Large-scale genetic analyses in an understudied disease: haemorrhoidal disease

Eleonora A M Festen, Rinse K Weersma

Approximately 20% of the US population have ever been diagnosed with haemorrhoidal disease, and out of those individuals, one in five reports to have had surgery for his or her haemorrhoids. Outpatient and emergency department visits to gastroenterologists and surgeons for haemorrhoids outnumber those for colon cancer, inflammatory bowel disease or irritable bowel syndrome. Historically accepted notions that constipation and a low-fibre diet contribute to disease development have not been confirmed in large US and UK survey and registry data studies.[1] Moreover, these studies found that generally accepted symptoms of haemorrhoids, like pruritus, pain, mucus discharge and bleeding, occur just as frequently in individuals without haemorrhoids. Haemorrhoidal disease is a classic example of a disease that has a large burden within the population, but for which scientific knowledge and research efforts are nonetheless very limited. A recent narrative review by Sandler and Peery thoroughly exposed the large gaps in our understanding of the underlying disease mechanisms and the lack of evidence-based treatment options in haemorrhoidal disease.[2]

In *Gut*, Zheng *et al* focused on the mechanisms underlying haemorrhoidal disease and set out to study the genetic components in its development.[3] The study is noteworthy and laudable for two reasons. First of all, it represents a state-of-the-art genome-wide association study (GWAS) making use of available data from (population based) biobanks resulting in a massive study involving just under a million participants. Second, the authors take the results of the initial GWAS and perform a series of downstream analyses, which give thorough insights into the genetics of haemorrhoids, and provide functional data that pinpoint the underlying disease mechanisms being smooth muscle, epithelial and connective tissue dysfunction.

While in the early days of GWAS dedicated disease cohorts were established and compared with healthy controls, the current study takes a different approach. Large-scale, well-phenotyped population biobanks, or biobanks from commercial entities like 23andMe, with genomic data from genotyping arrays, or even with whole genome sequencing data already available, allow for large population-based GWAS of relatively common phenotypes. Although the phenotypic resolution for the disease under study is very low (eg, International Statistical Classification of Diseases and Related Health Problems (ICD) codes in the UK biobank or self-reported haemorrhoids in 23andMe), the trade-off is the sheer number of participants, yielding extraordinary statistical power, and the possibility of replicating initial findings in independent cohorts.[4] We have previously called on the gastroenterology and hepatology research community to make more use of the possibilities that these biobanks offer: they provide unprecedented opportunities to study common disease phenotypes like haemorrhoids or diverticulosis without investing large amounts of scientific resources (diverticulosis has been studied in a similar manner by the same group).[5 6]

Another important aspect of the study is the inclusion of polygenic risk scores. Analyses of mutations in single genes for the identification of individuals at risk for developing monogenic diseases are increasingly being employed in clinical approaches. However, most common diseases are not monogenic but have a polygenic inheritance pattern involving many common genetic variants with small effect sizes. Recently, it has been proven that the combination of these common genetic variants in polygenic risk scores have similar predictive values for common polygenic diseases as those of single mutations for monogenic diseases.[7] Interestingly, Zheng *et al* put the polygenic risk scores derived from the GWAS analyses forward into independent cohorts and show that these identify individuals at risk for early development of haemorrhoids or having more severe disease (defined by requiring surgical interventions). In addition, the authors of the current study have performed cross trait analyses showing that almost half of the genetic variants they identified as risk variants for hemorrhoidal disease are also associated with diverticular disease and irritable bowel syndrome. The fact that an increasing number of individuals have their genomic data already available, in combination with the predictive value of polygenic risk scores for common polygenic disorders, and the pleiotropy of genetic risk variants within related common disease phenotypes, opens up the possibility for preventative measures.

Now back to the findings of Zheng *et al* in *Gut*: to what extend do these findings improve our understanding of haemorrhoidal disease? The authors show that there indeed is a hereditary component for haemorrhoids and implicate 102 independent genomic loci in the pathogenesis. Their main findings identify three categories of disease mechanisms underlying haemorrhoids: first, *the extracellular matrix, elasticity of the connective tissue and smooth muscle function*; second, the *vasculature and circulatory system*; and third, the involvement of *gut motility*. In the first two categories, the authors find several potential functional variants affecting genes with functions in extracellular matrix organisation and muscle contraction, but further interpretation is hampered because many of these genes remain poorly characterised. In the category of gut motility, the authors however report a compelling finding: the identification of a missense variant in the gene *ANO1* which is known to influence the function of anoctamin-1 which is an ion channel in the interstitial cells of Cajal involved in intestinal peristalsis, and is also implicated in irritable bowel syndrome. Finally, the authors report an intriguing finding that does not fall into either of these three categories. The second and third strongest GWAS signals were detected at loci linked to genes involved in

Department of Gastroenterology and Hepatology, University of Groningen and University Medical Center, Groningen, The Netherlands

Correspondence to Prof Dr Rinse K Weersma, Gastroenterology and Hepatology, University Medical Centre Groningen, Groningen, The Netherlands; r.k.weersma@umcg.nl

the determination of blood groups, which by itself is hard to interpret mechanistically. However, the authors indeed show that individuals with blood group O have an increased risk for haemorrhoidal disease, whereas those with blood groups A and B have a decreased risk.

All in all, Zheng *et al*, while not completely solving our knowledge gaps for haemorrhoidal disease, provide valuable new insights into its underlying biological mechanisms, making optimal use of the availability of large-scale population-based biobanks. The study also implies shared genetic and pathophysiological components between haemorrhoidal disease and other common GI diseases. While the development of haemorrhoids may not be the top priority for healthcare, improving general gut health, with broader preventative measures also directed at diverticular disease and irritable bowel syndrome certainly should be.

Contributors Manuscript writing: EAMF and RKW.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests RKW received unrestricted research grants from Takeda Pharmaceutical Company, Johnson and Johnson, Ferring and has acted as consultant for Takeda Pharmaceutical Company. EAMF received an unrestricted research grant from Takeda Pharmaceutical Company.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting or dissemination plans of this research.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Festen EAM, Weersma RK. *Gut* 2021;**70**:1429–1430.

Received 14 May 2021
Accepted 20 May 2021
Published Online First 25 May 2021

Linked

► http://dx.doi.org/10.1136/gutjnl-2020-323868

Gut 2021;**70**:1429–1430.
doi:10.1136/gutjnl-2021-324817

ORCID iD
Rinse K Weersma http://orcid.org/0000-0001-7928-7371

REFERENCES

1 LeClere FB, Moss AJ, Everhart JE, *et al*. Prevalence of major digestive disorders and bowel symptoms, 1989. *Adv Data* 1992;212:1–15.
2 Sandler RS, Peery AF. Rethinking what we know about hemorrhoids. *Clin Gastroenterol Hepatol* 2019;17:8–15.
3 Zheng T, Ellinghaus D, Juzenas S, *et al*. Genome-Wide analysis of 944 133 individuals provides insights into the etiology of haemorrhoidal disease. *Gut* 2021;70:1538–49.
4 DeBoever C, Tanigawa Y, Aguirre M, *et al*. Assessing digital phenotyping to enhance genetic studies of human diseases. *Am J Hum Genet* 2020;106:611–22.
5 Weersma RK, Parkes M. Diverticular disease: Picking pockets and population biobanks. *Gut* 2019;68:769–70.
6 Schafmayer C, Harrison JW, Buch S, *et al*. Genome-Wide association analysis of diverticular disease points towards neuromuscular, connective tissue and epithelial pathomechanisms. *Gut* 2019;68:854–65.
7 Khera AV, Chaffin M, Aragam KG, *et al*. Genome-Wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat Genet* 2018;50:1219–24.