A novel framework for assessing causal effect of microbiome on health: long-term antibiotic usage as pseudo-randomizer

Nele Taba[1], Krista Fischer[1,2], Elin Org[1], Oliver Aasmets[1]

[1] Estonian Genome Center, Institute of Genomics, University of Tartu, Tartu, Estonia

[2] Institute of Mathematics and Statistics, Faculty of Science and Technology, University of Tartu, Tartu, Estonia

Abstract

Gut microbiome (MB) has been widely shown to affect human health. Since MB in turn can be altered by various exposures, such as diet and medications, it holds immense potential for future treatments and healthy ageing. On the one hand, faecal microbiota transplantation and Mendelian Randomization have proven a causal link between treatment, MB and diseases. On the other hand, assessing the causality of the MB effects on health has remained challenging, since randomised trials in human subjects are often unethical or difficult to pursue, and Mendelian Randomization lacks valid instruments. Thus, novel analytical approaches are needed for inferring causal associations.

To overcome these barriers, we propose a novel framework of antibiotic instrumental variable regression (AB-IVR) for estimating the causal relationships between MB and various diseases. Our inspiration originates from the popular Mendelian Randomization method that uses genetic mutations as instruments in the instrumental variable analysis (IVR). Further, we rely on the recently shown results that antibiotic (AB) treatment has a cumulative long-term effect on MB, consequently pseudo-randomizing individuals with higher AB usage to have more perturbed MB. Thus, we developed a new AB-IVR framework to utilise the long-term AB usage as an instrument in the IVR for assessing the causal effect of MB on health.

We pursued a plethora of sensitivity analyses to explore the properties of our method: varying the sample's age group and maximum number of AB used; using a buffer-time for incident disease outcomes to account for feedback-mechanism; using subgroups of AB as instrument; and simulating data for disease outcomes. We detected several interesting causal effects of MB on health outcomes; some causal effects – such as MB effects on migraine, depression, irritable bowel syndrome, and several more – remain significant irrespective of the sensitivity analysis used. We believe that our AB-IVR framework has promising potential to be the new widely used method for assessing MB effect on health.

Introduction

Human microbiome studies have demonstrated that the gut microbiome can be affected by various exposures such as diet and medications (Tarini Shankar Ghosh *et al.*, 2020; Aasmets *et al.*, 2022) and changes in the gut microbiome composition have been associated with the prevalence or susceptibility to complex diseases such as type 2 diabetes (Gurung *et al.*, 2020; Aasmets *et al.*, 2021; Ruuskanen *et al.*, 2022), Crohn's disease (Pascal *et al.*, 2017) and different cancers (Wirbel *et al.*, 2019; Kartal *et al.*, 2022). However, with the primary interest in disease associations, determining the causal effects has remained challenging. Randomised trials are seen as the golden standard for causal discoveries in human subjects. However, they are often difficult or unethical to conduct due to the unexpected or possible harmful effects on health. Faecal microbiota transplantation methods have proven a causal mechanistic link between treatment, microbiome and health (Ser *et al.*, 2021), but to date the methods for determining such causal associations noninvasively are limited (Lv, Quan and Zhang, 2021). Mendelian Randomization (MR) is a promising computational method to tackle these questions, and has been used to show the causal role of several taxa on cancers (Long *et al.*, 2023), inflammatory bowel diseases (Liu *et al.*, 2022) and even depression (Chen *et al.*, 2022). However, MR is limited by the lack of valid genetic instruments - this might remain an issue even with larger sample sizes, since microbiome is shown to be with low heritability (Wade and Hall, 2019; Kurilshikov *et al.*, 2021). To fill this void, we propose a new analytical approach to assess the causal role of microbiome on health: antibiotic instrumental variable regression (AB-IVR).

Our inspiration comes from the Mendelian Randomisation method, where genetic variants are used as instruments in the instrumental variable analysis setting (Burgess and Thompson, 2015). Instead of genetic variants, we use the long-term history of antibiotic usage (hAB) as an instrument. Antibiotic usage is shown to affect microbiome in the long term in a cumulative manner, whereby larger number of antibiotics prescribed in the past results in a more perturbed microbiome composition (Forslund *et al.*, 2021; Aasmets *et al.*, 2022). This means that there are on average consistent differences in microbiome between the individuals with higher and lower long-term antibiotics usage. Importantly, this long-term usage effect is already evident in subjects who have taken more than only 3 courses of AB in the last 10 years before the sample collection. Such antibiotic consumption is highly common in the general population. This, and the fact that antibiotics are expected to affect health primarily via microbiome and not other pathways, made us ponder whether hAB - measured as number or antibiotics prescribed during 10 years prior to microbiome sampling - can serve as a natural experiment randomising individuals to have more or less perturbed microbiome. Thus, we propose that by using hAB as an instrument, we can compare the disease incidence between

groups that have consistent long-term differences in microbiome, and by this assess the causal role of microbiome in disease. We consider the AB-IVR method as an intermediate option between the observational studies and randomised trials: we aim to gather more information compared to observational studies, but do not see it as an equal alternative to the randomised trials.

Next sections introduce the AB-IVR methodology and its usage on the Estonian Biobank samples. Firstly, we introduce the method as a two-sample two-stage least squares procedure and the corresponding sample sets that we utilise for estimating the causal effects. Next, we introduce the concept of long-term antibiotics usage as an instrument. We discuss model assumptions and limitations that can bias the results such as the scope of the estimable effects, feedback loops arising from the bidirectional microbiome-disease crosstalk and effects of general health behaviour. Lastly, we assess the causal role of microbiome in several common diseases, test the validity of the findings through a series of sensitivity analysis and give insights for future developments. The AB-IVR method showed that microbiome has a causal effect for several common diseases such as irritable bowel syndrome, migraine and depression, and we believe that it opens up new possibilities for causal discovery in the human microbiome field.

Results

Study overview and data selection

For inferring the causal effect of microbiome on disease we utilised the method of Two-Sample Two-Stage Least Squares (TSTSLS), which is a special case of instrumental variable regression. In our case, the history of antibiotics usage (hAB) serves as the instrumental variable similarly to genetic variants serving as an instrument in the Mendelian randomization framework (Antibiotic instrumental variable regression, AB-IVR) (**Fig. 1A**). The TSTSLS methodology relies on two independent datasets which are used for 1) estimating the effect of hAB on disease (**Fig. 1A**, $\boldsymbol{\beta}_{\mathbf{D,hAB}}$) and 2) estimating the effect of hAB on the microbiome ($\boldsymbol{\beta}_{\mathbf{MB,hAB}}$). The two effects are combined to evaluate the causal effect of microbiome on disease ($\boldsymbol{\beta}_{\mathbf{D,MB}}$). For estimating the effect of hAB on disease, we leveraged the electronic health records data available for the Estonian Biobank (EstBB) participants (N > 210 000) (Leitsalu *et al.*, 2015). As prevalent diseases can have an effect on the microbiome and potentially create a feedback loop, which biases the results, we analysed the effect of previous antibiotics usage on the incident diseases instead (**Fig. 1B**) (See Methods). To estimate the effect of hAB on the microbiome, we used Estonian Microbiome (EstMB, a subcohort of EstBB) cohort (N > 2500) samples with shotgun metagenomics data available (**Fig. 1C**) (Aasmets *et al.*, 2022).

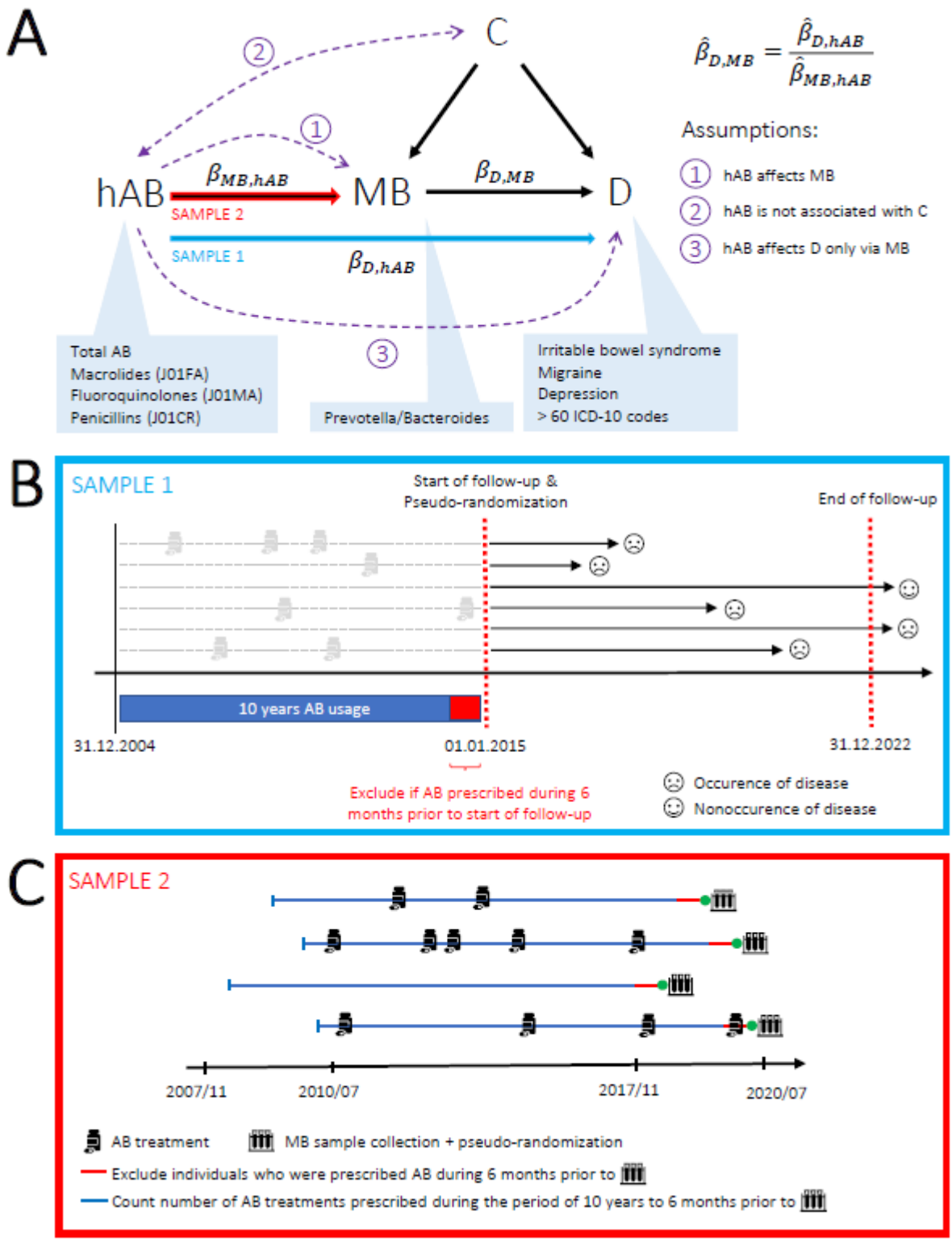


Figure 1. Graphical description of the study design. Upper panel **(A)** illustrates the instrumental variable regression and corresponding assumptions schematically in the context of our study. AB - antibiotic usage; MB - microbiome; D - disease; C - confounder. Middle panel **(B)** in the blue box describes the formation of Sample1, where disease outcomes were measured, and the lower panel **(C)** in the red box describes the formation of Sample2, where microbiome was measured. We define the moment of pseudo-randomization as the start of follow-up in Sample1 (01.01.2015) and the collection of the microbiome in Sample2 (arbitrary moment between November 2017 to July 2020) (See Methods).

In both samples, we recorded the AB usage as the number of AB prescribed (total and AB-subgroups) during the period of 10 years to 6 months preceding the pseudo-randomization, whereas individuals who were prescribed AB during the 6 months preceding the pseudo-randomization were excluded from the samples. Sample1 was followed up for incident disease outcomes (64 common diseases) until 31.12.2022, whereafter observations are right-censored irrespective of the future outcome. For each disease we excluded the prevalent cases from the sample prior to analysing the corresponding outcome. AB recording scheme in Sample2 (lower panel in red box): each individual had their own period when AB was measured, corresponding to when their MB sample was collected; starting from the top, the first individual was prescribed two antibiotics, second individual five antibiotics, third individual was not prescribed any antibiotics (but is included in the sample) and fourth individual was prescribed antibiotics during the last six months prior to MB sample collection and was thus excluded from the sample. The causal effect of MB on disease ($\beta_{D,MB}$) in the two-sample setting is estimated as the ratio of the effect of AB-usage on disease ($\beta_{D,hAB}$) in Sample1 (EstBB) and the effect of AB-usage on MB ($\beta_{MB,hAB}$) in Sample2 (EstMB).

Previous antibiotics usage as a valid instrument

Estimating the causal effects by instrumental variable regression analysis relies on several assumptions. Firstly, the hAB must be associated with the gut microbiome (**Fig. 1A**). We and others have previously shown that hAB is associated with the microbiome composition additively, meaning that higher hAB is associated with larger changes in the microbiome composition (Forslund *et al.*, 2021; Aasmets *et al.*, 2022). We extend this result by showing that this also holds for different classes of antibiotics, namely macrolides (J01FA), penicillins (J01CR) and fluoroquinolones (J01MA), which we will use for the sensitivity analysis (**Fig. 2A**). It must be highlighted that the different antibiotics classes are largely uncorrelated in their usage and they have differing long-term effects on the microbiome composition (**Fig. 2C, Fig. 2D**).

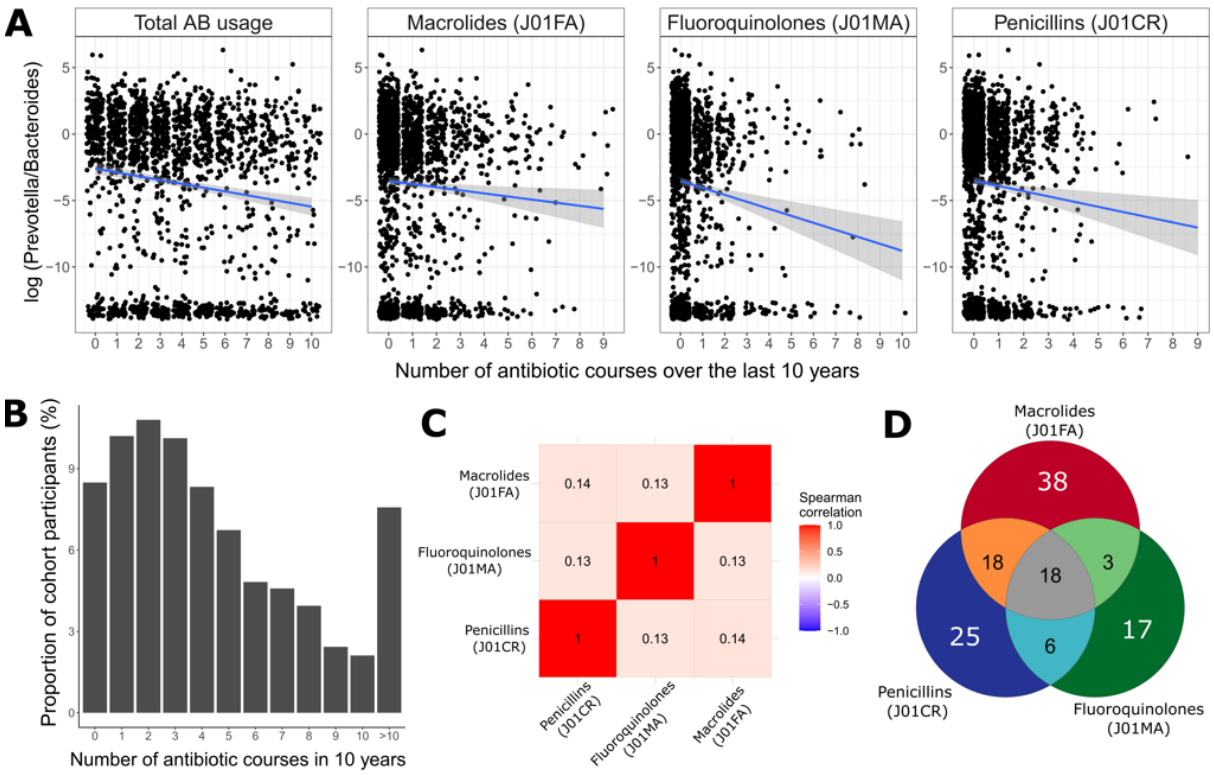


Figure 2. Antibiotic usage in the EstMB cohort. Panel **A** shows the association between the number of different antibiotics used during the last 10 years before the sample collection and *Prevotella-Bacteroides* ratio. The total AB usage and usage of all AB subclasses was strongly associated with *Prevotella/Bacteroides* ratio (combined AB usage p=2.04e-8, macrolides p=0.0192, fluoroquinolones p=3.03e-5, and penicillins p=0.0026). Panel **B** shows the proportion of cohort participants by the number of antibiotics used during the past 10 years. Panel **C** shows the Spearman correlation between the number of different antibiotics classes used during the last 10 years before the sample collection. Panel **D** shows the unique and shared hits of the univariate analyses associating the antibiotics usage history with the abundance of microbial species.

Secondly, the usage of antibiotics and the disease cannot have any common cause. This assumption can be easily violated when prevalent disease cases are analysed. Namely, the disease could for example weaken the immune system, which would in turn result in higher antibiotics consumption, often referred to as a “feedback loop”. We address this issue by analysing only incident disease cases. As a sensitivity analysis we additionally analyse disease cases that have been diagnosed more than 5 years after the sample collection, to eliminate the possibility of prevalent diseases that have not yet been diagnosed, but are already present and affecting the microbiome composition. Violation of the second assumption can also occur when unhealthy subjects who are more prone to comorbidities and taking antibiotics are studied. To address this concern, we focus our analysis only on younger people (ages 23-50), who have taken up to 5 courses of antibiotics during 10 years prior to pseudo-randomization (See Methods, **Fig. 1**). The lower limit for age refers to the minimum age in the EstMB cohort. The chosen threshold for the number of antibiotics courses is highly common in the population, thus it is meant to represent the healthier part of the population (**Fig. 2B**), allowing us to exclude individuals with chronic and extreme AB usage. As sensitivity analyses, we study older people, extending the age range to 89 (refers to the maximum age in EstMB cohort), and usage of antibiotics up to 10 courses.

Thirdly, antibiotics cannot have a direct effect on the disease that is not mediated by the microbiome composition. Although some authors have indicated that antibiotics can have such microbiome-independent effects, the evidence is currently weak and we consider that most of the effect is based on altering the microbiome composition (Patangia *et al.*, 2022).

Causal role of the microbiome on diverse disease groups confirmed with the novel AB-IVR framework

We applied the AB-IVR methodology to assess whether the microbiome has a causal role for the development of 64 conditions across diverse disease groups. We focused primarily on common chronic diseases and cancers with at least 50 incident cases. Due to the broad selection of diseases, we did not apply any disease specific inclusion and exclusion criteria, but future studies should take the disease specificity into account. Antibiotics have a complex

effect on MB via altering the abundances of several members of the composition simultaneously. Therefore, when the analysis is based on general hAB, various taxa are pseudo-randomized simultaneously and thus it is not possible to imply which specific species or genera are the culprit in disease formation. This is a limitation of our methodology and remains a challenge to be solved by future improvements. Therefore, instead of using the data of specific species or genera to characterise the changes in the microbiome, we analysed the effect of Prevotella/Bacteroides ratio (P/B ratio) that represents a summarised state of microbiome and is also strongly associated with the antibiotics usage history (**Fig. 2A**).

In total, the microbiome was identified to have a causal effect on 29 diseases (FDR <= 0.05) (**Fig. 3A**, **Supplementary Fig. 1, Supplementary Table 1**). These diseases include cardiometabolic diseases like cardiac arrhythmias, intestinal diseases like irritable bowel syndrome and gastro-esophageal reflux disease, skin diseases like atopic dermatitis, and several mental disorders like anxiety disorders and depression. The causal estimates for a change per standard deviation (sd) of the log *Prevotella*/*Bacteroides* ratio are highest for lactose intolerance, migraine and irritable bowel syndrome with the dominance of *Prevotella* defending against the disease progression. Notably, we did not identify the causal effect of the microbiome on any of the cancers studied, which is likely due to excluding older individuals from the analysis.

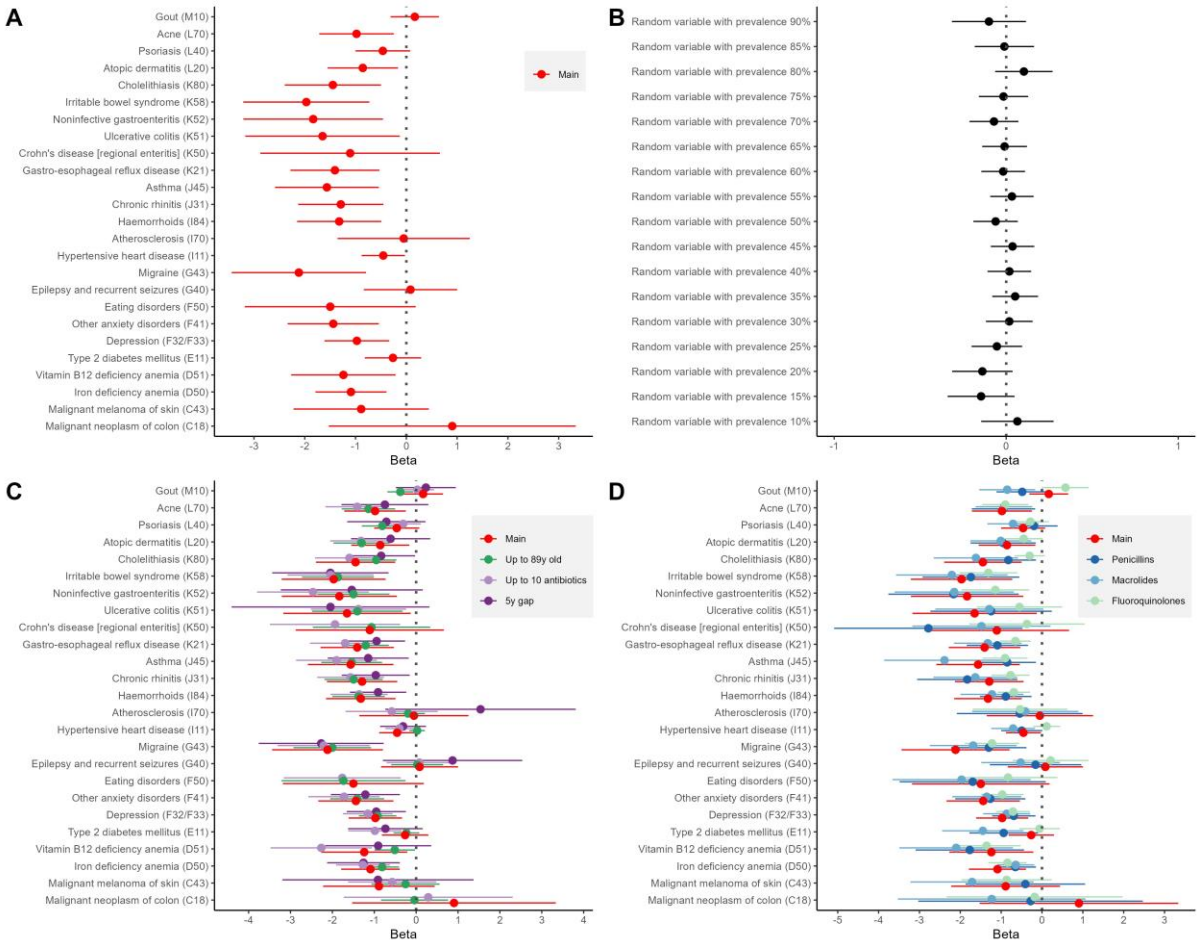


Figure 3. Results of the main analysis (A) and sensitivity analyses (B,C,D). On panels **A**,**C**, and **D** the causal effect of *Prevotella*/*Bacteroides* ratio on a selection of diseases is presented (full results comprising all diseases analysed can be viewed in **Supplementary Table 1**). In panel **A** the age is filtered as 23-50, maximum number of AB prescribed is five, and minimum number of cases per disease is 50. Panel **B** describes the sensitivity analysis where disease outcomes were randomly generated from the binomial distribution with fixed disease probability (originally ranging from 2.5% to 97.5%, for readability ranges from 10% to 90% are depicted); other settings were identical to the main analysis. Panel **C** represents the sensitivity analyses where sample formation varies, whereas the main analysis is depicted for comparison in red. Green shows the effect estimates, when age is filtered as 23-89; light-purple shows the effect estimates, when the maximum number of AB prescribed is 10; dark-purple describes the scenario where the first five years of incidence after the pseudo-randomization is considered as prevalent disease. Panel **D** represents the sensitivity analyses where information regarding subclasses of AB were used instead of the total amount of AB prescribed, whereas other settings were identical to the main analysis**.** Red corresponds to the main analysis (same as **A**), dark blue corresponds to the class of penicillins (J01CR), light-blue corresponds to the class of macrolides (J01FA) and light-green corresponds to the class of fluoroquinolones (J01MA).

Sensitivity analysis and assessment of different antibiotics as instruments

We carried out several sensitivity analyses to assess the validity of the model estimates. Firstly, we simulated random binary variables with different event probabilities to see how our approach behaves in a scenario where no effect is expected. Indeed, the method did not identify any effects and manages to control the false discovery rate perfectly (**Fig. 3B**). Next, we focused on different scenarios which would test the model behaviour when the study population includes subjects who are more likely to have health complications, thus testing the second model assumption. For that we considered subjects up to 89 years old (refers to the maximum age in EstMB cohort), subjects that had been prescribed antibiotics up to 10 times during 10 years prior to pseudo-randomization and considering a 5 year “gap” in defining the incident cases after the pseudo-randomization (**Fig. 3C, Supplementary Fig. 2, 3, Supplementary Tables 2, 3, 8**). The reasoning behind the chosen sensitivity analysis is discussed in detail in the Methods. We observed no major differences in the estimated causal effects for any of the tested scenarios. Most notably, the causal estimates for diseases that showed strongest effects in the primary analysis remained largely untouched by the scenarios tested. This includes diseases like migraine, irritable bowel syndrome, chronic rhinitis and depression. Lastly, we carried out sensitivity analysis to test the performance of different antibiotics subclasses as instruments (**Fig. 3D, Supplementary Tables 5, 6, 7**). Again, the results remained similar and the previously highlighted causal estimates remained largely unchanged and independent of the chosen instrument. Since different antibiotics classes are largely uncorrelated in their usage and they have differing long-term effects on the microbiome composition (**Fig. 2C, Fig. 2D**), the differences between the effect estimates can originate from the varying targets of antibiotic subclasses and can refer to divergent patterns in influenced subcommunities. This might be the case for gout (ICD10 code M10) where the past

usage of macrolides and fluoroquinolones lead to causal effects estimates with opposite directions, but the combined antibiotics usage estimates no causal effect (**Supplementary Fig. 3**). However, the overall similarity of the estimates for most of the diseases confirms the underlying causal effect of the microbiome as a whole. Lastly, since the logarithm of the P/B ratio is with a bimodal distribution, we additionally performed an analysis where the P/B ratio was inverse normal transformed. However, we did not observe any notable discrepancies compared to the main analysis (**Supplementary Fig. 1, Supplementary Table 4**).

Discussion

We introduce and demonstrate an analytical approach that expands the causal inference toolbox for microbiome studies by using long-term antibiotics usage as an instrument in an instrumental variable regression setting. We show that the history of antibiotics usage can be a valid instrument and that the microbiome has a causal role for several diseases, such as migraine, irritable bowel syndrome, depression and many others.

Antibiotics are meant to kill or inhibit the growth of bacteria and are commonly directly or indirectly effective against several members of the microbial community (Anthony *et al.*, 2022). Therefore, using long term usage of antibiotics as an instrument allows us to estimate the causal effect of the disturbed part of the community. This brings along two limitations for the proposed methodology: we cannot identify the effects of single members of the community; and we cannot rule out the potential causal effects of the members that are unaffected by the antibiotics used. Thus, the microbiome can still be causal for the development of diseases for which our proposed methodology did not show a causal effect. Tackling these limitations opens up several directions for future development. Since different antibiotic subgroups are intended to have an effect on different bacteria, one can use or design instruments that are specific to certain subcommunities or taxa to estimate the causal effects of interest. Also, several instruments such as long-term usage of different antibiotic subclasses and other drugs with persistent effects on microbiome can be combined and used simultaneously, which in turn opens the opportunity to utilise various sensitivity analyses that are developed for 2-sample Mendelian Randomization framework. For example, consumption of numerous host-targeted drugs are known to affect the microbiome (Maier *et al.*, 2018; Forslund *et al.*, 2021) and antidepressants have also been shown to have long term effects (McGovern, Hamlin and Winter, 2019), but as microbiome-related research is currently rapidly evolving, many more suitable instruments, not limited to medications, are likely emerging in the near future.

Here, we showed in principle how long-term effects of antibiotics can be used to identify the causal role of the microbiome in a large set of common diseases. However, certain limitations must be kept in mind while interpreting the results for a certain outcome. We did not focus

separately on any single disease, thus future studies should consider disease-specific inclusion-exclusion criteria when designing the study. An important consideration concerns the role of general health behaviour. It is possible that the overall awareness and interest in one's health promotes the usage of antibiotics, which also indirectly changes the disease risks. Variables such as gender and education might entail general health behaviour and consequently simultaneously affect AB-usage and health-outcomes (eg. individuals with higher education might seek more help from health-care professionals (thus have different AB-consumption habits), have better eating habits, smoke less, be more physically active etc). Therefore we advise any future studies using AB-IVR framework to account for such possible confounders. Lastly, the amount of antibiotic consumption is highly variable in different populations and antibiotics consumption in Estonia is among the lowest in Europe (Bruyndonckx *et al.*, 2021). Thus, the results need to be validated in different populations to account for the differences in overall burden of antibiotics usage.

We demonstrated a novel AB-IVR approach by assessing the effect of *Prevotella/Bacteroides* ratio on a large set of diseases. The results confirmed the utility of the introduced method and we believe it has large potential to become a new widely used framework that allows to assess analytically the causal effect of microbiome on health. Further research with disease-specific inclusion-exclusion criteria is warranted for drawing conclusions about specific health-effects and several exciting future development options arised.

Methods

Sample description

The proposed method combines the information from two different samples, whereas the association between the instrument and the exposure is analysed in one sample and the association between the instrument and the outcome in another sample (**Fig. 1**). In our case the association between antibiotics usage and microbiome is assessed in the Estonian Microbiome Cohort (Aasmets *et al.*, 2022) (**Fig. 1c**) and association between antibiotics usage and diseases in the Estonian Biobank data (Leitsalu *et al.*, 2015) (**Fig. 1b**).

The Estonian Biobank (EstBB) is a volunteer-based population cohort initiated in 1999 that currently includes over 210,000 genotyped adults (≥ 18 years old) across Estonia. Estonian Microbiome Cohort (EstMB) was initiated in 2017 when more than 2500 EstBB participants, who joined the EstBB at least 10 years before, provided stool, oral, and blood samples. The EstMB cohort, microbiome sample collection, stool bacterial DNA extraction and shotgun metagenomic sequencing are described in detail in (Aasmets *et al.*, 2022). For the current

project, taxonomic profiling of the gut microbiome was carried out using *Metaphlan3* (Beghini *et al.*, 2021). The *Prevotella-Bacteroides* ratio, which we considered as the primary indicator of the microbiome inter-individual variability in our analysis, was calculated on the *genus* level after imputing zeros with a pseudocount equal to half of the minimal non-zero relative abundance observed in the data. For both, EstMB and EstBB one of the major advantages is the possibility to use electronic health records (EHR) data and follow the participants' health both retrospectively and prospectively. Using EHR, the number of antibiotics used before the pseudo-randomization and the disease incidence were analysed. The EstMB participants were removed from the EstBB sample to ensure that the two samples are independent.

Pseudo-randomization

An important aspect in the following data-processing is the moment of pseudo-randomization (**Fig. 1**), which we consider to be the moment of MB-collection in the EstMB cohort and the start of follow-up in the EstBB cohort. In the EstMB cohort the moment of pseudo-randomization varies between individuals, whereas in the EstBB cohort it is set to 1st January 2015. The reason for the latter is that the history of antibiotics usage is not properly recorded before 2005 in the electronic health records and the chosen cut-off date allows us to analyse the antibiotics consumption during a 10-year period prior to follow-up in the EstBB cohort. From both samples we excluded the individuals who had been prescribed AB during the six months preceding the pseudo-randomization.

Defining antibiotics usage and incident diseases

To characterise the long-term antibiotics usage, we quantified the number of antibiotics treatments prescribed in the period between the last 6 months and last 10 years before the pseudo-randomization. All drugs with the Anatomical Therapeutic Chemical (ATC) classification code J01* were considered as antibiotics. We chose the 10-year period similar to Aasmets *et al. (2022)*, where we showed a cumulative long-term effect of antibiotics usage. Additionally, we considered separately the history of usage of penicillins (J01CR), macrolides (J01FA) and fluoroquinolones (J01MA) since these classes have the strongest long-term effects on gut microbiome (Aasmets *et al., unpublished data*).

We restricted the analysis to incident diseases, which we defined as the first occurrence of an ICD-10 code of interest in the EHR after the start of follow-up (1st January 2015) until the end of follow-up (31st December 2022). We focused primarily on the common diseases and cancers with at least 50 incident cases in the EstBB cohort. Summary of the selected diseases is shown in **Supplementary Table 1**. Furthermore, we excluded from any analysis of an incident disease the prevalent cases for the corresponding disease. We did not use any other disease-specific exclusion criteria for the method demonstration. However, we do encourage

to utilise disease-specific exclusion criteria when a more thorough analysis of a specific disease is of interest.

Main method

For inferring the causal effect of MB on disease we utilised the method of Two-Sample Two-Stage Least Squares (TSTSLS), which is a special case of instrumental variable regression for the two-stage estimation in a two-sample setting. A detailed description of TSTSLS can be found elsewhere (Pacini and Windmeijer, 2016). Rationale for using the two-sample setting is that the microbiome is often measured in smaller samples compared to incident disease outcomes available in the large biobank samples. Using TSTSLS allows us to leverage the information from such large samples. We denote in the following formulas the estimated effect of instrument (hAB) on exposure (MB) as $\beta_{MB,hAB}$, instrument on outcome (D) as $\beta_{D,hAB}$, and exposure on outcome, i.e. the causal relationship of interest, as $\beta_{D,MB}$. We estimate $\beta_{D,hAB}$ in the EstBB cohort (Sample1) using the logistic regression model and $\beta_{MB,hAB}$ in the EstMB cohort (Sample2) using the linear regression model. Thereafter we calculate the effect-estimate for the causal relationship as a ratio of coefficients:

$$\beta_{D,MB}=\frac{\beta_{D,hAB}}{\beta_{MB,hAB}}.$$

Further, we calculate the standard deviation for the estimated $\beta_{D,MB}$ based on (Pacini and Windmeijer, 2016) as:

$$\sigma_{D,MB}=\sqrt{\frac{\sigma^2_{D,hAB}+\beta^2_{D,MB}*\sigma^2_{MB,hAB}}{\beta^2_{MB,hAB}}},$$

and obtained the p-values based on z-scores calculated as $\beta_{D,MB}/\sigma_{D,MB}$, and standard normal distribution. Further, we calculated the confidence intervals for the estimates as $\beta_{D,MB} \pm 1.96 * \sigma_{D,MB}$.

To account for multiple testing, we used the Benjamini-Hochberg procedure to control for the false discovery rate (FDR). We set the significance threshold for FRD-adjusted p-value to 0.05.

Model assumptions

The obtained $\beta_{D,MB}$ is a valid causal estimate if the following assumptions are met (Burgess and Thompson, 2015):

- AB usage history has an effect on MB;
- AB-usage history is not associated with any of the confounders of the microbiome-disease association;
- There is no direct effect of AB history on incident disease outcomes outside of the pathway via microbiome.

Possible violations of these assumptions are described in the Results section.

Sensitivity analyses

In the main analysis, we set the maximum number of antibiotics prescribed during the period of interest to 5, minimum age to 23 and maximum age to 50. The lower limit for age refers to the minimum age in the EstMB cohort. Subjects with impaired health might be more prone to comorbidities and higher consumption of antibiotics. Thus, the aim of the initial restriction of the age range and number of antibiotics used is to focus on the healthier part of the population to address potential confounding. However, we pursued several sensitivity analyses to assess the validity and robustness of the method:

1. To account for the feedback-mechanism of the disease whereby the disease might actually already be prevalent for a period of time prior to diagnosis and could thus already have an effect on the MB, we pursued a sensitivity analysis where we considered the first 5 years of incidence after the pseudo-randomization as prevalent **(Supplementary Table 8)**. With such sensitivity analysis we aim to eliminate bias originating from the feedback mechanism of a disease that was already developing and affecting MB prior to the diagnosis;
2. Since the logarithm of the *Prevotella-Bacteroides* (P/B) ratio is with a bimodal distribution, we additionally performed an analysis where we applied inverse normal transformation on the P/B ratio prior to analysis using the *RNOmni* R package (McCaw *et al.*, 2020). **(Supplementary Table 4)**
3. We assessed the performance of the method in the setting, where we expect there to be no association. For that we created pseudo-variables indicating disease presence/absence for each individual in the EstBB sample by randomly sampling from binomial distribution with predefined probability. In total we created 39 pseudo-variables with disease-probabilities ranging from 2.5% to 97.5% (**Supplementary Table 9)**. We assumed the P/B ratio to be significantly associated with less than 5% of the generated variables.
4. We assessed the relationships between MB and diseases while using specific AB subgroups as instruments instead of total AB-usage. More specifically, we defined three additional variables as instruments based on the usage of AB with following ATC-

codes: J01FA (macrolides) **(Supplementary Table 5)**, J01MA (fluoroquinolones) **(Supplementary Table 6)**, and J01CR (penicillins) **(Supplementary Table 7)**. These antibiotics classes were chosen since they have the strongest long-term effects on gut microbiome (Aasmets *et al., unpublished data*). All three variables were uncorrelated with each-other (all Spearman correlations in EstMB and EstBB between the three AB-subgroups below 0.2 (**Fig. 2C**)), thus representing independent instruments. We assumed to see in general similar patterns of associations between MB and diseases, since a strong causal effect should not depend on the instrument used, provided that the instrument is valid. For this sub-analysis we excluded all the individuals who were prescribed any AB irregardless of the ATC-code during 6 months prior to pseudo-randomization. These analyses were done in two settings: maximum number of AB=5, age 23-50; and maximum number of AB=10, age 23-89.

Data availability

The metagenomic data generated in this study have been deposited in the European Genome-Phenome Archive database (https://www.ebi.ac.uk/ega/) under accession code EGAS00001008448. The phenotype data contain sensitive information from healthcare registers and they are available under restricted access through the Estonian biobank upon submission of a research plan and signing a data transfer agreement. All data access to the Estonian Biobank must follow the informed consent regulations of the Estonian Committee on Bioethics and Human Research, which are clearly described in the Data Access section at https://genomics.ut.ee/en/content/estonian-biobank. A preliminary request for raw metagenome and phenotype data must first be submitted via the email address releases@ut.ee

All participants included in the EstBB cohort provided informed consent for the data and samples to be used for scientific purposes. This study was approved by the Research Ethics Committee of the University of Tartu (approval No. 266/T10) and by the Estonian Committee on Bioethics and Human Research (Estonian Ministry of Social Affairs; approval No. 1.1-12/17 and 1.1-12/624). All participants have joined the Estonian Biobank on a voluntary basis and have signed a broad consent form, which allows to receive participant's personal and health data from national registries and databases. Rights of gene donors are regulated by Human Genes Research Act (HGRA) § 9 – Voluntary nature of gene donation (https://www.riigiteataja.ee/en/eli/ee/531102013003/consolide/ current).

Acknowledgments

The authors would like to thank Mari-Liis Tammeorg, Marili Palover, Anu Reigo, Neeme Tõnisson, Liis Leitsalu, and Esta Pintsaar for participating in the sample collection process of the Estonian Microbiome cohort. We thank Steven Smit, Rita Kreevan, Martin Tootsi, and Triinu Temberg for the DNA extraction process. We thank Reidar Andreson for bioinformatic support. We also thank all the EstBB study participants.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

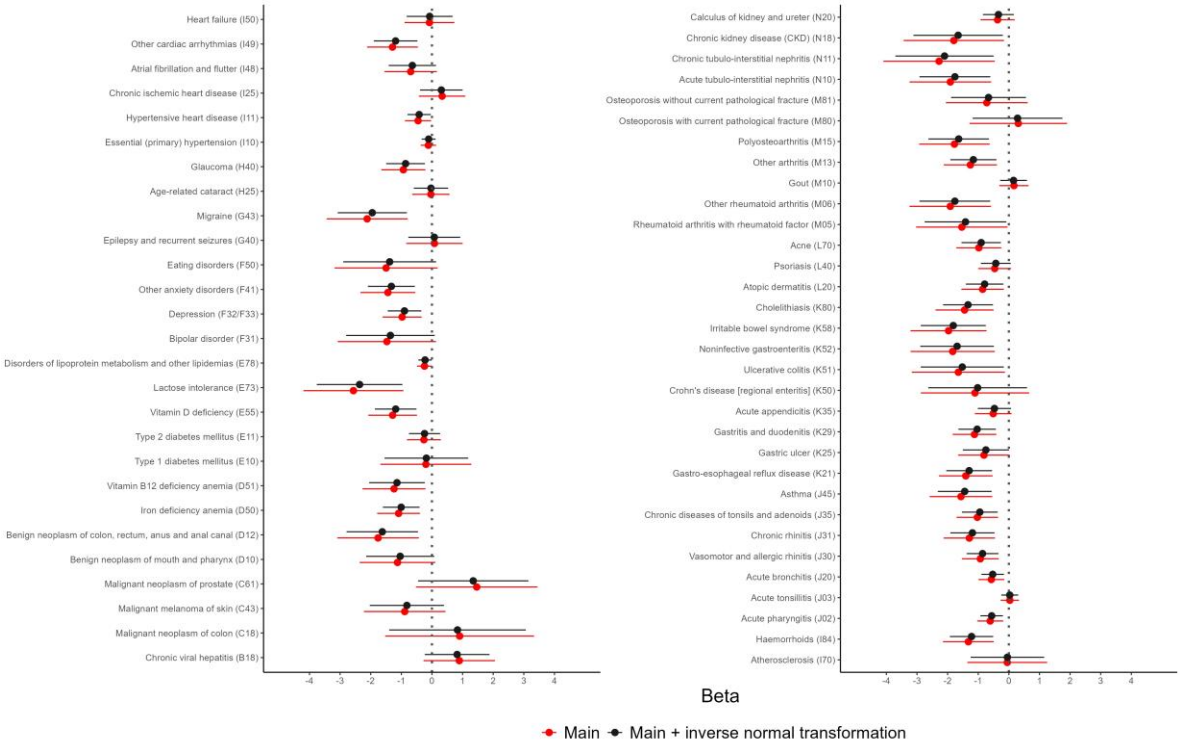
Author Contributions

O.A., E.O. and K.F. conceptualized and supervised the study. O.A and N.T. designed the study and performed the data analysis. N.T. and O.A. interpreted the data and prepared the figures and wrote the manuscript. All authors read and approved the final paper.

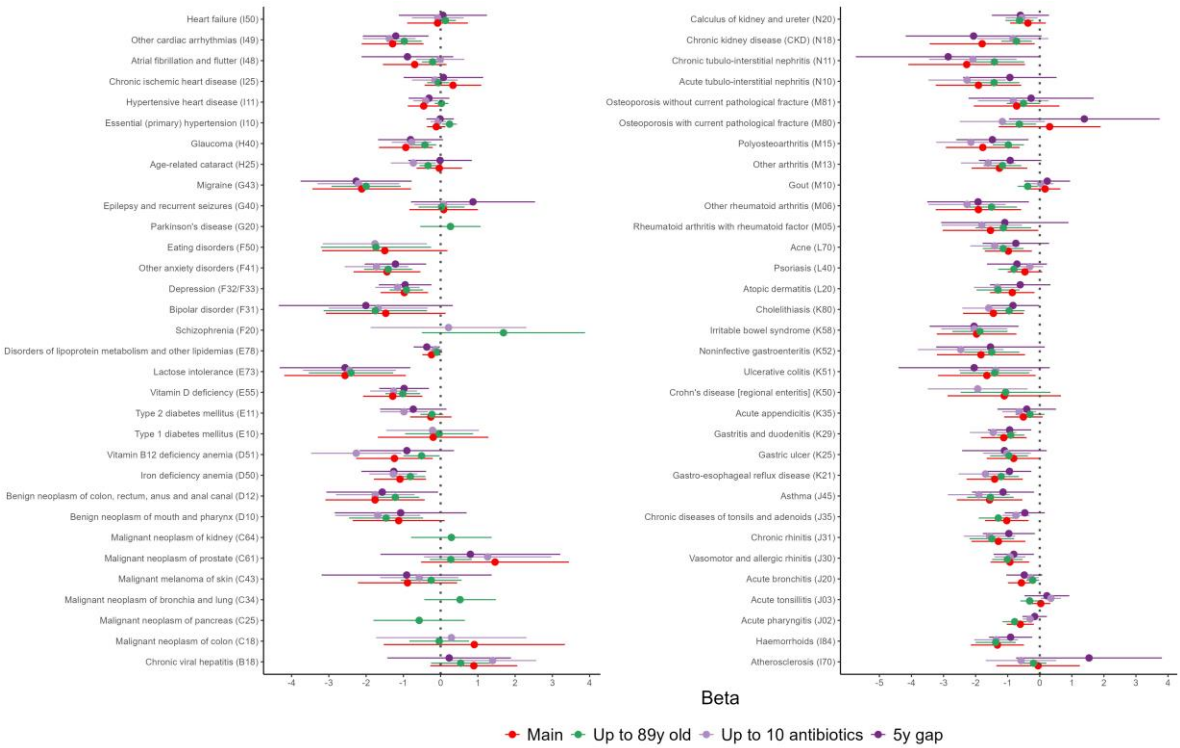
Funding

This work was funded by Estonian Research Council grants (PRG1414 to E.O., N.T. and O.A. and PRG1197 to K.F.) and an EMBO Installation grant (No. 3573 to E.O., N.T. and O.A.).

Supplementary figures

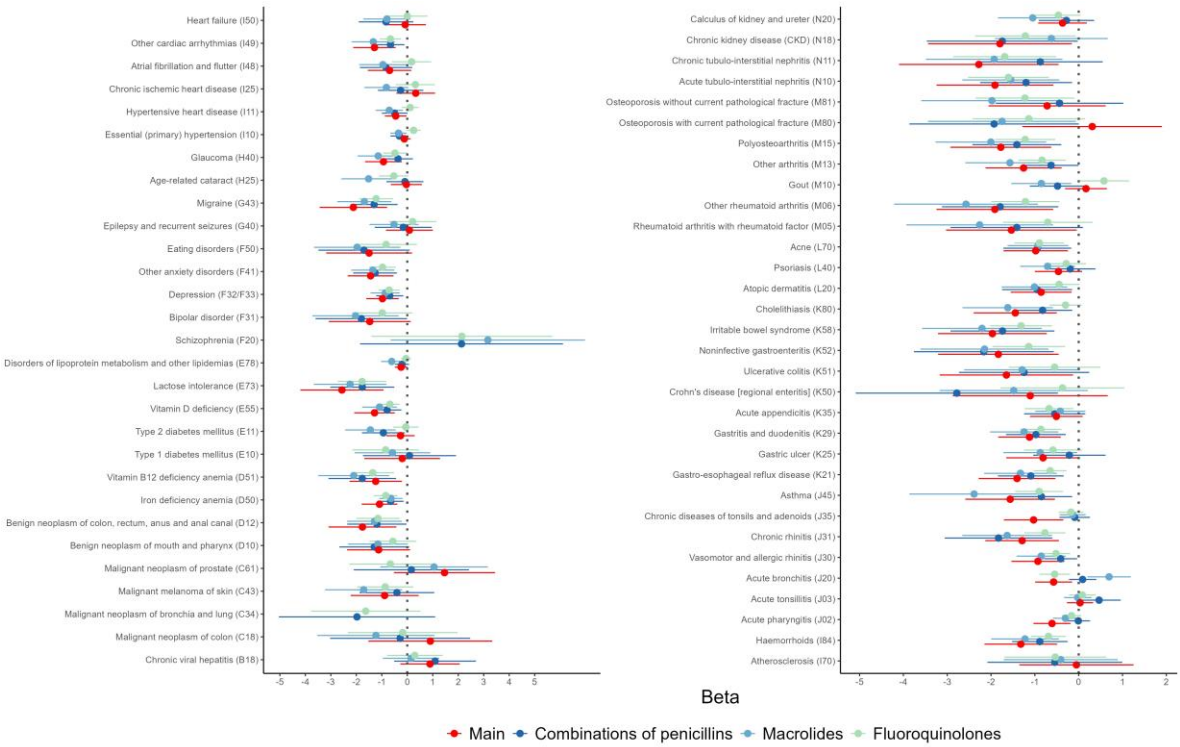


Supplementary figure 1. Results of the main analysis for all the 59 studied diseases. Colours represent different data transformations for the *Prevotella/Bacteroides* ratio. Sensitivity analysis results with the inverse normal transformation is shown in black.



Supplementary figure 2. Results of the sensitivity analysis analysis for all the 64 studied diseases. Green shows the effect estimates, when age is filtered as 23-89; light-purple shows the effect estimates, when the maximum number of AB prescribed is 10; dark-purple describes the scenario

where the first five years of incidence after the pseudo-randomization is considered as prevalent disease.



Supplementary figure 3. Results of the sensitivity analysis for all the 64 studied diseases. Information regarding subclasses of AB were used instead of the total amount of AB prescribed, whereas other settings were identical to the main analysis**.** Red corresponds to the main analysis with the antibiotics combined, dark blue corresponds to the class of penicillins (J01CR), light-blue corresponds to the class of macrolides (J01FA) and light-green corresponds to the class of fluoroquinolones (J01MA).

References

1. Aasmets, O. *et al.* (2021) 'Machine Learning Reveals Time-Varying Microbial Predictors with Complex Effects on Glucose Regulation', *mSystems*. Edited by M. Arumugam, 6(1), pp. e01191-20. Available at: https://doi.org/10.1128/mSystems.01191-20.
2. Aasmets, O. *et al.* (2022) 'Gut metagenome associations with extensive digital health data in a volunteer-based Estonian microbiome cohort', *Nature Communications*, 13(1), p. 869. Available at: https://doi.org/10.1038/s41467-022-28464-9.
3. Anthony, W.E. *et al.* (2022) 'Acute and persistent effects of commonly used antibiotics on the gut microbiome and resistome in healthy adults', *Cell Reports*, 39(2), p. 110649. Available at: https://doi.org/10.1016/j.celrep.2022.110649.
4. Beghini, F. *et al.* (2021) 'Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with bioBakery 3', *eLife*, 10, p. e65088. Available at: https://doi.org/10.7554/eLife.65088.
5. Bruyndonckx, R. *et al.* (2021) 'Consumption of antibiotics in the community, European Union/European Economic Area, 1997–2017', *Journal of Antimicrobial Chemotherapy*, 76(Supplement_2), pp. ii7–ii13. Available at: https://doi.org/10.1093/jac/dkab172.

6. Burgess, S. and Thompson, S.G. (2015) *Mendelian randomization: methods for using genetic variants in causal estimation*. CRC Press.
7. Chen, M. *et al.* (2022) 'Gut microbiota and major depressive disorder: A bidirectional Mendelian randomization', *Journal of Affective Disorders*, 316, pp. 187–193. Available at: https://doi.org/10.1016/j.jad.2022.08.012.
8. Forslund, S.K. *et al.* (2021) 'Combinatorial, additive and dose-dependent drug–microbiome associations', *Nature*, 600(7889), pp. 500–505. Available at: https://doi.org/10.1038/s41586-021-04177-9.
9. Gurung, M. *et al.* (2020) 'Role of gut microbiota in type 2 diabetes pathophysiology', *EBioMedicine*, 51, p. 102590. Available at: https://doi.org/10.1016/j.ebiom.2019.11.051.
10. Kartal, E. *et al.* (2022) 'A faecal microbiota signature with high specificity for pancreatic cancer', *Gut*, 71(7), pp. 1359–1372. Available at: https://doi.org/10.1136/gutjnl-2021-324755.
11. Kurilshikov, A. *et al.* (2021) 'Large-scale association analyses identify host factors influencing human gut microbiome composition', *Nature Genetics*, 53(2), pp. 156–165. Available at: https://doi.org/10.1038/s41588-020-00763-1.
12. Leitsalu, L. *et al.* (2015) 'Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu', *International Journal of Epidemiology*, 44(4), pp. 1137–1147. Available at: https://doi.org/10.1093/ije/dyt268.
13. Liu, B. *et al.* (2022) 'Two-Sample Mendelian Randomization Analysis Investigates Causal Associations Between Gut Microbial Genera and Inflammatory Bowel Disease, and Specificity Causal Associations in Ulcerative Colitis or Crohn's Disease', *Frontiers in Immunology*, 13, p. 921546. Available at: https://doi.org/10.3389/fimmu.2022.921546.
14. Long, Y. *et al.* (2023) 'Causal relationship between gut microbiota and cancers: a two-sample Mendelian randomisation study', *BMC Medicine*, 21(1), p. 66. Available at: https://doi.org/10.1186/s12916-023-02761-6.
15. Lv, B.-M., Quan, Y. and Zhang, H.-Y. (2021) 'Causal Inference in Microbiome Medicine: Principles and Applications', *Trends in Microbiology*, 29(8), pp. 736–746. Available at: https://doi.org/10.1016/j.tim.2021.03.015.
16. Maier, L. *et al.* (2018) 'Extensive impact of non-antibiotic drugs on human gut bacteria', *Nature*, 555(7698), pp. 623–628. Available at: https://doi.org/10.1038/nature25979.
17. McCaw, Z.R. *et al.* (2020) 'Operating characteristics of the rank-based inverse normal transformation for quantitative trait analysis in genome-wide association studies', *Biometrics*, 76(4), pp. 1262–1272. Available at: https://doi.org/10.1111/biom.13214.
18. McGovern, A.S., Hamlin, A.S. and Winter, G. (2019) 'A review of the antimicrobial side of antidepressants and its putative implications on the gut microbiome', *Australian & New Zealand Journal of Psychiatry*, 53(12), pp. 1151–1166. Available at: https://doi.org/10.1177/0004867419877954.
19. Pacini, D. and Windmeijer, F. (2016) 'Robust inference for the Two-Sample 2SLS estimator', *Economics Letters*, 146, pp. 50–54. Available at: https://doi.org/10.1016/j.econlet.2016.06.033.
20. Pascal, V. *et al.* (2017) 'A microbial signature for Crohn's disease', *Gut*, 66(5), pp. 813–822. Available at: https://doi.org/10.1136/gutjnl-2016-313235.
21. Patangia, D.V. *et al.* (2022) 'Impact of antibiotics on the human microbiome and consequences for host health', *MicrobiologyOpen*, 11(1). Available at: https://doi.org/10.1002/mbo3.1260.
22. Ruuskanen, M.O. *et al.* (2022) 'Gut Microbiome Composition Is Predictive of Incident Type 2 Diabetes in a Population Cohort of 5,572 Finnish Adults', *Diabetes Care*, 45(4), pp. 811–818. Available at: https://doi.org/10.2337/dc21-2358.
23. Ser, H.-L. *et al.* (2021) 'The Use of Fecal Microbiome Transplant in Treating Human Diseases: Too Early for Poop?', *Frontiers in Microbiology*, 12, p. 519836. Available

at: https://doi.org/10.3389/fmicb.2021.519836.

24. Tarini Shankar Ghosh *et al.* (2020) ‘Mediterranean diet intervention alters the gut microbiome in older people reducing frailty and improving health status: the NU-AGE 1-year dietary intervention across five European countries’, *Gut*, 69(7), p. 1218. Available at: https://doi.org/10.1136/gutjnl-2019-319654.
25. Wade, K.H. and Hall, L.J. (2019) ‘Improving causality in microbiome research: can human genetic epidemiology help?’, *Wellcome Open Research*, 4, p. 199. Available at: https://doi.org/10.12688/wellcomeopenres.15628.3.
26. Wirbel, J. *et al.* (2019) ‘Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer’, *Nature Medicine*, 25(4), pp. 679–689. Available at: https://doi.org/10.1038/s41591-019-0406-6.