**Siyi Wang**

**Abstract**

abstract

# 1 Introduction

The pleiotropy is a historical research topic that matters to our traits. The first paper on pleiotropy was published 115 years ago, and this term was formally founded by Ludwig Plate from Germany. The application of pleiotropy across diverse research fields has made significant contributions to human knowledge, particularly in medical genetics and trait analysis. Etymologically, the term "pleiotropic" derives from "pleio" (many) and "tropic" (affecting). In scientific terms, pleiotropic analysis can study how single genetic variants or a single locus affect two or more seemingly unrelated phenotypic traits, or, more formally, it is often identified as a single mutation that affects two or more wild-type traits. This pleiotropy analysis helps us understand the shared genetic architecture of two traits by identifying significant genetic loci associated with several phenotypes. [Stearns, 2010] Therefore, we want to perform a two-way pleiotropy analysis between colorectal cancer (CRC) and inflammatory bowel disease (IBD). It is important to note that ulcerative colitis(UC) commonly affects the colon, which is a primary type of IBD. Past experiments indicate that, even after adjusting for patient age and tumor stage, the risk of developing CRC remains increased following a diagnosis of UC. [Jess et al., 2012, Lutgens et al., 2013, Olén et al., 2020]

The current research shows clinical associations but does not clarify whether shared genetic variants drive both diseases through pleiotropy. There is already some evidence showing that CRC has an association with patients with PBC, indicating an increased risk of CRC from a population-based retrospective cohort study [Chen et al., 2025]. In addition, IBD has been identified as commonly associated with liver disease. Co-occurrence of IBD and primary biliary cholangitis (PBC) has been increasingly observed, and a paper published last year detected shared genetic architecture between IBD and PBC. The results provide evidence of a causal relationship between PBC and IBD [Huang et al., 2024].

Sometimes, csmGmm lacks significant findings in the three-way analysis, so I will perform a two-way pleiotropy analysis using composite null hypotheses for CRC-IBD, following Sun's methodology. [Sun et al., 2025] Thus, I will set K = 2 for the two-trait pleiotropy analysis. Statistically, I will reject the null hypothesis that both $\alpha = 0, \beta = 0$ to demonstrate that both effects are non-zero, thereby showing pleiotropy across the two traits as a composite-null task. However, traditional testing approaches focus on single-trait associations based on genome-wide association studies (GWAS), which reject the hypothesis that a variant affects neither trait to prove that the variant affects at least one trait. This simplified approach can lead to misspecification, making it difficult to determine whether a genetic variant simultaneously influences multiple traits based solely on summary z-scores [Kim et al., 2015]. Also, in the past

experiments, LFDR estimation method is required to have trenmendous of gene into consideration and the whole biomarkers population, which is infeasible to achieve for most cases. [Li et al., 2013]

To address these issues, the conditionally symmetric multidimensional Gaussian mixture model (csmGmm) provides a unified multiple empirical-Bayes framework for inference on composite nulls, including mediation, pleiotropy, and replication. This approach imposes symmetry on the multivariate Gaussian distributions and assumes that truly associated SNPs cluster at higher effect magnitudes, which is used to compute local false discovery rates (lfdr) while controlling the false discovery rate (FDR). [Sun et al., 2025] Most importantly, for CRC and IBD, we approach two-trait pleiotropy as a large-scale composite-null problem. We adapt the csmGmm method to work with K=2 traits, which gives us interpretable results based on local false discovery rates. These clear z-statistic results with controlled false discovery rate can generate robust findings for pleiotropic SNPs affecting both CRC and IBD.

# References

Pei-Tzu Chen, Wu-Chien Chien, Chi-Hsiang Chung, Peng-Jen Chen, Tien-Yu Huang, and Hsuan-Wei Chen. Risk of colon polyps and colorectal cancer in primary biliary cholangitis, a population-based retrospective cohort study in taiwan. *Journal of the Formosan Medical Association*, 2025. ISSN 0929-6646. doi: https://doi.org/10.1016/j.jfma.2025.04.007. URL `https://www.sciencedirect.com/science/article/pii/S0929664625001767`.

Wentao Huang, Rui Jiang, Sitao Li, Ruijie Zeng, Yang Li, Yajie Zhang, Shuangshuang Tong, Yanlin Lyu, Jiaxuan Wang, Qizhou Lian, Felix W. Leung, Ruibang Luo, Weihong Sha, and Hao Chen. Investigating shared genetic architecture between inflammatory bowel diseases and primary biliary cholangitis. *JHEP Reports*, 6(6):101037, 2024. ISSN 2589-5559. doi: https://doi.org/10.1016/j.jhepr.2024.101037. URL `https://www.sciencedirect.com/science/article/pii/S2589555924000387`.

Tine Jess, Jacob Simonsen, Kristian Tore Jørgensen, Bo Vestergaard Pedersen, Nete Munk Nielsen, and Morten Frisch. Decreasing risk of colorectal cancer in patients with inflammatory bowel disease over 30 years. *Gastroenterology*, 143(2):375–381.e1, 2012. ISSN 0016-5085. doi: https://doi.org/10.1053/j.gastro.2012.04.016. URL `https://www.sciencedirect.com/science/article/pii/S0016508512005677`.

Junghi Kim, Yun Bai, and Wei Pan. An adaptive association test for multiple phenotypes with gwas summary statistics. *Genetic Epidemiology*, 39(8):651–663, 12 2015. doi: 10.1002/gepi.21931. URL `https://doi.org/10.1002/gepi.21931`.

Zuojing Li, Qing Li, Lulu Geng, Xiaohui Chen, and Kaishun Bi. Use of the local false discovery rate for identification of metabolic biomarkers in rat

urine following genkwa flos-induced hepatotoxicity. *PloS one*, 8(7):e67451, 2013. doi: 10.1371/journal.pone.0067451. URL https://doi.org/10.1371/journal.pone.0067451.

Maurice W. M. D. Lutgens, Martijn G. H. van Oijen, Geert J. M. G. van der Heijden, Frank P. Vleggaar, Peter D. Siersema, and Bas Oldenburg. Declining risk of colorectal cancer in inflammatory bowel disease: An updated meta-analysis of population-based cohort studies. *Inflammatory Bowel Diseases*, 19 (4):789–799, 02 2013. ISSN 1078-0998. doi: 10.1097/MIB.0b013e31828029c0. URL https://doi.org/10.1097/MIB.0b013e31828029c0.

Ola Olén, Rune Erichsen, Michael C Sachs, Lars Pedersen, Jonas Halfvarson, Johan Askling, Anders Ekbom, Henrik Toft Sørensen, and Jonas F Ludvigsson. Colorectal cancer in ulcerative colitis: a scandinavian population-based cohort study. *The Lancet*, 395(10218):123–131, 2020. ISSN 0140-6736. doi: https://doi.org/10.1016/S0140-6736(19)32545-0. URL https://www.sciencedirect.com/science/article/pii/S0140673619325450.

Frank W Stearns. One hundred years of pleiotropy: A retrospective. *Genetics*, 186(3):767–773, 11 2010. ISSN 1943-2631. doi: 10.1534/genetics.110.122549. URL https://doi.org/10.1534/genetics.110.122549.

Ryan Sun, Zachary R. McCaw, and Xihong Lin. Testing a large number of composite null hypotheses using conditionally symmetric multidimensional gaussian mixtures in genome-wide studies. *Journal of the American Statistical Association*, 120(550):605–617, 2025. doi: 10.1080/01621459.2024.2422124. URL https://doi.org/10.1080/01621459.2024.2422124.