Taking the example of educational attainment (a complex, biologically distal and poorly measured phenotype not dissimilar to BMI), a large-scale GWAS identified three genetic variants reliably correlated with education (96) but these signals represent less than a tenth of the expected difference between girls and boys in educational attainment (97). Faced with the lack of a strong genetic proxy for substantive MR study, an alternative strategy is to generate composite genetic proxy measures from collections of genomewide data (easily done through blind application of refined software interfaces such as PLINK (98)). In this example specifically, a composite genetic proxy measure for educational attainment can explain up to ~3% of the variance in this exposure would therefore be a valuable tool for MR studies focused on compelling hypotheses such as the impact of education on income/ lifetime earning ability.

However, the formation of genetic proxy measures in this way can have complex flaws. Through the combined impact of genetic contributions from many different biological pathways and the possible biasing effects of pleiotropy, the use of genomewide proxy measures can produce effect estimates that are biased towards the confounded observational estimates MR is attempting to avoid. Furthermore, with the expansion of GWAS study

sample size and power and the consequent discovery of increasingly distal contributions to outcome variance, looks set to introduce these complications even in the presence of apparently robust genetic association discoveries. The expansion of genetic association consortia for the analysis of BMI is now spilling way over $n=300000$ and with targets of up to $n=1000000$ in a single meta-analysis, the abundance of genetic proxy measures for BMI is set to grow. It is therefore with these limitations in mind that we should approach the use of novel findings that carry with them as much complication as clarification.

This is not to remove applied genetic epidemiology and MR as a logical extension to the analysis of causal relationships, rather to suggest that in an era of proliferation for genetic analysis, we should remain sceptical of the performance of any one analysis type. Triangulation of evidence should be sought where possible and MR viewed as a valuable contributor rather than a sole answer(37, 38). What is clear is that the success of MR and its move to mainstream analysis should not become the worst enemy of this approach. Furthermore, the growing presence of high quality functional biological data to help understand genetic associations and novel statistical approaches(99) to undertaking MR will help to relieve some of the problems mentioned above.

Conclusion

This review has considered major contributions to non-genetic approaches to assessing the causal impact of BMI on human health and current knowledge concerning the common, population based, genetic contribution to BMI and how this has fed into the developing field of applied epidemiology. We have revised complications of complex genetic aetiology and phenotypic measurement, and considered potential development and application of multiple omic data sources to help unpick the largely misunderstood relationships between BMI and human health and disease. Lastly, we have brought to attention the importance of appropriate use of applied genetic analyses in that whilst potentially complex, the ability to de-confound and add clarity to the prevailing weight of evidence is a superb possibility in suitable conditions.

Obesity and adiposity, measured principally via the faithful stand-in BMI, is of course a major risk factor when considering variance in risk for all sorts of health outcomes. There have now been a series of established study designs (prospective observational studies and MR analyses in particular) which have supported the notion of BMI as a causal agent in the formation of disease risk. For any given patient, however, it is unlikely to be the label “ 33kg/m^2 ” that causes morbidity or mortality. Understanding the detailed routes from the biology reported (on average) by BMI to disease by employing new measurement techniques and through advanced causal analysis methods will be crucial for future preventative medicine. Combinatorial investigations incorporating multi-omic examination of patients going through radical changes in BMI via surgical intervention, population based analyses of BMI affect through MR and analyses aimed at identifying modifiable risk factors able to modify exposure will be essential to the future breakdown and understanding of how BMI exerts a causal effect on health.

Funding

Authors are supported by the Medical Research Council (MC_UU_12013/3).

References

1. Hill JO, Peters JC. Environmental Contributions to the Obesity Epidemic. *Science*. 1998; 280:1371–1374. [PubMed: 9603719]
2. Swinburn B, Egger G, Raza F. Dissecting Obesogenic Environments: The Development and Application of a Framework for Identifying and Prioritizing Environmental Interventions for Obesity. *Preventive Medicine*. 1999; 29:563–570. [PubMed: 10600438]
3. Wolf AM, Colditz GA. Current estimates of the economic cost of obesity in the United States. *Obes Res*. 1998; 6:97–106. [PubMed: 9545015]
4. Dombrowski SU, Knittle K, Avenell A, Araújo-Soares V, Snihotta FF. Long term maintenance of weight loss with non-surgical interventions in obese adults: systematic review and meta-analyses of randomised controlled trials. *BMJ*. 2014; 348
5. The Look AHEAD Research Group. Cardiovascular Effects of Intensive Lifestyle Intervention in Type 2 Diabetes. *New England Journal of Medicine*. 2013; 369:145–154. [PubMed: 23796131]
6. Gray LJ, Cooper N, Dunkley A, Warren FC, Ara R, Abrams K, et al. A systematic review and mixed treatment comparison of pharmacological interventions for the treatment of obesity. *Obesity Reviews*. 2012; 13:483–498. [PubMed: 22288431]
7. Yanovski SZ, Yanovski JA. Long-term drug treatment for obesity: a systematic and clinical review. *JAMA*. 2014; 311:74–86. [PubMed: 24231879]
8. Picot J, Jones J, Colquitt J, Loveman E, Clegg A. Weight Loss Surgery for Mild to Moderate Obesity: A Systematic Review and Economic Evaluation. *OBES SURG*. 2012; 22:1496–1506. [PubMed: 22926715]
9. Prospective Studies Collaboration. Body-mass index and cause-specific mortality in 900000 adults: collaborative analyses of 57 prospective studies. *The Lancet*. 2009; 373:1083–1096.
10. Zheng W, McLerran DF, Rolland B, Zhang X, Inoue M, Matsuo K, et al. Association between Body-Mass Index and Risk of Death in More Than 1 Million Asians. *New England Journal of Medicine*. 2011; 364:719–729. [PubMed: 21345101]
11. Bhaskaran K, Douglas I, Forbes H, dos-Santos-Silva I, Leon DA, Smeeth L. Body-mass index and risk of 22 specific cancers: a population-based cohort study of 5.24 million UK adults. *The Lancet*. 2014; 384:755–765.
12. Wake M, Lycett K, Clifford SA, Sabin MA, Gunn J, Gibbons K, et al. Shared care obesity management in 3-10 year old children: 12 month outcomes of HopSCOTCH randomised trial. *BMJ*. 2013; 346
13. Willi SM, Hirst K, Jago R, Buse J, Kaufman F, El ghormli L, et al. Cardiovascular risk factors in multiethnic middle school students: the HEALTHY primary prevention trial. *Pediatric Obesity*. 2012; 7:230–239. [PubMed: 22461375]
14. The Look AHEAD Research Group. Long Term Effects of a Lifestyle Intervention on Weight and Cardiovascular Risk Factors in Individuals with Type 2 Diabetes: Four Year Results of the Look AHEAD Trial. *Archives of internal medicine*. 2010; 170:1566–1575. [PubMed: 20876408]
15. Dansinger ML, Gleason J, Griffith JL, Selker HP, Schaefer EJ. Comparison of the atkins, ornish, weight watchers, and zone diets for weight loss and heart disease risk reduction: A randomized trial. *JAMA*. 2005; 293:43–53. [PubMed: 15632335]
16. Roberts CK, Ng C, Hama S, Eliseo AJ, Barnard RJ. Effect of a short-term diet and exercise intervention on inflammatory/anti-inflammatory properties of HDL in overweight/obese men with cardiovascular risk factors. *Journal of Applied Physiology*. 2006; 101:1727–1732. [PubMed: 16902063]
17. Groeneveld IF, Proper KI, van der Beek AJ, van Mechelen W. Sustained body weight reduction by an individual-based lifestyle intervention for workers in the construction industry at risk for cardiovascular disease: Results of a randomized controlled trial. *Preventive Medicine*. 2010; 51:240–246. [PubMed: 20692282]

18. Davidson MH, Hauptman J, DiGirolamo M, Foreyt JP, Halsted CH, Heber D, et al. Weight control and risk factor reduction in obese subjects treated for 2 years with orlistat: a randomized controlled trial. *JAMA*. 1999; 281:235–242. [PubMed: 9918478]
19. Richelsen B, Tonstad S, Rossner S, Toubro S, Niskanen L, Madsbad S, et al. Effect of orlistat on weight regain and cardiovascular risk factors following a very-low-energy diet in abdominally obese patients: a 3-year randomized, placebo-controlled study. *Diabetes Care*. 2007; 30:27–32. [PubMed: 17192328]
20. Rodgers RJ, Tschöp MH, Wilding JPH. Anti-obesity drugs: past, present and future. *Disease Models & Mechanisms*. 2012; 5:621–626. [PubMed: 22915024]
21. Christou NV, Sampalis JS, Liberman M, Look D, Auger S, McLean APH, et al. Surgery Decreases Long-term Mortality, Morbidity, and Health Care Use in Morbidly Obese Patients. *Annals of Surgery*. 2004; 240:416–424. [PubMed: 15319713]
22. Maciejewski ML, Livingston EH, Smith VA, et al. SURvival among high-risk patients after bariatric surgery. *JAMA*. 2011; 305:2419–2426. [PubMed: 21666276]
23. Sjöström L, Narbro K, Sjöström CD, Karason K, Larsson B, Wedel H, et al. Effects of Bariatric Surgery on Mortality in Swedish Obese Subjects. *New England Journal of Medicine*. 2007; 357:741–752. [PubMed: 17715408]
24. Knop FK, Taylor R. Mechanism of Metabolic Advantages After Bariatric Surgery: It's all gastrointestinal factors versus it's all food restriction. *Diabetes Care*. 2013; 36:S287–S291. [PubMed: 23882061]
25. Picot J, Jones J, Colquitt JL, Gospodarevskaya E, Loveman E, Baxter L, Clegg AJ. The clinical effectiveness and cost-effectiveness of bariatric (weight loss) surgery for obesity: a systematic review and economic evaluation. *Clinical Governance: An International Journal*. 2010; 15:1.
26. McCarthy MI, Abecasis GR, Cardon LR, Goldstein DB, Little J, Ioannidis JPA, et al. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat Rev Genet*. 2008; 9:356–369. [PubMed: 18398418]
27. The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007; 447:661–678. [PubMed: 17554300]
28. Zeggini E, Weedon MN, Lindgren CM, Frayling TM, Elliott KS, Lango H, et al. Replication of genome-wide association signals in UK samples reveals risk loci for type 2 diabetes. *Science*. 2007; 316:1336–1341. [PubMed: 17463249]
29. Frayling TM, Timpson NJ, Weedon MN, Zeggini E, Freathy RM, Lindgren CM, et al. A Common Variant in the FTO Gene Is Associated with Body Mass Index and Predisposes to Childhood and Adult Obesity. *Science*. 2007; 316:889–894. [PubMed: 17434869]
30. Marchini J, Howie B. Genotype imputation for genome-wide association studies. *Nat Rev Genet*. 2010; 11:499–511. [PubMed: 20517342]
31. Ioannidis JPA. Non-Replication and Inconsistency in the Genome-Wide Association Setting. *Human Heredity*. 2007; 64:203–213. [PubMed: 17551261]
32. Ioannidis JP, Boffetta P, Little J, O'Brien TR, Uitterlinden AG, Vineis P, et al. Assessment of cumulative evidence on genetic associations: interim guidelines. *Int J Epidemiol*. 2008; 37:120–132. [PubMed: 17898028]
33. Locke AE, Kahali B, Berndt SI, Justice AE, Pers TH, Day FR, et al. Genetic studies of body mass index yield new insights for obesity biology. *Nature*. 2015; 518:197–206. [PubMed: 25673413]
34. Walter K, Min JL, Huang J, Crooks L, Memari Y, et al. The UK10K project identifies rare variants in health and disease. *Nature*. 2015; 526:82–90. [PubMed: 26367797]
35. Davey Smith G, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S. Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. *PLoS Medicine*. 2007; 4:e352. [PubMed: 18076282]
36. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. *Stat Med*. 2008; 27:1133–1163. [PubMed: 17886233]

37. Davey Smith G, Ebrahim S. 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? *Int J Epidemiol.* 2003; 32:1–22. [PubMed: 12689998]
38. Davey Smith G, Ebrahim S. Mendelian randomization: prospects, potentials, and limitations. *Int J Epidemiol.* 2004; 33:30–42. [PubMed: 15075143]
39. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. *Human Molecular Genetics.* 2014; 23:R89–R98. [PubMed: 25064373]
40. Lawlor DA, Timpson NJ, Harbord RM, Leary S, Ness A, McCarthy MI, et al. Exploring the developmental overnutrition hypothesis using parental-offspring associations and FTO as an instrumental variable. *PLoS Medicine.* 2008; 5:e33. [PubMed: 18336062]
41. Richmond RC, Al-Amin A, Davey Smith G, Relton CL. Approaches for drawing causal inferences from epidemiological birth cohorts: A review. *Early Human Development.* 2014; 90:769–780. [PubMed: 25260961]
42. Freathy RM, Timpson NJ, Lawlor DA, Pouta A, Ben-Shlomo Y, Ruokonen A, et al. Common variation in the FTO gene alters diabetes-related metabolic traits to the extent expected given its effect on BMI. *Diabetes.* 2008; 57:1419–1426. [PubMed: 18346983]
43. Timpson NJ, Harbord R, Davey Smith G, Zacho J, Tybjaerg-Hansen A, Nordestgaard BG. Does greater adiposity increase blood pressure and hypertension risk?: Mendelian randomization using the FTO/MC4R genotype. *Hypertension.* 2009; 54:84–90. [PubMed: 19470880]
44. Holmes Michael V, Lange Leslie A, Palmer T, Lanktree Matthew B, North Kari E, Almoguera B, et al. Causal Effects of Body Mass Index on Cardiometabolic Traits and Events: A Mendelian Randomization Analysis. *The American Journal of Human Genetics.* 2014; 94:198–208. [PubMed: 24462370]
45. Welsh P, Polisecki E, Robertson M, Jahn S, Buckley BM, de Craen AJM, et al. Unraveling the Directional Link between Adiposity and Inflammation: A Bidirectional Mendelian Randomization Approach. *The Journal of Clinical Endocrinology and Metabolism.* 2010; 95:93–99. [PubMed: 19906786]
46. Nordestgaard BG, Palmer TM, Benn M, Zacho J, Tybjaerg-Hansen A, Davey Smith G, et al. The Effect of Elevated Body Mass Index on Ischemic Heart Disease Risk: Causal Estimates from a Mendelian Randomisation Approach. *PLoS Medicine.* 2012; 9:e1001212. [PubMed: 22563304]
47. Lawlor DA, Harbord RM, Tybjaerg-Hansen A, Palmer TM, Zacho J, Benn M, et al. Using genetic loci to understand the relationship between adiposity and psychological distress: a Mendelian Randomization study in the Copenhagen General Population Study of 53 221 adults. *Journal of Internal Medicine.* 2011; 269:525–537. [PubMed: 21210875]
48. Kivimäki M, Jokela M, Hamer M, Geddes J, Ebmeier K, Kumari M, et al. Examining Overweight and Obesity as Risk Factors for Common Mental Disorders Using Fat Mass and Obesity-Associated (FTO) Genotype-Instrumented Analysis: The Whitehall II Study, 1985–2004. *American Journal of Epidemiology.* 2011
49. Granell R, Henderson AJ, Evans DM, Smith GD, Ness AR, Lewis S, et al. Effects of BMI, Fat Mass, and Lean Mass on Asthma in Childhood: A Mendelian Randomization Study. *PLoS Med.* 2014; 11:e1001669. [PubMed: 24983943]
50. Timpson NJ, Sayers A, Davey-Smith G, Tobias JH. How does body fat influence bone mass in childhood? A Mendelian randomization approach. *J Bone Miner Res.* 2009; 24:522–533. [PubMed: 19016587]
51. Palmer TM, Nordestgaard BG, Benn B, Tybjærg-Hansen A, Davey Smith G, Lawlor DA, et al. Association of plasma uric acid with ischaemic heart disease and blood pressure: mendelian randomisation analysis of two large cohorts. *BMJ.* 2013; 347
52. Brennan P, McKay J, Moore L, Zaridze D, Mukeria A, Szeszenia-Dabrowska N, et al. Obesity and cancer: Mendelian randomization approach utilizing the FTO genotype. *Int J Epidemiol.* 2009; 38:971–975. [PubMed: 19542184]
53. Timpson NJ, Nordestgaard BG, Harbord RM, Zacho J, Frayling TM, Tybjærg-Hansen A, et al. C-reactive protein levels and body mass index: elucidating direction of causation through reciprocal Mendelian randomization. *Int J Obes.* 2011; 35:300–308.

54. Timpson NJ, Nordestgaard BG, Harbord RM, Zacho J, Frayling TM, Tybjaerg-Hansen A, et al. C-reactive protein levels and body mass index: elucidating direction of causation through reciprocal Mendelian randomization. *Int J Obes.* 2011; 35:300–308.
55. Richmond RC, Davey Smith G, Ness AR, den Hoed M, McMahon G, Timpson NJ. Assessing Causality in the Association between Child Adiposity and Physical Activity Levels: A Mendelian Randomization Analysis. *PLoS Med.* 2014; 11:e1001618. [PubMed: 24642734]
56. CRP CHD Genetics Collaboration. Collaborative pooled analysis of data on C-reactive protein gene variants and coronary disease: judging causality by Mendelian randomisation. *Eur J Epidemiol.* 2008; 23:531–540. [PubMed: 18425592]
57. CRP CHD Genetics Collaboration. Association between C reactive protein and coronary heart disease: mendelian randomisation analysis based on individual participant data. *BMJ.* 2011; 342:d548. [PubMed: 21325005]
58. Danesh J, Whincup P, Walker M, Lennon L, Thomson A, Appleby P, et al. Low grade inflammation and coronary heart disease: prospective study and updated meta-analyses. *Bmj.* 2000; 321:199–204. [PubMed: 10903648]
59. Timpson NJ, Lawlor DA, Harbord RM, Gaunt TR, Day INM, Palmer LJ, et al. C-reactive protein and its role in metabolic syndrome: mendelian randomisation study. *Lancet.* 2005; 366:1954–1959. [PubMed: 16325697]
60. Zacho J, Tybjaerg-Hansen A, Jensen JS, Grande P, Sillesen H, Nordestgaard BG. Genetically elevated C-reactive protein and ischemic vascular disease. *New England Journal of Medicine.* 2008; 359:1897–1908. [PubMed: 18971492]
61. Hung CF, Rivera M, Craddock N, Owen MJ, Gill M, Korszun A, et al. Relationship between obesity and the risk of clinically significant depression: Mendelian randomisation study. *Br J Psychiatry.* 2014; 205:24–28. [PubMed: 24809401]
62. Lawlor DA, Harbord RM, Tybjaerg-Hansen A, Palmer TM, Zacho J, Benn M, et al. Using genetic loci to understand the relationship between adiposity and psychological distress: a Mendelian Randomization study in the Copenhagen General Population Study of 53,221 adults. *Journal of Internal Medicine.* 2011; 269:525–537. [PubMed: 21210875]
63. Lawlor DA, Timpson NJ, Harbord RM, Leary S, Ness A, McCarthy MI, et al. Exploring the developmental overnutrition hypothesis using parental-offspring associations and FTO as an instrumental variable. *PLoS Medicine.* 2008; 5:e33. [PubMed: 18336062]
64. Shungin D, Cornelis MC, Divaris K, Holtfreter B, Shaffer JR, Yu YH, et al. Using genetics to test the causal relationship of total adiposity and periodontitis: Mendelian randomization analyses in the Gene-Lifestyle Interactions and Dental Endpoints (GLIDE) Consortium. *Int J Epidemiol.* 2015; 44:638–650. [PubMed: 26050256]
65. Panoutsopoulou K, Metrustry S, Doherty SA, Laslett LL, Maciewicz RA, Hart DJ, et al. The effect of FTO variation on increased osteoarthritis risk is mediated through body mass index: a Mendelian randomisation study. *Ann Rheum Dis.* 2014; 73:2082–2086. [PubMed: 23921993]
66. Millard LA, Davies NM, Timpson NJ, Tilling K, Flach PA, Smith GD. MR-PheWAS: hypothesis prioritization among potential causal effects of body mass index on many outcomes, using Mendelian randomization. *Sci.* 2015; 5:16645.
67. Burgess S, Daniel RM, Butterworth AS, Thompson SG, Consortium tE-I. Network Mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways. *Int J Epidemiol.* 2015; 44:484–495. [PubMed: 25150977]
68. Varbo A, Benn M, Smith GD, Timpson NJ, Tybjærg-Hansen A, Nordestgaard BG. Remnant Cholesterol, Low-Density Lipoprotein Cholesterol, and Blood Pressure as Mediators From Obesity to Ischemic Heart Disease. *Circulation Research.* 2015; 116:665–673. [PubMed: 25411050]
69. Relton CL, Davey Smith G. Two-step epigenetic Mendelian randomization: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol.* 2012; 41:161–176. [PubMed: 22422451]
70. Ala-Korpela M, Kangas A, Soininen P. Quantitative high-throughput metabolomics: a new era in epidemiology and genetics. *Genome Med.* 2012; 4:1–5. [PubMed: 22257447]
71. Nordström A, Lewensohn R. Metabolomics: Moving to the Clinic. *J Neuroimmune Pharmacol.* 2010; 5:4–17. [PubMed: 19399626]

72. Zhang G-F, Sadhukhan S, Tochtrop GP, Brunengraber H. Metabolomics, Pathway Regulation, and Pathway Discovery. *Journal of Biological Chemistry*. 2011; 286:23631–23635. [PubMed: 21566142]
73. Patti GJ, Yanes O, Siuzdak G. Innovation: Metabolomics: the apogee of the omics trilogy. *Nat Rev Mol Cell Biol*. 2012; 13:263–269. [PubMed: 22436749]
74. Davies M, Volta M, Pidsley R, Lunnon K, Dixit A, Lovestone S, et al. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biology*. 2012; 13:R43. [PubMed: 22703893]
75. Davies MN, Volta M, Pidsley R, Lunnon K, Dixit A, Lovestone S, et al. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. *Genome Biology*. 2012; 13:R43. [PubMed: 22703893]
76. Patti GJ, Yanes O, Siuzdak G. Innovation: Metabolomics: the apogee of the omics trilogy. *Nature Reviews Molecular Cell Biology*. 2012; 13:263–269. [PubMed: 22436749]
77. Zhang GF, Sadhukhan S, Tochtrop GP, Brunengraber H. Metabolomics, pathway regulation, and pathway discovery. *Journal of Biological Chemistry*. 2011; 286:23631–23635. [PubMed: 21566142]
78. Nordstrom A, Lewensohn R. Metabolomics: moving to the clinic. *Journal Of Neuroimmune Pharmacology: The Official Journal Of The Society On NeuroImmune Pharmacology*. 2010; 5:4–17. [PubMed: 19399626]
79. Ala-Korpela M, Kangas AJ, Soininen P. Quantitative high-throughput metabolomics: a new era in epidemiology and genetics. *Genome Med*. 2012; 4:36. [PubMed: 22546737]
80. Relton C, Davey Smith G. Two step epigenetic Mendelian randomisation: a strategy for establishing the causal role of epigenetic processes in pathways to disease. *Int J Epidemiol*. 2011 **AOP**.
81. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, et al. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature*. 2011; 472:57–63. [PubMed: 21475195]
82. Wang TJ, Larson MG, Vasani RS, Cheng S, Rhee EP, McCabe E, et al. Metabolite profiles and the risk of developing diabetes. *Nature Medicine*. 2011; 17:448–453.
83. Shin SY, Fauman EB, Petersen AK, Krumsiek J, Santos R, Huang J, et al. An atlas of genetic influences on human blood metabolites. *Nat Genet*. 2014; 46:543–550. [PubMed: 24816252]
84. Kettunen J, Tukiainen T, Sarin A-P, Ortega-Alonso A, Tikkanen E, Lyytikäinen L-P, et al. Genome-wide association study identifies multiple loci influencing human serum metabolite levels. *Nat Genet*. 2012; 44:269–276. [PubMed: 22286219]
85. Würtz P, Wang Q, Kangas AJ, Richmond RC, Skarp J, Tiainen M, et al. Metabolic Signatures of Adiposity in Young Adults: Mendelian Randomization Analysis and Effects of Weight Change. *PLoS Med*. 2014; 11:e1001765. [PubMed: 25490400]
86. Quetelet, A. A treatise on man and the development of his faculties. Burt Franklin; New York: 1842.
87. Keys A, Fidanza F, Karvonen MJ, Kimura N, Taylor HL. Indices of relative weight and obesity. *Journal of Chronic Diseases*. 1972; 25:329–343. [PubMed: 4650929]
88. Wells JC. Commentary: The paradox of body mass index in obesity assessment: not a good index of adiposity, but not a bad index of cardio-metabolic risk. *Int J Epidemiol*. 2014; 43:672–674. [PubMed: 24691953]
89. Heymsfield SB, Gallagher D, Mayer L, Beetsch J, Pietrobelli A. Scaling of human body composition to stature: new insights into body mass index. *The American Journal of Clinical Nutrition*. 2007; 86:82–91. [PubMed: 17616766]
90. Yajnik CS, Yudkin JS. The Y-Y paradox. *Lancet*. 2004; 363:163. [PubMed: 14726172]
91. Krakauer NY, Krakauer JC. A new body shape index predicts mortality hazard independently of body mass index. *PLoS ONE*. 2012; 7:e39504. [PubMed: 22815707]
92. Stergiakouli E, Gaillard R, Tavaré JM, Balthasar N, Loos RJ, Taal HR, et al. Genome-wide Association Study of Height-adjusted BMI in Childhood Identifies Functional Variant in ADCY3. *Obesity*. 2014; 22:2252–2259. [PubMed: 25044758]

93. Waddington CH. Canalisation of development and the inheritance of acquired characters. *Nature*. 1942; 150:563–565.
94. Aschard H, Vilhjálmsson Bjarni J, Joshi Amit D, Price Alkes L, Kraft P. Adjusting for Heritable Covariates Can Bias Effect Estimates in Genome-Wide Association Studies. *The American Journal of Human Genetics*. 2015; 96:329–339. [PubMed: 25640676]
95. Burgess S, Thompson SG. Use of allele scores as instrumental variables for Mendelian randomization. *Int J Epidemiol*. 2013; 42:1134–1144. [PubMed: 24062299]
96. Rietveld CA, Medland SE, Derringer J, Yang J, Esko T, Martin NW, et al. GWAS of 126,559 Individuals Identifies Genetic Variants Associated with Educational Attainment. *Science*. 2013; 340:1467–1471. [PubMed: 23722424]
97. Ward ME, McMahon G, St Pourcain B, Evans DM, Rietveld CA, Benjamin DJ, et al. Genetic Variation Associated with Differential Educational Attainment in Adults Has Anticipated Associations with School Performance in Children. *PLoS ONE*. 2014; 9:e100248. [PubMed: 25032841]
98. Chang C, Chow C, Tellier L, Vattikuti S, Purcell S, Lee J. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*. 2015; 4:7. [PubMed: 25722852]
99. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. *Int J Epidemiol*. 2015; 44:512–525. [PubMed: 26050253]
100. Loos RJF, Lindgren CM, Li S, Wheeler E, Zhao JH, Prokopenko I, et al. Common variants near MC4R are associated with fat mass, weight and risk of obesity. *Nat Genet*. 2008; 40:768–775. [PubMed: 18454148]
101. Willer CJ, Speliotes EK, Loos RJF, Li S, Lindgren CM, Heid IM, et al. Six new loci associated with body mass index highlight a neuronal influence on body weight regulation. *Nature genetics*. 2009; 41:25–34. [PubMed: 19079261]
102. Speliotes EK, Willer CJ, Berndt SI, Monda KL, Thorleifsson G, Jackson AU, et al. Association analyses of 249,796 individuals reveal 18 new loci associated with body mass index. *Nat Genet*. 2010; 42:937–948. [PubMed: 20935630]

Study Importance Questions

1. What major reviews have already been published on this subject?
 - a) Loos, R.J.F. (2012) Genetic determinants of common obesity and their value in prediction, *Best Practice & Research Clinical Endocrinology & Metabolism*, 26(2): 211-226.

This paper reviews the discovery of loci associated with obesity-related traits, and subsequently focuses on the body mass index (BMI) loci in particular to explore whether there is sufficient evidence for these loci to be used as clinical predictors. It does not consider use of loci in causal analyses.
 - b) Burgess, S. *et al.* (2015) Mendelian randomization: where are we now and where are we going? *International Journal of Epidemiology*, 44(2): 378-388.

This editorial considers developments in the methodology and application of Mendelian randomization to study causal mechanisms in health and disease over the past decade. It does not consider adiposity, BMI or obesity specifically.

2. What does our study add?

Progress in the field of applied genetic epidemiology and in particular in the application of Mendelian randomisation has been rapid in the last few years, driven largely by developments in a variety of omics technologies. This review provides a reflection on what has been achieved so far in dissecting the causal relationship between body mass index (BMI) and disease and gives comment on the likely future directions of this field. We also present a sobering discussion of the potential limitations of these approaches which are becoming commonplace in the field of complex trait analysis, especially for BMI in light of large-scale consortium science.

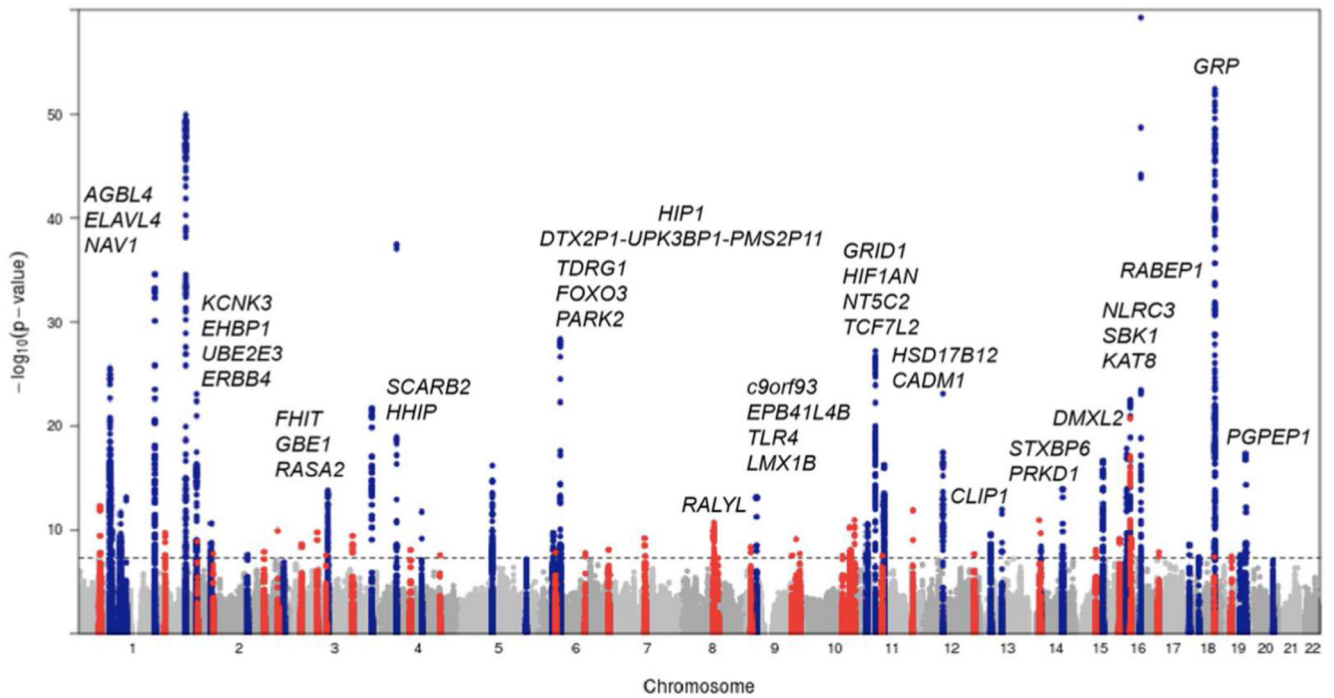


Figure 1. Manhattan plot showing body mass index (BMI)-associated variants with loci identified prior to 2015 in blue and novel loci identified by Locke *et al.* (33) in red.

Novel loci are labelled with the nearest gene, and the y-axis is truncated to allow easier observation of novel associations. This plot is reproduced from Locke *et al.* (33) with the permission of the authors.

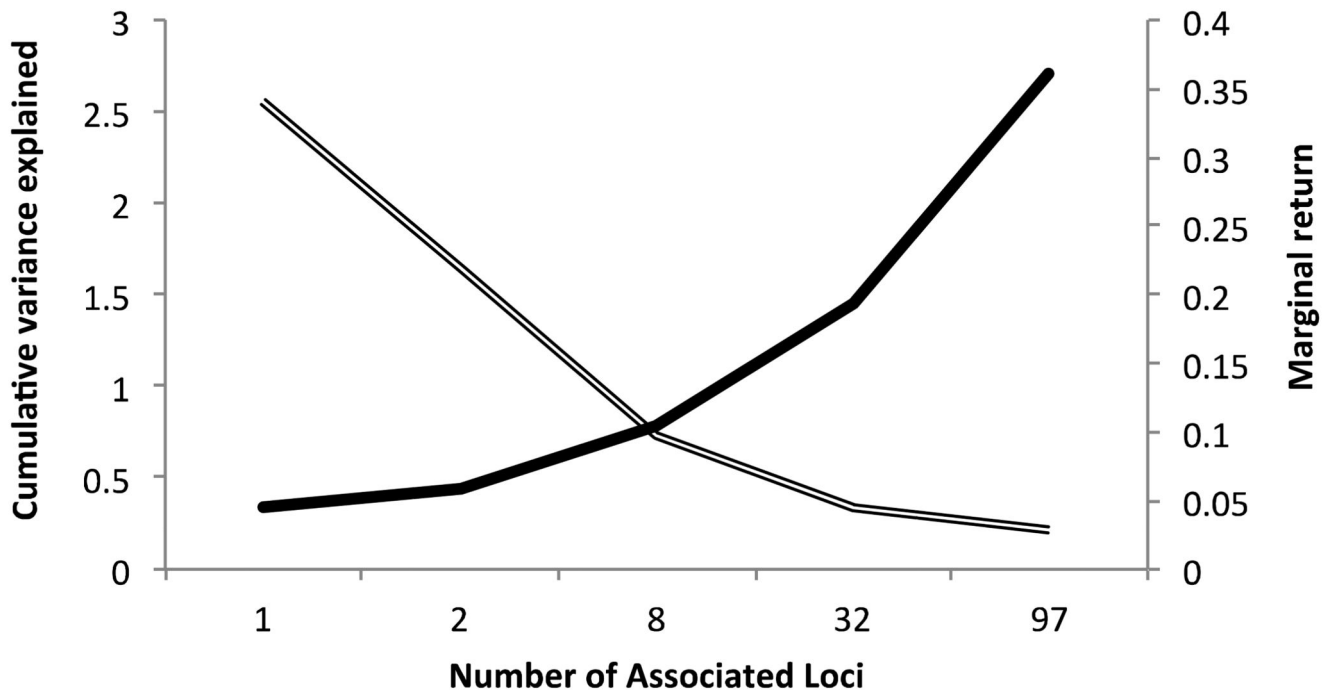


Figure 2. The interplay between increased variance explained and diminishing marginal return as the number of confirmed body mass index (BMI)-associated genetic loci has increased. The single line represents the cumulative variance explained and the double line the marginal return, calculated as the cumulative variance explained divided by the number of loci (29, 33, 100, 101, 102).

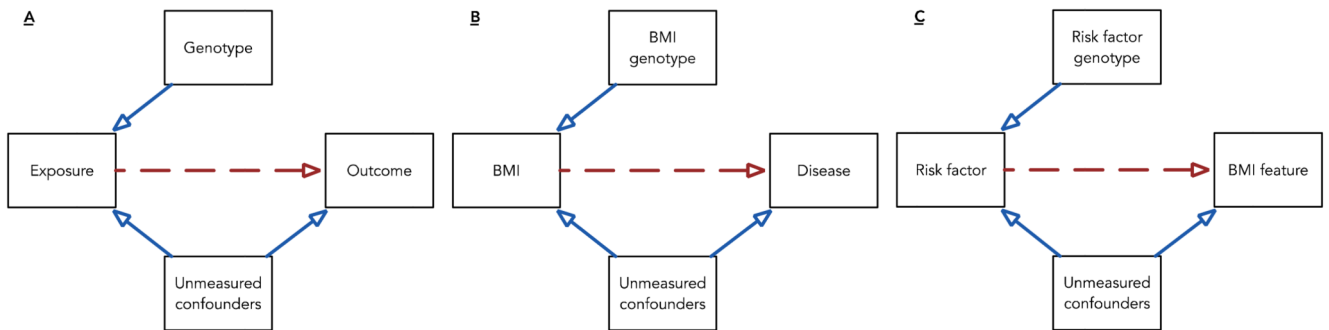


Figure 3. Mendelian randomisation; the use of genetic proxy measures of risk factors to allow causal inference.

(A) In general, a genotype of use to this study is associated with the exposure, is independent of measured or unmeasured confounders and can only influence outcome via the causal effect of the exposure.

(B) The presence or absence of association between the BMI associated genotype and disease risk (from existing genomewide association study data sets) give evidence that the BMI is a causal risk factor for disease.

(C) Here genotype acts as a proxy measure for an exposure potentially affecting the BMI in a reciprocal analysis. This type of reciprocal analysis allows for a triangulation or network approach to the assessment of the effects of and effects on BMI.

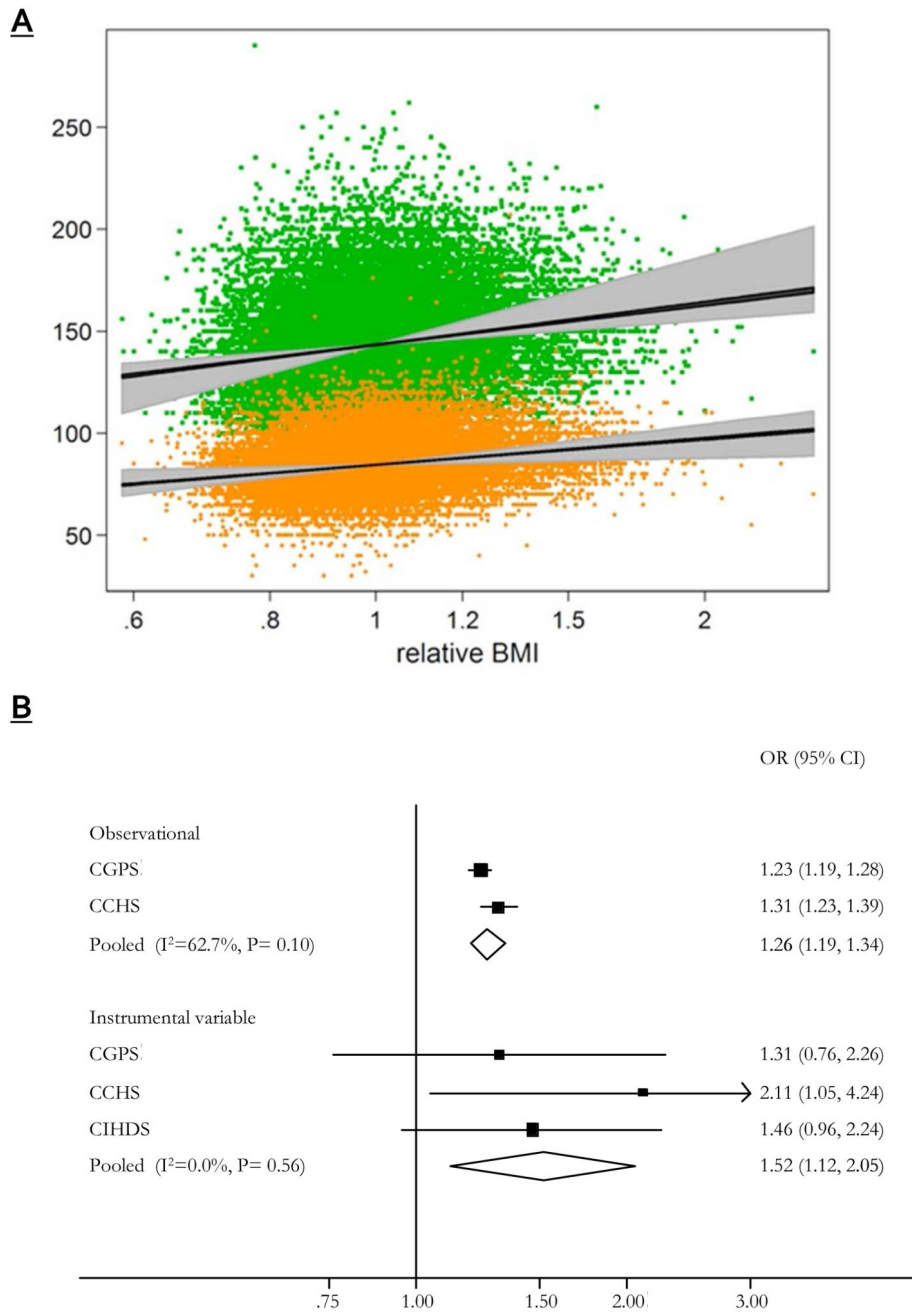


Figure 4. The comparison of observational and Mendelian randomisation derived estimates for blood pressure and ischaemic heart disease.

(A) Linear relationships between body mass index and blood pressure derived from observational and Mendelian randomization analyses. Upper scatter indicates systolic blood pressure and the lower diastolic. Grey areas around the estimated relationships indicate 95% CI for Mendelian randomisation estimates and in black those for observational estimates (plot generated from analysis for (43)). Note that for this analysis the log of body mass index was regressed on sex, age, age squared, log(height), and an age-sex interaction

and exponentiated to give an individual's "relative BMI," that is, the ratio between his or her actual BMI and that expected for his or her sex, age, and height.

(B) Meta-analysis forest plots of observational and instrumental variable estimates of the relationship between ischaemic heart disease and body mass index. Odds ratios are for a $4\text{kg}/\text{m}^2$ increase in body mass index (plot generated from analysis for (46)).

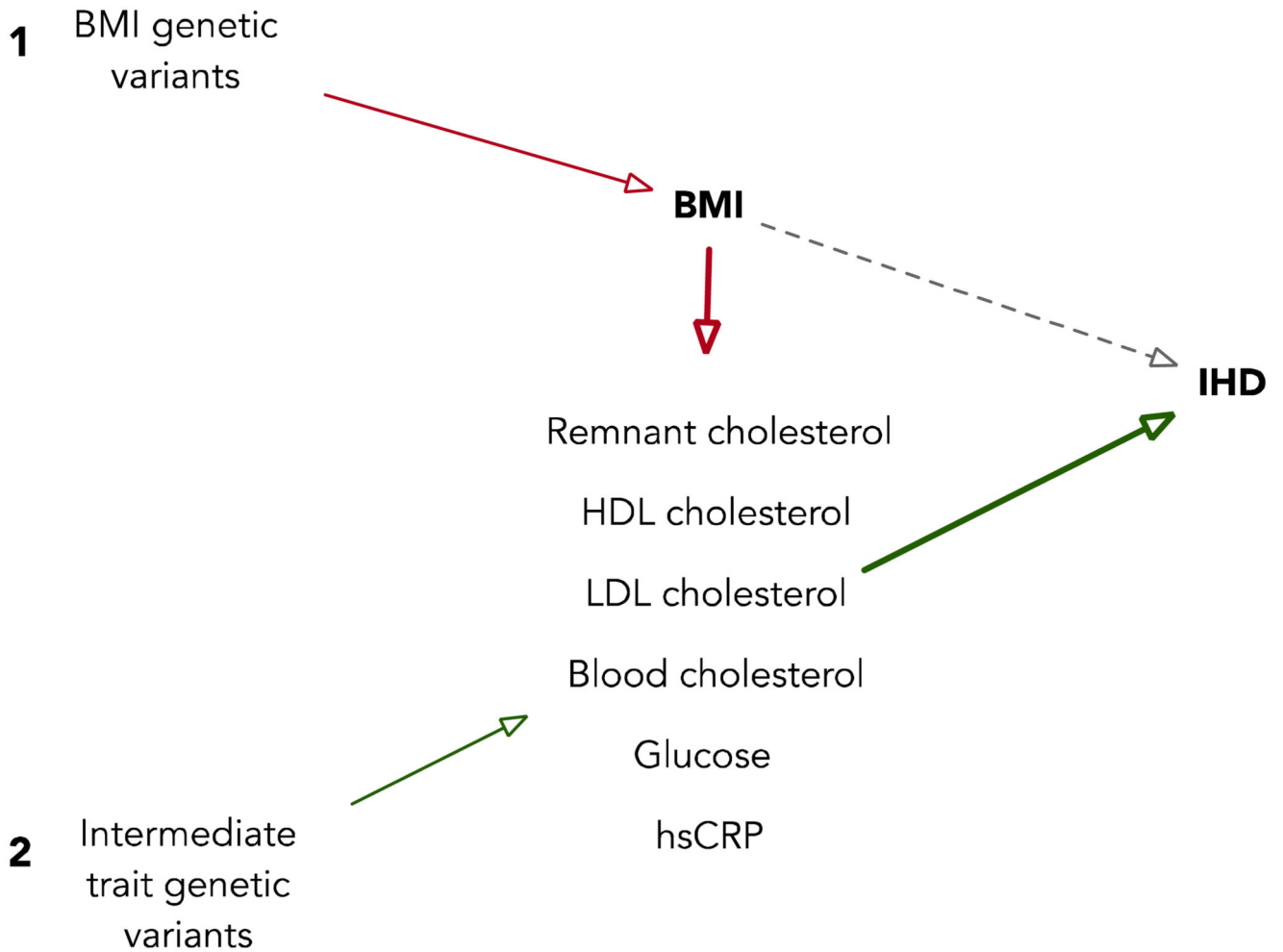


Figure 5. A two-step Mendelian randomization design applied to intermediate phenotype analysis in body mass index (BMI) and ischaemic heart disease (IHD).

In Step 1 (shown in red), BMI-associated variants are used to estimate the causal effect of BMI on relevant intermediates. In Step 2 (shown in green), variants associated with each of the intermediate traits are used to estimate the causal effect for those traits on IHD.