Accepted Manuscript

Variation in Interleukin 6 Receptor Gene Associates with Risk of Crohn's Disease and Ulcerative Colitis

C.A. Parisinos, S. Serghiou, M. Katsoulis, M.J. George, R.S. Patel, H. Hemingway, A.D. Hingorani

PII: S0016-5085(18)34542-6
DOI: 10.1053/j.gastro.2018.05.022
Reference: YGAST 61888

To appear in: *Gastroenterology*
Accepted Date: 6 May 2018

Please cite this article as: Parisinos C, Serghiou S, Katsoulis M, George M, Patel R, Hemingway H, Hingorani A, Variation in Interleukin 6 Receptor Gene Associates with Risk of Crohn's Disease and Ulcerative Colitis, *Gastroenterology* (2018), doi: 10.1053/j.gastro.2018.05.022.

Variation in Interleukin 6 Receptor Gene Associates with Risk of Crohn's Disease and Ulcerative Colitis

Parisinos CA[1*], Serghiou S[2], Katsoulis M[1], George MJ[3], Patel RS[4], Hemingway H[1], Hingorani AD[4].

* Corresponding author contacts: email: c.parisinos@ucl.ac.uk tel: +447899786998

1) Institute of Health Informatics Research, Faculty of Population Health Sciences
University College London
222 Euston Road
United Kingdom
NW1 2DA

2) Health Research and Policy, Epidemiology
Stanford University
450 Serra Mall
Stanford
United States of America
CA 94305 – 2004

3) Centre for Clinical Pharmacology, Division of Medicine
University College London
Gower Street
United Kingdom
WC1E 6BT

4) Institute of Cardiovascular Science, Faculty of Population Health Sciences
University College London
222 Euston Road
United Kingdom
NW1 2DA

CAP: Study concept and design, data acquisition, analysis and interpretation of data

SS: Statistical analysis and interpretation of data

MK: Statistical analysis

GMJ: Discussion

RSP: Critical review of manuscript, study supervision

HH: Critical review of manuscript, study supervision

ADH: Critical review of manuscript, interpretation of data, study supervision

Conflicts of interest for all authors: None

Abstract for Gastroenterology brief report

Abstract: Interleukin 6 (IL6) is an inflammatory cytokine; signaling via its receptor (IL6R) is believed to contribute to development of inflammatory bowel diseases (IBD). The single nucleotide polymorphism rs2228145 in *IL6R* associates with increased levels of soluble IL6R (s-IL6R), as well as reduced IL6R signaling and risk of inflammatory disorders; its effects are similar to those of a therapeutic monoclonal antibody that blocks IL6R signaling. We used the effect of rs2228145 on s-IL6R level as an indirect marker to investigate whether reduced IL6R signaling associates with risk of ulcerative colitis (UC) or Crohn's disease (CD). In a genome-wide meta-analysis of 20,550 patients with CD, 17647 patients with UC, and more than 40,000 individuals without IBD (controls), we found that rs2228145 (scaled to a 2-fold increase in s-IL6R) was associated with reduced risk of CD (odds ratio, 0.876; 95% CI, 0.822–0.933; *P*=.00003) or UC (odds ratio, 0.932; 95% CI, 0.875–0.996; *P*=.036). These findings indicate that therapeutics designed to block IL6R signaling might be effective in treatment of IBD.

KEY WORDS: SNP, genetics, Mendelian randomization, drug target validation

Brief report

Crohn's disease (CD) and ulcerative colitis (UC) are chronic inflammatory diseases affecting the gastrointestinal tract. Effective novel therapeutics remain a priority for these disabling conditions.

Interleukin-6 (IL6) is a pro-inflammatory cytokine that can exert its biological effect via two mechanisms: classical signaling through its membrane-bound IL6 receptor (IL6R), and trans-signaling, by binding to a soluble form of IL6R (s-IL6R) and subsequently to the membrane-bound transducer glycoprotein 130 (gp130)[1]. Whether IL6R is an attractive drug target for the management of both UC and CD is unknown.

Mendelian randomization (MR) can provide information on causality between an exposure and disease and has been successfully adopted for drug target validation[2]. This method relies on a simple principle; if a modifiable exposure (e.g. a biomarker, a complex trait, or an environmental risk factor such as alcohol intake) is causal for a disease, then the genetic variants associated with (or that mirror the biological effects of) that exposure will also be associated with disease risk (Figure 1). The causal inference is possible due to the fundamental nature of the genome; variants are randomly allocated at meiosis, balancing confounders, and reverse causation, another important source of bias in observational studies, is not possible since the sequence of the germline is generally not modifiable by disease. MR can therefore be considered in many ways analogous to a randomized controlled trial (RCT) (Figure 2).

The *IL6R* SNP rs2228145 leads to increased proteolysis of its product and a reduction in classical signaling and has similar directional effects to an existing anti-IL6R monoclonal antibody, tocilizumab (licensed for the treatment of rheumatoid arthritis); these include accumulation of circulating IL6 and s-IL6R levels, most likely due to reduced clearance of IL6 via its receptor in the liver and termination of negative feedback mechanisms, as well as a reduction in downstream inflammatory biomarkers such as c-reactive protein (CRP), platelets and fibrinogen[3]. Despite important differences between tocilizumab and rs2228145 (the former inhibits both IL6R and s-IL6R, whereas the latter induces proteolysis specific to IL6R), the above similarities make rs2228145 an attractive genetic instrument for drug target validation of IL6R inhibition (Figure 2)[4]. Previous MR studies have demonstrated protective associations between rs2228145 and inflammatory conditions including coronary heart disease and rheumatoid arthritis[3].

We aimed to evaluate and quantify the effect of IL6R signaling on the risk of developing UC and CD using a 2-sample MR design (see Supplementary Material). We used the effect of rs2228145 on circulating levels of s-IL6R (and IL6, sensitivity analysis, see Supplementary Material) as an indirect marker for our exposure, reduced membrane bound IL6R and classical IL6R signaling, as described elsewhere (Figure 1)[2,5].

The SNP (rs2228145) - biomarker (s-IL6R, IL6) associations estimated in 1650 individuals were used as genetic instruments; rs2228145 elevates serum s-IL6R levels by 34% (IL6 by 15%)[4]. SNP - inflammatory bowel disease (IBD) associations

were extracted from the largest IBD Genetics Consortium meta-analysis to date[6]. The ratio MR method was used to obtain individual exposure estimates by dividing the SNP - outcome by the SNP - biomarker effect estimates.

In a combined total of 20550 patients with CD (41642 controls) and 17647 patients with UC (47179 controls), rs2228145 was associated with decreased odds of CD (OR 0.948, 95% CI, 0.925-0.972, p = .00003) and UC (OR 0.973, 95% CI, 0.948-0.998, p = .038) per effect allele. When applying the ratio MR method to quantify the association between the indirect marker (s-IL6R) and the outcome, a two-fold genetic elevation of s-IL6R was associated with decreased odds of CD (OR 0.876, 95% CI, 0.822-0.933, p = .00003) and UC (OR 0.932, 95% CI, 0.875-0.996, p = .036) (see Supplementary material). As a point of reference, tocilizumab increases s-IL6R levels by approximately ten-fold[2].

About 90% of drugs that enter clinical development fail; genetic evidence for a therapeutic target doubles the clinical success rate of such drugs[7]. Our findings are consistent with 2 RCTs of antibodies targeting IL6 signaling for the treatment of CD; a small Phase I study of tocilizumab suggested higher clinical response rates in CD than the placebo group[8]. Additionally, a recent RCT (ANDANTE) of an anti-IL6 antibody (PF-04236921) yielded higher clinical response and remission rates in patients with refractory CD versus placebo[9]. Rare cases of gastrointestinal perforation however in patients treated with both of the above antibodies remain a concern, since IL6 signaling may also contribute to epithelial repair of the intestinal mucosa[10]. Avoiding use in at risk individuals including patients with diverticulitis and active fistulae should still be advocated.

In the MR paradigm, genetic associations are generally free from confounding; previous studies have demonstrated no association between rs2228145 and age, birth weight and education[11]. Another concern for potential bias occurs when variants influence biomarkers on distinct causal pathways (horizontal pleiotropy). Genetic instruments however designed to model a protein (compared to more distal exposures such as complex traits) are protected from such pleiotropy[12]. *IL6R* affects multiple downstream biomarkers on the same causal pathway, in a similar manner to pharmacological blockade, a (vertical) form of pleiotropy that does not lead to bias (Figure 2).

Interpretation of our results requires caution. A reduction in the risk of developing IBD does not necessarily translate to a reduction in disease progression, since genetic contributions to prognosis may differ to those of susceptibility[13]. An incorrect conclusion would be that “increased s-IL6R reduces the risk of developing IBD”; s-IL6R was used as an indirect marker for reduced membrane-bound IL6R and subsequent classical signaling. Results should be interpreted as “a genetic reduction of IL6R and its classical signaling, sufficient to double s-IL6R reduces the risk of developing IBD”[2].

Our study suggests that a reduction in IL6R and subsequent classical signaling reduces IBD risk. One possible consequence is an increase in s-IL6R/IL6 complexes and subsequent trans-signaling via membrane-bound gp130. However, even if this occurs, it does not fully compensate for the impaired classical signaling, as

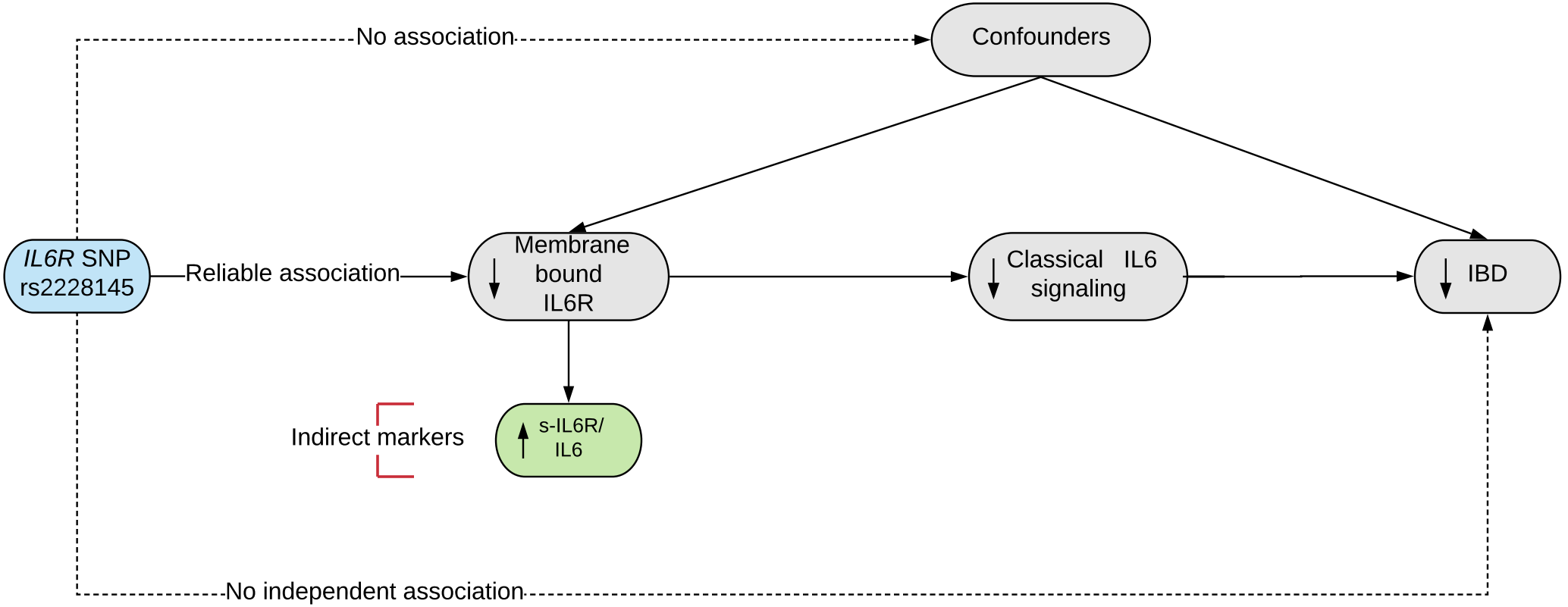
evidenced by the net reduction in downstream biomarkers such as CRP, platelets and fibrinogen. Alternatively, it has been proposed that the increase in s-IL6R may actually further reduce trans-signaling by enhancing the buffering potential of the abundant soluble glycoprotein 130 (s-gp130)[14]. Further work is required to determine the effect of rs2228145 on IL6 trans-signaling and whether trans-signaling is in itself a potential therapeutic target in IBD.

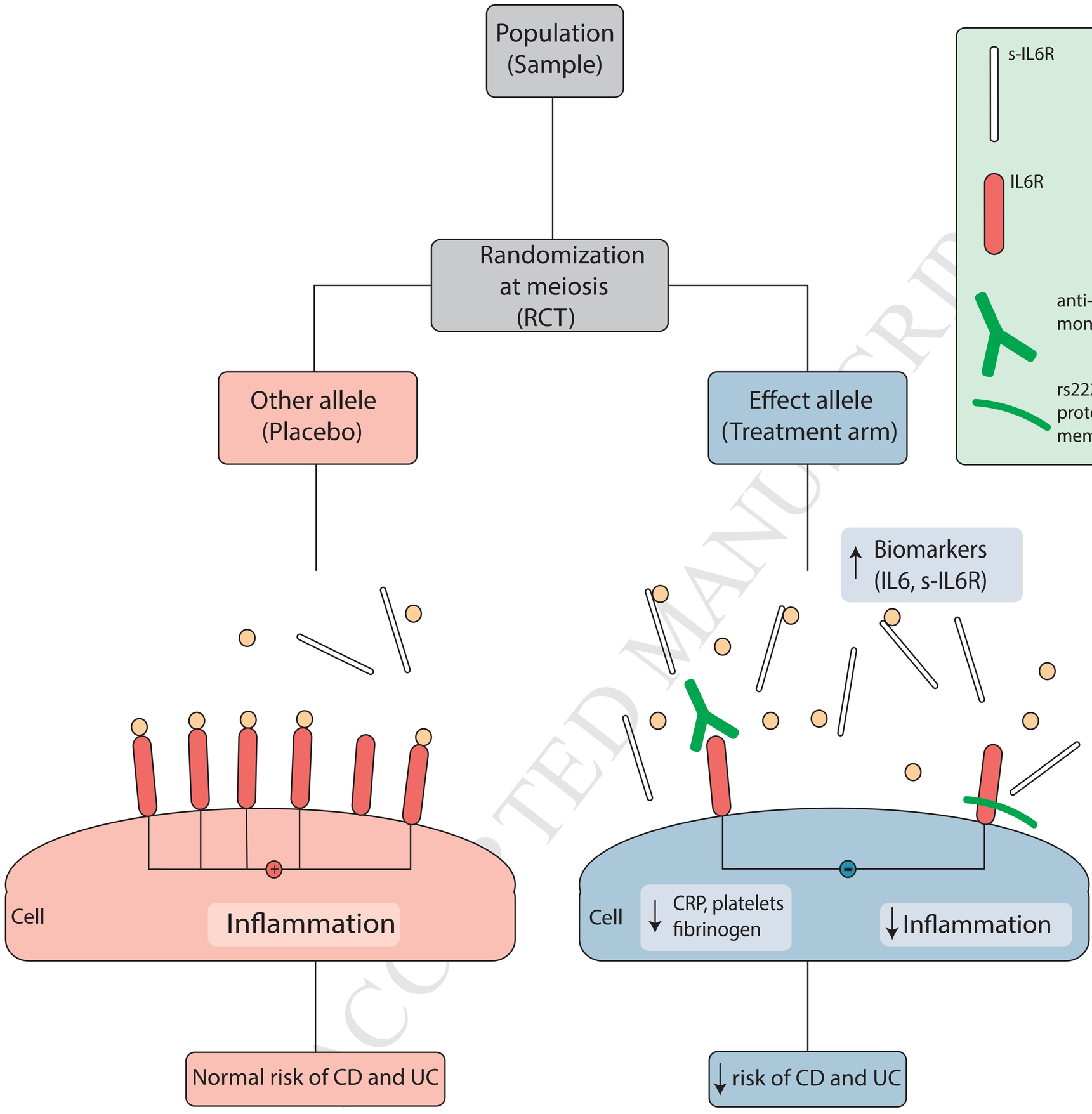
On the basis of genetic evidence in humans, IL6R signaling seems to have a causal role in the development of both CD and UC; Suitably powered RCTs of new and existing therapeutics targeting this complex pathway are required in both conditions, alongside ongoing focus on the pathophysiology underlying rare complications such as gastrointestinal perforations.

Figure 1. The MR model using a variant that disrupts normal function of the exposure (IL6R classical signaling), indirectly measured through increased levels of s-IL6R and IL6. The three principles of MR analysis are: a genetic instrument is robustly associated with the exposure (assumption 1, continuous arrow) but not with confounders (assumption 2, dotted arrow). The genetic variant is associated with the disease only through its effects on the exposure (assumption 3, dotted arrow).

Figure 2. Schematic demonstrating how MR can be considered a natural analogue of the classical randomized controlled trial (corresponding RCT steps in brackets). *IL6R* SNP rs2228145 has similar directional biomarker effects with tocilizumab.

1. **Lissilaa R,** et al. J Immunol 2010;185(9):5512-5521. doi:10.4049/jimmunol.1002015.
2. **Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6R MR) Consortium,** et al. Lancet 2012;379(9822):1214-1224.
3. **Ferreira RC,** et al. PLoS Genet. 2013;9(4):e1003444. doi:10.1371/journal.pgen.1003444.
4. **IL6R Genetics Consortium Emerging Risk Factors Collaboration,** Lancet 2012;379(9822):1205-1213.
5. **Hartwig FP,** et al. JAMA Psychiatry 2017. doi:10.1001/jamapsychiatry.2017.3191.
6. **Liu JZ,** et al. Nat Genet 2015;47(9):979-986.
7. **Nelson MR,** et al. Nat Genet 2015;47(8):856-860.
8. **Ito H,** et al. Gastroenterology 2004;126(4):989-996.
9. **Danese S,** et al. Gut 2017:gutjnl-2017-314562.
10. **Aden K,** et al. Oncogenesis. 2016;5(11):e270-e270.
11. **Khandaker GM**, et al. Brain Behav Immun. December 2017.
12. **Swerdlow DI,** et al. Int J Epidemiol 2016;45(5):1600-1616.
13. **Lee JC,** et al. Nat Genet 2017;49(2):262-268.
14. **Scheller J,** et al. The interleukin 6 pathway and atherosclerosis. Lancet 2012;380(9839):338.


No association
Confounders
IL6R SNP
rs2228145
Reliable association
Membrane
bound
IL6R
Classical IL6
signaling
IBD
Indirect markers
s-IL6R/
IL6
No independent association


Population
(Sample)
Randomization
at meiosis
(RCT)
Other allele
(Placebo)
Effect allele
(Treatment arm)
s-IL6R
IL6
IL6R
s-IL6R/
IL6
complex
anti-IL6R
monoclonal antibody
rs2228145 promotes
proteolysis of
membrane bound IL6R
Biomarkers
(IL6, s-IL6R)
Cell
Inflammation
Normal risk of CD and UC
Cell
CRP, platelets
fibrinogen
Inflammation
risk of CD and UC

Supplementary Materials and Methods

Mendelian Randomization

Mendelian randomization (MR) requires that the genetic instrument is associated with a modifiable exposure of interest, and any association between the instrument and the outcome is mediated by the exposure through its downstream effects. MR assesses the causal effect of an exposure (in this report, we use increased soluble IL-6 receptor levels (s-IL6R) as an indirect measure for our exposure, reduced IL6R and subsequent impaired IL-6 receptor (IL6R) signaling, as per previous studies[1,2]) on an outcome (e.g. inflammatory bowel disease (IBD)) by grouping individuals in the population according to the possession of the genetic variants that modify an exposure of interest; genetic variants should be free from conventional confounding since they are "assigned" randomly at meiosis; reverse causality bias is impossible since the germline genome is generally non-modifiable by disease. MR can be thought of as analogous to a randomized controlled trial (RCT) that uses genetic variation as the method of randomization; the key difference is that MR can be performed at any time using routine genotype data, without the exposure of patients to the drug. Advances on MR methodology have allowed for the measurement of genetic effects on exposure and outcomes in independent samples (2 sample MR); this has enabled the use of publically available results from very large genome wide association studies (GWAS) for both risk factor "exposures" and disease "outcomes" to evaluate causal relationships[3]. By using genetic variants that effect known drug targets and have quantifiable effects, MR offers a relatively unbiased approach for

the prediction of both intended (e.g. RCT endpoints) and unintended (e.g. drug repurposing opportunities, adverse events) drug effects.

Genetic datasets used

SNP-Crohn's disease (CD), SNP-ulcerative colitis (UC), ln (odds ratio) and standard errors were downloaded from the most recent IBD Genetics Consortium meta-analysis (https://www.ibdgenetics.org/downloads.html).

Study outcome

The primary outcome was the risk of developing a) CD b) UC per two-fold elevation in circulating s-IL6R as an indirect measure for our exposure (genetically impaired IL6R signaling).

Statistical analysis

The SNP- IBD associations were extracted as described above and presented per effect allele. The SNP - s-IL6R were estimated in 1,650 individuals included in a large collaborative MR study[4]. These estimates were presented in percentage differences, which were converted to ln- transformed units.

Single-nucleotide polymorphism (SNP) - s-IL6R and IL6 associations were collected in ln-transformed units from previous publications (Table 1).

The ratio MR method was used. This method was used to obtain individual SNP estimates by dividing the SNP – outcome by the SNP – biomarker effect estimates. Analysis were performed using R package MR-Base, version 3.2.4 (http://www.r-project.org).[5]

Odds ratio (OR) estimates of UC and CD were transformed (for ease of interpretation). Results were presented per two - fold increase in circulating s-IL6R and IL6 levels:

$$(\sqrt[e]{OR})^2$$

Where OR is the odds ratio estimate per 1-ln increment in biomarker levels and *e* is the base of the natural logarithm.

Sensitivity Analysis

When applying the ratio MR method to model the association between the indirect marker IL6 and the outcome, a two-fold genetic elevation of IL6 was associated with decreased odds of CD (OR 0.75, 95% CI, 0.81-0.94, p = .00003) and UC (OR 0.861, 95% CI, 0.71-1, p = .038).

Table 1. The SNP- s-IL6R and SNP- IL6 associations (per Effect Allele) generated from a previous study were used as genetic instruments in this report. SNP: single nucleotide polymorphism. CRP: C-reactive protein. IL6R: Interleukin-6 receptor. EA: effect allele. OA: other allele. OR: odds ratio.

Gene	SNP	EA	OA	Biomarker	ln(levels)	SE
IL6R	rs2228145	C	A	s-IL6R	0.2949	0.0148
IL6	rs2228145	C	A	IL6	0.1362	0.0176

Table 2. SNP-Crohn's disease and SNP-ulcerative colitis (per Effect Allele) associations downloaded from available GWAS summary statistics, presented here after data harmonization. SNP: single nucleotide polymorphism. CD: Crohn's disease. UC: Ulcerative colitis. CRP: C-reactive protein. IL6R: Interleukin-6 receptor. OR: odds ratio. SE: Standard error. EA: effect allele. OA: other allele.

Study	Disease	SNP	EA	OA	ln(levels)	SE
Liu[6]	CD	rs2228145	C	A	-0.0530323	0.0126631
Liu[6]	UC	rs2228145	C	A	-0.0276154	0.0131298

1. **Harrison SC,** et al. Eur Heart J 2013;34(48):3707-3716.
2. **Interleukin-6 Receptor Mendelian Randomisation Analysis (IL6R MR) Consortium,** et al. Lancet 2012; 379(9822):1214-1224.
3. **Pierce BL,** et al. Am J Epidemiol. 2013;178(7):1177-1184.
4. **IL6R Genetics Consortium Emerging Risk Factors Collaboration,** et al. Lancet 2012;379(9822):1205-1213.
5. **Hemani Gibran,** et al. bioRxiv. doi:doi: https://doi.org/10.1101/078972.
6. **Liu JZ,** et al. Nat Genet 2015;47(9):979-986.