Research Techniques Made Simple: Using Genetic Variants for Randomization

Ashley Budu-Aggrey and Lavinia Paternoster

Observational epidemiological studies have identified associations between a number of modifiable exposures and outcomes, including in dermatology, such as between smoking and psoriasis. However, it is challenging to determine if such relationships are causal because of the potential of confounding and reverse causation. Mendelian randomization (MR) is a statistical method that can be used to investigate the causal relationships between an exposure and outcome by using a genetic instrument that proxies the exposure. The resulting estimate (under certain assumptions) can be interpreted as the causal estimate, free of confounding and reverse causation. In this review, we provide an overview of how to undertake an MR analysis, with examples from the dermatology literature. We also discuss the challenges and future directions of this method.

Journal of Investigative Dermatology (2019) **139,** 1416–1421; doi:10.1016/j.jid.2019.03.1138

CME Activity Dates: 20 June 2019
Expiration Date: 19 June 2020
Estimated Time to Complete: 1 hour

Planning Committee/Speaker Disclosure: Lavinia Paternoster is a consultant/advisor for Merck. All other authors, planning committee members, CME committee members and staff involved with this activity as content validation reviewers have no financial relationships with commercial interests to disclose relative to the content of this CME activity.

Commercial Support Acknowledgment: This CME activity is supported by an educational grant from Lilly USA, LLC.

Description: This article, designed for dermatologists, residents, fellows, and related healthcare providers, seeks to reduce the growing divide between dermatology clinical practice and the basic science/current research methodologies on which many diagnostic and therapeutic advances are built.

Objectives: At the conclusion of this activity, learners should be better able to:
- Recognize the newest techniques in biomedical research.
- Describe how these techniques can be utilized and their limitations.
- Describe the potential impact of these techniques.

CME Accreditation and Credit Designation: This activity has been planned and implemented in accordance with the accreditation requirements and policies of the Accreditation Council for Continuing Medical Education through the joint providership of Beaumont Health and the Society for Investigative Dermatology. Beaumont Health is accredited by the ACCME to provide continuing medical education for physicians. Beaumont Health designates this enduring material for a maximum of 1.0 *AMA PRA Category 1 Credit(s)*™. Physicians should claim only the credit commensurate with the extent of their participation in the activity.

Method of Physician Participation in Learning Process: The content can be read from the Journal of Investigative Dermatology website: http://www.jidonline.org/current. Tests for CME credits may only be submitted online at https://beaumont.cloud-cme.com/RTMS-Jul19 – click 'CME on Demand' and locate the article to complete the test. Fax or other copies will not be accepted. To receive credits, learners must review the CME accreditation information; view the entire article, complete the post-test with a minimum performance level of 60%; and complete the online evaluation form in order to claim CME credit. The CME credit code for this activity is: 21310. For questions about CME credit email cme@beaumont.edu.

INTRODUCTION

Observational epidemiological studies have uncovered relationships between disease and various explanatory factors known as *exposures* (Table 1). Notable examples in dermatology include the association of psoriasis with smoking (Armstrong et al., 2014) and, more recently, the association of atopic dermatitis with cardiovascular traits (Standl et al., 2017). However, traditional observational studies are prone to biases such as confounding, where the observed association may be due to the exposure being related to other lifestyle or socioeconomic factors that have a casual influence on disease. Furthermore, the observed associations may be due to reverse causation, where disease is actually influencing the assumed exposure (Lawlor et al., 2008); for example, having psoriasis could influence an individual's propensity to smoke. Mendelian randomization (MR) presents a method for evaluating causality in an observational study setting. We aim to provide an overview of the principle of MR and the statistical methods used.

Medical Research Council Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

Correspondence: Lavinia Paternoster, Medical Research Council Integrative Epidemiology Unit, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK. E-mail: L.Paternoster@bristol.ac.uk

Abbreviations: BMI, body mass index; GWAS, genome-wide association study; MR, Mendelian randomization

SUMMARY POINTS

- Mendelian randomization (MR) is a statistical method for investigating causality between exposure and outcome variables in observational epidemiology.
- Unlike traditional observational studies, MR uses genetic variants as instruments (or proxies) for the exposure, hence avoiding confounding and reverse causation.
- Application of such methods in the field of dermatology is a promising area of research.
- Future directions and developments will allow MR to be a valuable tool for investigating causal pathways for disease, as well as providing insight into therapeutic interventions.

THE PRINCIPLE OF MR

MR is a form of instrumental variable analysis whereby genetic variants are used as instruments (or proxies) for an exposure of interest (Table 1). Because genetic variants are randomly segregated at conception and cannot be influenced by confounding factors or the outcome itself, they can be used to estimate the causal effect of the exposure upon an outcome (Lawlor et al., 2008) (Figure 1).

Performing MR requires two pieces of information: (i) the effect of the genetic instrument on the exposure (β_{XZ}) and (ii) the effect of the genetic instrument on the outcome (β_{YZ}). These can then be used to estimate the causal effect of the exposure on the outcome (causal β_{YX}) with the following ratio (Wald, 1940): causal $\beta_{YX} = \frac{\beta_{YZ}}{\beta_{XZ}}$.

For a genetic variant to qualify as an instrumental variable, three core assumptions must be satisfied: the variants (i) must be truly associated with the exposure of interest, (ii) must not be associated with confounders of the exposure-outcome relationship, and (iii) must affect only the outcome via the exposure and not through an alternative pathway (Zheng et al., 2017). The use of genetic variants in an MR framework can be compared with a randomized controlled trial, where genotypes are used to randomize individuals to different subgroups (Lawlor et al., 2008). The effect of the genetic instrument on the outcome (β_{YZ}) is analogous to an intention-to-treat effect from an association between randomization and an outcome in a randomized controlled trial (Burgess and Thompson, 2015).

Because MR requires estimates of the associations between genetic variants and the exposure and genetic variants and the outcome, the rise of genome-wide association studies (GWASs) (Tsoi et al., 2018) provides a wealthy resource of genetic instruments for MR. Published summary GWAS data can be obtained from various sources such as the GWAS catalogue (www.ebi.ac.uk/gwas/) and MR-base (www.mrbase.org) or directly from the authors of the GWAS (Figure 2). Commonly, independent single-nucleotide polymorphisms (SNPs) that have been reported to be associated with an exposure on a genome-wide significance level (P-value $< 5 \times 10^{-8}$) are used as genetic instruments for the exposure (Zheng et al., 2017), but MR analyses can be conducted by using just a single genetic variant or even using all variants in the genome (appropriately weighted by their effect

Table 1. Glossary

Term	Description
Confounder	A variable that is a common cause of both the exposure and the outcome.
Exposure	An explanatory variable used to explain or predict an outcome variable, such as a trait or disease.
F-statistic	Obtained from the regression of a response variable on a predictor variable, for example, the regression of the exposure of interest on an instrumental variable (IV). This can be used as a measure of the strength of association between an IV and the exposure, thereby giving an indication of the strength of the instrument. The further away the *F*-statistic is from 1, the stronger the instrument. The *F*-statistic also depends on the size of the sample.
GWAS	Genome-wide association study. Involves analyzing genetic variants across the genome, such as single-nucleotide polymorphisms for association with a disease or trait of interest.
Instrumental variable (IV)	A variable that is associated with an exposure of interest but not the outcome. In MR studies, genetic variants are used as IVs. A valid IV must also be independent of confounders of the exposure-outcome association and must affect only the outcome via the exposure.
Mendelian randomization	A method for assessing the causal effect of an exposure on an outcome by using genetic variants as instruments or proxies for the exposure variable.
MR-base	A centralized database of summary GWAS data and an analytical platform to perform Mendelian randomization and sensitivity analyses.
PheWAS	Phenome-wide association study. Involves analyzing the association between genetic variants and multiple phenotypic variables (on a phenome-wide scale) rather than a single phenotype.
Pleiotropy	Occurs when a genetic instrument is independently associated with multiple risk factors for the outcome, in addition to the exposure of interest. This results in the third IV assumption being violated, which assumes that the genetic instrument affects only the outcome via the exposure.
Reverse causality	Where an association is due to the assumed outcome variable influencing the exposure variable rather than the exposure influencing the outcome.
Sensitivity analysis	Performed to assess the robustness of the main analysis or the validity of the main results.

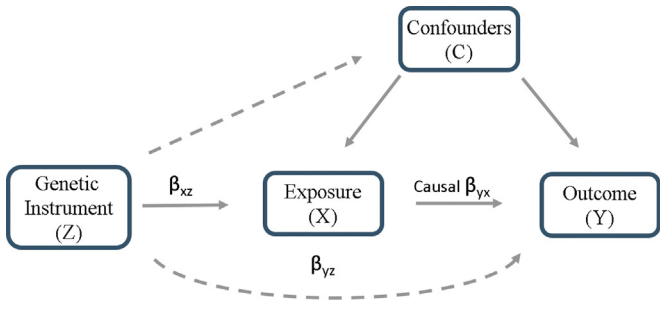

Figure 1. Illustrative diagram of standard Mendelian randomization (MR) analysis. A valid genetic instrument (Z) must be truly associated with the exposure (X), must not be associated with confounders (C), and should have an effect only on the outcome (Y) via the exposure. Dashed arrows represent violations of these MR assumptions.

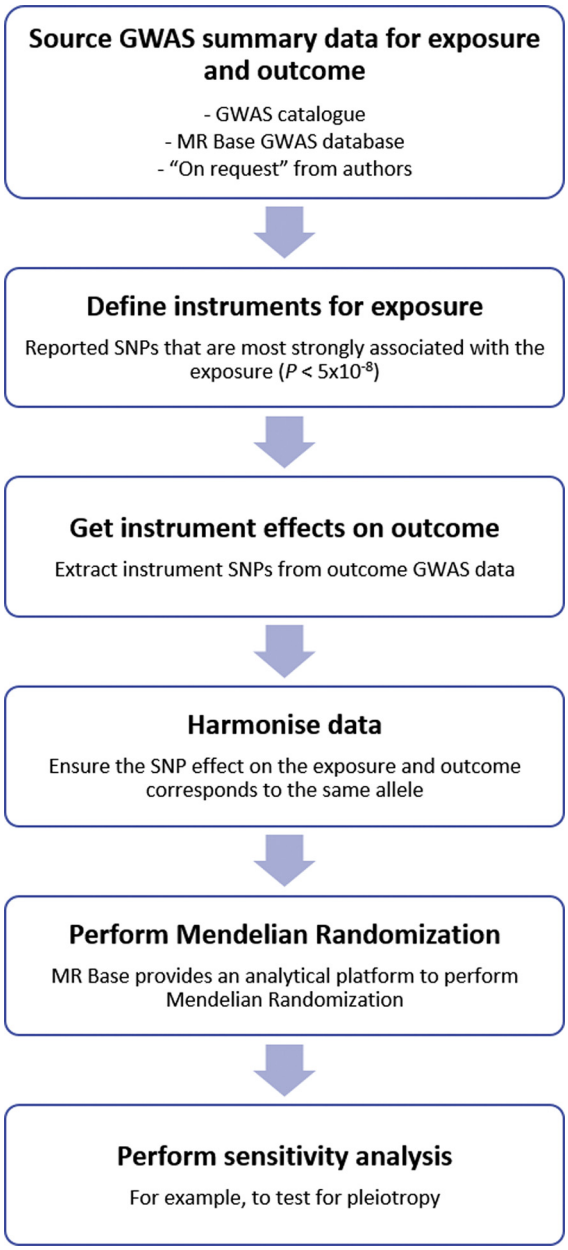

Figure 2. Workflow for performing two-sample MR analysis. Summary GWAS data provide a wealthy resource of genetic instruments to perform MR. Various MR methods and sensitivity analyses can be performed with analytical platforms such as MR-base. Adapted from Hemani et al. (2018). MR, Mendelian randomization; SNP, single-nucleotide polymorphism.

on the exposure). Published MR studies in dermatology include those investigating causal relationships between fatty acids and melanoma (Liyanage et al., 2018), vitamin D levels and AD risk (Manousaki et al., 2017) as well as skin aging (Noordam et al., 2017), and, most recently, body mass index (BMI) and psoriasis risk (Budu-Aggrey et al., 2019), which will be referred to throughout this review.

MR APPROACHES AND STATISTICAL METHODS

MR study designs

A basic MR study design involves obtaining all information required from the same set of individuals, meaning that the genetic, exposure, and outcome data are all available from the same study. This is known as *one-sample MR* (Table 2). Large population-based studies such as the UK Biobank provide ideal data sets for such analyses to be carried out. However, it may not always be possible to gather exposure and outcome measures from the same data set. *Two-sample MR* is therefore more commonly adopted, whereby the effect of genetic variants on the exposure is obtained from one sample, and the effect of genetic variants on the outcome is obtained from another (Table 2). This approach has been greatly facilitated by the increasing availability of summary GWAS data, as well as analytical platforms to perform two-sample MR, such as MR-base. The steps for a two-sample MR are shown in Figure 2 (Hemani et al., 2018).

We recently investigated causality between BMI and psoriasis using both one-sample MR with individual-level data from the UK Biobank and Nord-Trøndelag Health Study (i.e., HUNT) and two-sample MR with published summary GWAS data. Consistent results were obtained from both analyses. The combined causal estimate suggested a 9% increase in the risk of psoriasis per 1 unit increase in BMI (Budu-Aggrey et al., 2019) (Figure 3). This finding supports previous reports of weight loss improving the prognosis of psoriasis (Maglio et al., 2017) and could suggest weight control as an intervention to prevent or treat psoriasis.

A bidirectional MR approach can also be adopted that investigates causal effects in both directions (Table 2). This requires suitable genetic instruments to be available for both traits. Such analysis can help uncover the direction of causality that explains the observational association. For example, when considering the relationship between BMI and psoriasis, we performed bidirectional MR and found evidence that the observational relationship is largely due to the causal effect of higher BMI on psoriasis risk rather than a causal effect of psoriasis influencing BMI (Budu-Aggrey et al., 2019).

MR statistical methods

The simplest method to perform MR involves dividing the effect of the genetic instrument on the outcome by the effect of the genetic instrument on the exposure. This is commonly termed the *ratio of coefficients method* or the *Wald ratio method* (as described earlier) and can be performed with either summarized or individual-level data (Burgess et al., 2017). Two-stage methods can also be applied, such as two-stage least squares, as used in the BMI and psoriasis article by Budu-Aggrey et al. (2019) (Table 2). This method involves regressing the exposure on the genetic instruments and then regressing the outcome on the genetically predicted values from the first regression, which allows for the true standard error to be estimated. Additional MR methods have been previously discussed elsewhere (Burgess et al., 2017).

Combining multiple variants

Where multiple genetic instruments are available for an exposure, these can be combined into a genetic risk score and used as a single instrument to perform MR (Zheng et al., 2017). Alternatively, an inverse-variance–weighted approach can be applied, whereby the ratio estimate from each independent genetic variant is combined by using a fixed-effect meta-analysis model, where each variant is assumed to provide independent information, and the contribution of each variant is the inverse of the variance of its effect on the outcome (Burgess et al., 2013) (Table 2).

Sensitivity methods

One major potential problem with MR occurs when the genetic instrument affects the outcome through an alternative pathway that is distinct from the exposure of interest (termed *pleiotropy*) (Table 1), which violates the third assumption (as outlined earlier). Various

Table 2. Methods and approaches for MR analysis

Category	Description
MR study design	
One-sample MR	Performed with genetic instruments, exposure and outcome data that have been measured in the same sample population.
Two-sample MR	The effect of the genetic instruments on the exposure and the effect of the genetic instruments on the outcome are obtained from a non-overlapping sample populations.
Bidirectional MR	The causal relationship between two traits is investigated in both directions. This approach can be applied to one-sample or two-sample MR.
Statistical methods	
Wald ratio method	Performed with a single genetic instrument (or genetic risk score) by dividing the coefficient of the outcome-instrument association by the coefficient of the exposure-instrument association.
Two-stage least squares (2SLS) regression	Involves two regression stages where the exposure is regressed on the genetic instruments. The outcome is then regressed on the genetically predicted exposure values from the first-stage regression.
Combining multiple variants	
Inverse-variance weighted (IVW) estimator	Combination of ratio estimates from individual variants in a fixed-effect meta-analysis. The contribution of each instrument is the inverse of the variance of its effect on the outcome.
Genetic risk score (GRS)	Multiple genetic instruments for an exposure are combined into a genetic risk score. This can then be used as a single instrument to perform MR.
Sensitivity analysis	
MR-Egger regression	Sensitivity analysis to perform MR with multiple instruments. This can be used to detect pleiotropy and provide a causal estimate that is robust to pleiotropy.
Weighted-median estimator	Sensitivity analysis to perform MR with multiple instruments. Will provide consistent causal estimates when at least 50% of the information in the analysis comes from valid genetic instruments.
Mode-based estimator	An MR sensitivity analysis that will provide a robust causal estimate in the presence of pleiotropy, if the most common pleiotropy value is zero across the genetic instruments.
Latent causal variable analysis	Distinguishes between genetic correlation and causation by mediating the genetic correlation between two traits with a latent causal variable that itself has a causal effect on each trait.

Abbreviation: Mendelian randomization.

sensitivity methods have been developed to detect and address pleiotropy, including MR-Egger regression, weighted-median analysis, the mode-based estimate, and the latent causal variable method (Table 2). These methods have different assumptions, but they aim to estimate the true causal effect in the presence of modest levels of pleiotropy (O'Connor and Price, 2018; Zheng et al., 2017).

Challenges and limitations of MR studies

Although MR has proven to be a useful tool for estimating causality, there are instances where MR may be limited or the instrumental variable assumptions may be violated. In some cases, there may be only weak genetic instruments available for the exposure of interest. Genetic instruments that explain very little of the variance in exposure can result in weak instrument bias, where the causal estimates can be biased toward the null in a two-sample MR setting and toward the observational estimate in a one-sample MR setting (Zheng et al., 2017). This highlights the need for GWASs to uncover associated variants and strong, reliable instruments to perform MR. The *F*-statistic from the regression of the exposure on the genetic instrument indicates the strength of the instrument (Table 1). It is recommended that genetic variants with an *F*-statistic greater than 10 be used (Burgess et al., 2013; Lawlor et al., 2008). Because the *F*-statistic is dependent on sample size, weak instrument bias can also be addressed by using larger sample sizes (Burgess and Thompson, 2015). Additionally, combining individual variants into a genetic risk score increases the instrument strength. The instrument for BMI in our psoriasis analysis had an *F*-statistic of 7,091, indicating a strong instrument for BMI (Budu-Aggrey et al., 2019).

Although it is assumed that a genetic instrument is independent of confounders, this cannot be tested for all potential confounders. However, it is sensible to test for association between the genetic instrument and any available measured potential confounders.

Applications and future directions for MR

MR is commonly performed to investigate the causality of established observational associations. However, a "hypothesis-free" approach can also be adopted to uncover novel causal relationships. This involves performing MR on a phenome-wide scale, known as *MR-pheWAS*, where the effect of a single exposure on multiple outcomes is evaluated. This has been shown by Haycock et al. (2017), who found that telomere length increased the risk of several cancers and reduced the risk of nonneoplastic diseases.

MR can also be applied to investigate the causal role of molecular traits, such as gene expression, methylation, and protein biomarkers, on disease. In doing so, genetic variants associated with expression (expression quantitative trait loci), methylation (methylation

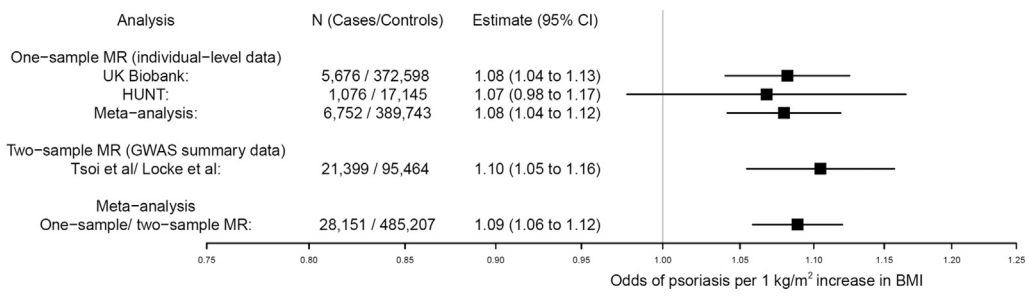

Figure 3. One-sample and two-sample MR estimates give evidence of increased psoriasis risk with 1 unit increase in BMI (kg/m^2). One sample MR has been performed with individual- level data. Two-sample MR has been performed with summary GWAS data. Adapted from Budu-Aggrey et al. (2019). BMI, body mass; CI, confidence interval; HUNT, Nord-Trøndelag Health Study; GWAS, genome-wide association study; MR, Mendelian randomization.

MULTIPLE CHOICE QUESTIONS

1. Which of the following is a limitation of observational studies that can be addressed with MR?
   A. Publication bias
   B. Selection bias
   C. Confounding
   D. Inadequate sample size
2. Which of the following is NOT an assumption for a valid MR instrument?
   A. The instrument must be truly associated with the exposure and the outcome.
   B. The instrument must be truly associated with the exposure.
   C. The instrument must not be associated with confounders of the exposure-outcome relationship.
   D. The instrument must affect only the outcome via the exposure.
3. Which of the following can be used to uncover the direction of a causal relationship?
   A. Two-sample MR
   B. Observational analysis
   C. One-sample MR
   D. Bidirectional MR
4. Which of the following can be used to address pleiotropy in MR?
   A. Wald ratio method
   B. MR-Egger regression
   C. Inverse-variance weighted estimator
   D. Two-stage least squares
5. Which of the following statements is FALSE?
   A. MR can be performed in a hypothesis-free manner.
   B. MR estimates represent the effect of long-term exposures.
   C. Pleiotropic genetic instruments cannot be included in MR analyses.
   D. MR can be used to investigate the causal role of molecular phenotypes.

quantitative trait loci), or plasma protein levels (protein quantitative trait loci) are used as genetic instruments for the exposure and can provide insight into the causal pathways that underlie disease. This has been shown for AD, where MR analysis with protein quantitative trait loci gave evidence that IL1RL2 and IL18R1 are causal proteins for AD risk (Sun et al., 2018).

Many MR studies are performed in cohorts with limited ethnic variation. As shown by Ogawa et al. (2018), transethnic MR studies can make the causal estimate more robust to confounding by population stratification and more generalizable to broader ethnic backgrounds (Ogawa et al., 2018).

We also expect that MR methods will begin to be applied to outcomes of disease progression (as opposed to onset), to enable them to be more informative for the treatment of patients (Paternoster et al., 2017). Such studies have begun to emerge in other disease areas, such as Parkinson disease (Simon et al., 2014), and could potentially uncover novel therapeutic targets or drug repurposing opportunities in dermatology.

CONCLUSION

MR has proven to be a robust statistical method to infer causal relationships in observational studies. In this review, we have presented strategies for performing MR, as well as the limitations and promising extensions of this method. As large GWAS summary statistics and open-access data sets become increasingly available and additional methods continue to be developed, the potential for MR analysis to produce further evidence of causality for dermatological traits will increase. This, in turn, will aid in the understanding of underlying mechanisms of disease and inform disease prevention and treatment.

CONFLICT OF INTEREST

LP has received personal fees from Merck for Scientific Input Engagement related to MR methodology.

ACKNOWLEDGMENTS

AB-A is funded by a grant awarded by the British Skin Foundation (8010 Innovative Project), awarded to LP. AB-A and LP work in a research unit funded by the UK Medical Research Council (MC_UU_00011/1).

SUPPLEMENTARY MATERIAL

Supplementary material is linked to this paper. Teaching slides are available as supplementary material.

REFERENCES

Armstrong AW, Harskamp CT, Dhillon JS, Armstrong EJ, Armstrong AW. Psoriasis and smoking: a systematic review and meta-analysis funding sources. Br J Dermatol 2014;170:304–14.

Budu-Aggrey A, Brumpton B, Tyrrell J, Watkins S, Modalsli EH, Celis-Morales C, et al. Evidence of a causal relationship between body mass index and psoriasis: a Mendelian randomization study. PLoS Med 2019;16(1):e1002739.

Burgess S, Butterworth A, Thompson SG. Mendelian randomization analysis with multiple genetic variants using summarized data. Genet Epidemiol 2013;37:658–65.

Burgess S, Small DS, Thompson SG. A review of instrumental variable estimators for Mendelian randomization. Stat Methods Med Res 2017;26:2333–55.

Burgess S, Thompson SG. Mendelian randomization: methods for using genetic variants in causal estimation, https://books.google.co.uk/books?id=WYSbBgAAQBAJ&printsec=frontcover&dq=Mendelian+Randomization+Methods+for+using+Genetic+variants+in+Casual+Estimation+CRC+Press&hl=en&sa=X&ved=0ahUKEwjbwuT9k5reAhUqBMAKHYvmBlMQ6AEIMjAB#v=onepage&q&f=false; 2015 (accessed 22 Oct 2018).

Haycock PC, Burgess S, Nounu A, Zheng J, Okoli GN, Bowden J, et al. Association between telomere length and risk of cancer and non-neoplastic diseases. JAMA Oncol 2017;3:636.

Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-base platform supports systematic causal inference across the human phenome. eLife 2018;7:e34408.

Lawlor DA, Harbord RM, Sterne JAC, Timpson N, Davey Smith G. Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med 2008;27:1133–63.

Liyanage UE, Law MH, Ong JS, Cust AE, Mann GJ, Ward SV, et al. Polyunsaturated fatty acids and risk of melanoma: a Mendelian randomisation analysis. Int J Cancer 2018;143:508–14.

Maglio C, Peltonen M, Rudin A, Carlsson LMS. Bariatric surgery and the incidence of psoriasis and psoriatic arthritis in the Swedish obese subjects study. Obesity (Silver Spring) 2017;25:2068–73.

Manousaki D, Paternoster L, Standl M, Moffatt MF, Farrall M, Bouzigon E, et al. Vitamin D levels and susceptibility to asthma, elevated immunoglobulin E levels, and atopic dermatitis: a Mendelian randomization study. PLoS Med 2017;14(5):e1002294.

Noordam R, Hamer MA, Pardo LM, van der Nat T, Kiefte-de Jong JC, Kayser M, et al. No causal association between 25-hydroxyvitamin D and features of skin aging: evidence from a bidirectional Mendelian randomization study. J Invest Dermatol 2017;137:2291–7.

O'Connor LJ, Price AL. Distinguishing genetic correlation from causation across 52 diseases and complex traits. Nat Genet 2018;50:1728–34.

Ogawa K, Stuart PE, Tsoi LC, Suzuki K, Nair RP, Mochizuki H, et al. A transethnic Mendelian randomization study identifies causality of obesity on risk of psoriasis [e-pub ahead print]. J Invest Dermatol 2019;139:1397–400.

Paternoster L, Tilling K, Davey Smith G. Genetic epidemiology and Mendelian randomization for informing disease therapeutics: conceptual and methodological challenges. PLoS Genet 2017;13(10):e1006944.

Simon KC, Eberly S, Gao X, Oakes D, Tanner CM, Shoulson I, et al. Mendelian randomization of serum urate and Parkinson disease progression. Ann Neurol 2014;76:862–8.

Standl M, Tesch F, Baurecht H, Rodríguez E, Müller-Nurasyid M, Gieger C, et al. Association of atopic dermatitis with cardiovascular risk factors and diseases. J Invest Dermatol 2017;137:1074–81.

Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, et al. Genomic atlas of the human plasma proteome. Nature 2018;558:73–9.

Tsoi LC, Patrick MT, Elder JT. Research techniques made simple: using genome-wide association studies to understand complex cutaneous disorders. J Invest Dermatol 2018;138(3):e23–9.

Wald A. The fitting of straight lines if both variables are subject to error. Ann Math Stat 1940;11:284–300.

Zheng J, Baird D, Borges M-C, Bowden J, Hemani G, Haycock P, et al. Recent developments in mendelian randomization studies. Curr Epidemiol Rep 2017;4:330–45.

DETAILED ANSWERS

1. **Which of the following are limitations of observational studies that can be addressed with MR?**

Correct answer: C. Confounding

Traditional observational studies are limited by confounding, reverse causation, and measurement error. MR can be used to evaluate causality in observational studies while avoiding these limitations.

2. **Which of the following is NOT an assumption for a valid MR instrument?**

Correct answer: A. The instrument must be truly associated with the exposure and the outcome.

A valid MR instrument must satisfy three core assumptions. The instrument must be truly associated with the exposure, must not be associated with confounders of the exposure-outcome relationship, and must affect only the outcome via the exposure and not through an alternative pathway.

3. **Which of the following can be used to uncover the direction of a causal relationship?**

Correct answer: D. Bidirectional MR

Bidirectional MR involves investigating the causal effect of an exposure on an outcome, as well as evaluating the effect in the reverse direction of the outcome on the exposure. In doing so, the direction of the causal relationship can be determined.

4. **Which of the following can be used to address pleiotropy in MR?**

Correct answer: B. MR-Egger regression

MR-Egger regression can be performed to detect the presence of pleiotropy and also to obtain a causal estimate that is robust to pleiotropy.

5. **Which of the following statements is FALSE?**

Correct answer: C. Pleiotropic genetic instruments cannot be included in MR analyses.

MR can be performed on a phenome-wide scale to investigate the causal effect of a single exposure on multiple outcomes with MR-pheWAS. MR estimates also represent the effect of long-term exposures rather than short-term interventions. In addition, MR can be extended to investigate the causal effect of molecular traits on disease, where expression quantitative trait loci, methylation quantitative trait loci, or protein quantitative trait loci are used as genetic instruments. Genetic instruments that are pleiotropic are not valid for MR analysis; however, MR methods have been developed to address pleiotropy which allows for both unpleiotropic and pleiotropic variants to be included. These include MR-Egger regression, weighted-median analysis, and the mode-based estimate.