Supplementary Information (SI) for Food & Function.
This journal is © The Royal Society of Chemistry 2024

STROBE-MR checklist of recommended items to address in reports of Mendelian randomization studies[1]

Item No.	Section	Checklist item	Page No.	Relevant text from manuscript
1	**TITLE and ABSTRACT**	Indicate Mendelian randomization (MR) as the study's design in the title and/or the abstract if that is a main purpose of the study	P2	**TITLE:** Dietary Influences on Urinary Tract Infections: Unraveling the Gut Microbiota Connection **ABSTRACT:** We investigated the intricate links between dietary patterns, gut microbiota, and This study explores the causal links between dietary patterns, gut microbiota, and urinary tract infections ...
	INTRODUCTION			
2	**Background**	Explain the scientific background and rationale for the reported study. What is the exposure? Is a potential causal relationship between exposure and outcome plausible? Justify why MR is a helpful method to address the study question	P4-5	**What is the exposure?** UTIs typically result from the retrograde infection of bacteria from the fecal microbiota. Given that the composition of gut microbiota is substantially influenced by dietary choices, alterations in diet may potentially modulate the risk of UTIs. While the connection between diet and UTIs, as well as the relationship between gut microbiota and UTIs, has garnered considerable attention from researchers, there has been a lack of direct Mendelian randomization studies concerning the causal relationships between dietary factors, gut microbiota, and UTIs. **Is a potential causal relationship between exposure and outcome plausible?** Prior case-control studies have indicated that the consumption of cranberry products and fermented dairy containing beneficial probiotic strains, such as Lactobacillus, effectively reduces the risk of recurrent UTIs. Additionally, prospective research has explored the association between vegetarian diets and UTIs. **Justify why MR is a helpful method to address the study question** Establishing a link between diet and UTIs has been challenging. While randomized controlled trials (RCTs) have long been considered the gold standard for establishing causality in multicenter studies, most long-term nutritional epidemiological investigations rely on food frequency questionnaires to gauge food and nutrient consumption. Such an approach is susceptible to bias due to self-reported measurement errors. Furthermore, the widespread consumption of fortified foods and vitamin supplements adds complexity to assessing nutritional intake. Observational studies can introduce biases due to confounding factors and reverse causation, particularly when causal inferences are involved.
3	**Objectives**	State specific objectives clearly, including pre-specified causal hypotheses (if any). State that MR is a method that, under specific assumptions, intends to estimate causal effects	P4-5	MR is a method for assessing causal relationships between risk factors and diseases. It utilizes one or more genetic variants associated with the exposure of interest as instrumental variables to evaluate the association between the exposure and the outcome. As alleles are randomly allocated during conception, genetic variations remain unaffected by measurement biases or biases arising from reverse causation. Furthermore, MR offers a feasible approach for inferring correlations between specific dietary intake and diseases, utilizing the uniqueness of genotypes to investigate causal relationships between exposures and outcomes. It employs genetically correlated instrumental variables (IVs), closely related to the exposure, to mimic a randomized controlled

				setting. The MR design can mitigate potential residual confounding effects and counteract reverse causation biases. Leveraging MR-based research designs allows the investigation of exposures that cannot be subjected to randomized controlled trials. While the connection between diet and UTIs, as well as the relationship between gut microbiota and UTIs, has garnered considerable attention from researchers, there has been a lack of direct Mendelian randomization studies concerning the causal relationships between dietary factors, gut microbiota, and UTIs. Hence, we employ MR analysis to explore the associations between dietary factors and gut microbiota with UTIs. Through MR analysis, our objective is to identify the various associations between dietary habits and urinary tract infections and the impact of different gut microbiota on urinary tract infections. This research aims to provide novel insights into the prevention and treatment of urinary tract infections.
	METHODS			
4	**Study design and data sources**	Present key elements of the study design early in the article. Consider including a table listing sources of data for all phases of the study. For each data source contributing to the analysis, describe the following:		
		a) Setting: Describe the study design and the underlying population, if possible. Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection, when available.	P6	Regarding urinary tract infection biomarkers, C-Reactive protein levels were obtained from MR-Base . This study measured serum CRP in mg/L using standard laboratory techniques and transformed the values by natural log. Individuals with autoimmune diseases, those taking immune-modulating agents (if this information was available), and those with CRP values 4 SD or more from the mean were excluded from all analyses. Neutrophil gelatinase-associated lipocalin and Procalcitonin levels were sourced from the Genomic Atlas of the Human Plasma Proteome, Interleukin-1-beta levels from the Genome-wide Association Study Identifies 27 Loci Influencing Concentrations of Circulating Cytokines and Growth Factors , and Interleukin-6, Interleukin-8, and Myeloperoxidase levels from the Genomic and Drug Target Evaluation of 90 Cardiovascular Proteins in 30,931 Individuals. Elevated erythrocyte sedimentation rate and abnormality of plasma viscosity data were obtained from the FinnGen biobank analysis. For these specific inflammatory biomarkers, all datasets excluded individuals with recent major illnesses (e.g., myocardial infarction, stroke, cancer, HIV, hepatitis B or C) and recent infections, as detailed in the cited literature.
		b) Participants: Give the eligibility criteria, and the sources and methods of selection of participants. Report the sample size, and whether any power or sample size calculations were carried out prior to the main analysis	P6	（Figure1) & 2.2 Data Sources

		c) Describe measurement, quality control and selection of genetic variants	P7	We established a set of IVs for dietary factors and gut microbiota, selecting SNPs with a genome-wide significance threshold of p < 5 e-8. We also applied an exclusion criterion, removing SNPs with an r2 > 0.001 (using a clumping window of 10,000 kb) to account for linkage disequilibrium. To refine our SNP selection, we further filtered each SNP based on potential confounders related to age, weight, gender, ethnicity, as well as factors related to the outcome or potential confounding factors within the dietary factors using the PhenoScanner website (Kamat et al., 2019). Subsequently, we computed the F-statistic to assess the strength of each individual SNP as an instrument. SNPs with an F-statistic greater than 10 were considered sufficiently strong to mitigate potential weak instrument bias, while SNPs with an F-statistic below 10 were excluded from the analysis.
		d) For each exposure, outcome, and other relevant variables, describe methods of assessment and diagnostic criteria for diseases	P5	2.1 Study Design
		e) Provide details of ethics committee ap relevant	P7	Ethical approval was not required for the current analysis, as all incorporated GWAS data were publicly available and had already received approval from their respective ethics review boards.
5	**Assumptions**	Explicitly state the three core IV assumptions for the main analysis (relevance, independence and exclusion restriction) as well assumptions for any additional or sensitivity analysis	P7	We established a set of IVs for dietary factors and gut microbiota, selecting SNPs with a genome-wide significance threshold of p < 5 e-8. We also applied an exclusion criterion, removing SNPs with an r2 > 0.001 (using a clumping window of 10,000 kb) to account for linkage disequilibrium. To refine our SNP selection, we further filtered each SNP based on potential confounders related to age, weight, gender, ethnicity, as well as factors related to the outcome or potential confounding factors within the dietary factors using the PhenoScanner website. Subsequently, we computed the F-statistic to assess the strength of each individual SNP as an instrument. SNPs with an F-statistic greater than 10 were considered sufficiently strong to mitigate potential weak instrument bias, while SNPs with an F-statistic below 10 were excluded from the analysis.
6	**Statistical methods: main analysis**	Describe statistical methods and statistics used		

a) **Describe how quantitative variables were handled in the analyses (i.e., scale, units, model)**
Quantitative variables in the analyses were handled through various statistical models and methods. For example, the inverse variance weighted method was primarily used to calculate causal effects. The odds ratios and their 95% confidence intervals were used to quantify the relationships between dietary habits, gut microbiota, and urinary tract infections. The F-statistic was used to assess the strength of each individual single nucleotide polymorphism as an instrument.

		b)	**Describe how genetic variants were handled in the analyses and, if applicable, how their weights were selected** Genetic variants, specifically SNPs, were used as instrumental variables in the analyses. SNPs associated with specific dietary habits were selected from the UK Biobank and FinnGen studies. SNPs were chosen based on a genome-wide significance threshold of $p < 5\ e^{-8}$ and linkage disequilibrium was accounted for by removing SNPs with an $r^2 > 0.001$ (using a clumping window of 10,000 kb). The strength of each SNP as an instrument was assessed using the F-statistic, with SNPs having an F-statistic greater than 10 considered sufficiently strong to mitigate potential weak instrument bias
		c)	**Describe the MR estimator (e.g. two-stage least squares, Wald ratio) and related statistics. Detail the included covariates and, in case of two-sample MR, whether the same covariate set was used for adjustment in the two samples** The Mendelian randomization (MR) estimator used in this study was the IVW method, which is often used in two-sample MR analyses to detect causal relationships. Sensitivity analyses were conducted using the Weighted Median and MR-Egger methods. The Cochrane's Q statistic was used to detect heterogeneity among the chosen IVs. If heterogeneity was present, a random-effects model was adopted; otherwise, a fixed-effects model was used. And each SNP was filtered based on potential confounders related to age, weight, gender, ethnicity, as well as factors related to the outcome or potential confounding factors within the dietary factors using the PhenoScanner website.
		d)	**Explain how missing data were addressed** As the study utilized publicly available data from the UK Biobank, the FinnGen study, and previously published Genome-Wide Association Studies, it is that any missing data issues would have been addressed in the original data collection and analysis processes.
		e)	If applicable, indicate how multiple testing was addressed
7	**Assessment of assumptions**		**Describe any methods or prior knowledge used to assess the assumptions or justify their validity** In Materials and Methods section
8	**Sensitivity analyses and additional analyses**		**Describe any sensitivity analyses or additional analyses performed (e.g. comparison of effect estimates from different approaches, independent replication, bias analytic techniques, validation of instruments, simulations)** In Materials and Methods section
9	**Software and preregistration**		
		a)	**Name statistical software and package(s), including version and settings used** In Materials and Methods section

		b)	**State whether the study protocol and details were pre-registered (as well as whennd where)** Not applicable to this study
	RESULTS		
10	**Descriptive data**		
		a)	**Report the numbers of individuals at each stage of included studies and reasons for exclusion. Consider use of a flow diagram** In Materials and Methods section
		b)	**Report summary statistics for phenotypic exposure(s), outcome(s), and other relevant variables (e.g. means, SDs, proportions)** In supplementary chart & Figure2-4
		c)	**If the data sources include meta-analyses of previous studies, provide the assessments of heterogeneity across these studies** In supplementary chart
		d)	**For two-sample MR:** **i. Provide justification of the similarity of the genetic variant-exposure associations between the exposure and outcome samples** In supplementary chart **ii. Provide information on the number of individuals who overlap between the exposure and outcome studies** In supplementary chart
11	**Main results**		
		a)	**Report the associations between genetic variant and exposure, and between genetic variant and outcome, preferably on an interpretable scale** In Results section.
		b)	**Report MR estimates of the relationship between exposure and outcome, and the measures of uncertainty from the MR analysis, on an interpretable scale, such as odds ratio or relative risk per SD difference** In Results section.
		c)	**If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period** Not applicable to this study
		d)	**Consider plots to visualize results (e.g. forest plot, scatterplot of associations between genetic variants and outcome versus between genetic variants and exposure)** In supplementary Figure
12	**Assessment of assumptions**		

		a)	**Report the assessment of the validity of the assumptions** We reported strong associations between the selected genetic variants and the dietary exposure, with all SNPs reaching genome-wide significance ($p < 5e^{-8}$), thereby satisfying the first assumption. The selected SNPs were not associated with potential confounders such as age, sex, or other lifestyle factors, as evidenced by our analyses. This supports the second assumption. We used MR-Egger regression to test the third assumption, which can detect horizontal pleiotropy (when a genetic variant affects the outcome through pathways other than the exposure). The intercept term from the MR-Egger regression was not significantly different from zero, suggesting no evidence of horizontal pleiotropy. This supports the third assumption. Additionally, we performed a leave-one-out analysis to check for the influence of individual SNPs on the overall estimate. The MR estimates remained consistent when each SNP was excluded in turn, suggesting no undue influence from any particular SNP. We also used the Cochrane's Q statistic to test for heterogeneity among the IVs. The results showed no significant heterogeneity, suggesting that the IVs were consistent in their effects.
		b)	**Report any additional statistics (e.g., assessments of heterogeneity across genetic variants, such as I^2, Q statistic or E-value)** In supplementary chart
13	**Sensitivity analyses and additional analyses**		
		a)	**Report any sensitivity analyses to assess the robustness of the main results to violations of the assumptions** In supplementary chart
			Report results from other sensitivity analyses or additional analyses
		b)	We also performed a multivariable MR analysis to adjust for potential pleiotropic effects of the genetic variants. The results remained largely unchanged, providing further evidence for the robustness of the findings.
		c)	**Report any assessment of direction of causal relationship (e.g., bidirectional MR)** Given the study's design and the clinical context aimed at addressing urinary tract infections, we opted for a unidirectional Mendelian randomization analysis. This choice was made to focus on elucidating a specific direction of causal relationship in accordance with the research objectives and the nature of the investigation
		d)	**When relevant, report and compare with estimates from non-MR analyses** In our discussion section, we address the reasons behind similarities or disparities with estimates from traditional observational studies. The Mendelian randomization (MR) estimates in our study were slightly lower, potentially attributed to the intricate confounding factors inherent in dietary habits. This observation serves as a starting point for further research and underscores the complexity of the relationships under investigation.
		e)	Consider additional plots to visualize results (e.g., leave-one-out analyses) In supplementary Figure
	DISCUSSION		
14	**Key results**		Summarize key results with reference to study objectives
15			**Limitations Discuss limitations of the study, taking into account the validity of the IV assumptions, other sources of potential bias, and imprecision. Discuss both direction and magnitude of any potential bias and any efforts to address them** In limitation section

16	**Interpretation**	
	a) **Meaning: Give a cautious overall interpretation of results in the context of their limitations and in comparison with other studies**	p9-11 Discussion & Limitations
	b) **Mechanism: Discuss underlying biological mechanisms that could drive a potential causal relationship between the investigated exposure and the outcome, and whether the gene-environment equivalence assumption is reasonable. Use causal language carefully, clarifying that IV estimates may provide causal effects only under certain assumptions**	P10-11 In this MR analysis, we assessed the causal relationships between dietary intake and gut microbiota species with susceptibility to UTIs. We observed potential protective effects of oily fish and cheese intake against UTIs, while high alcohol intake frequency appeared to increase the risk of UTIs. Furthermore, we identified 15 associations between gut microbiota and UTIs. Numerous studies have confirmed that alcohol impairs the functionality of various components of the immune defense system. Prolonged alcohol consumption disrupts various aspects of acquired immune responses, including cell-mediated and humoral responses, rendering individuals more susceptible to viral and bacterial infections, as well as sterile inflammation. Existing research suggests that alcohol can alter the balance and interactions between the host immune system and the host microbiota to influence immune function. Alcohol consumption disrupts the gut barrier by increasing oxidative stress in the gut, leading to increased gut permeability and dysbiosis. This disruption can eventually result in the translocation of Gram-negative bacterial products, and the observed increase in CRP levels in our results aligns well with these findings. Moreover, as previously reported , alcohol consumption can lead to a decrease in the abundance of Bacteroidetes phyla (synonym Bacteroidota) bacteria and an increase in Firmicutes phylum (synonym Bacillota), specifically the Bacilli class. In our study, the protective effect of the family Prevotellaceae, which belongs to the Bacteroidetes phylum (Bacteroidota), against UTIs, and the increased risk

associated with genera Anaerotruncus and Dorea (both of Firmicutes phylum: Bacilli class), corroborate this perspective. Therefore, the elevated frequency of alcohol consumption may enhance susceptibility to UTIs by affecting the immune system and altering the gut microbiota composition.

Oily fish exhibit a potential protective effect against UTIs. In comparison to white fish, oily fish are rich in omega-3 polyunsaturated fatty acids (PUFAs), specifically eicosapentaenoic acid (EPA) and docosahexaenoic acid, both of which belong to the ω-3 class. The biochemical and physiological actions of these PUFAs largely depend on their conversion to 20-carbon or 22-carbon acids, followed by subsequent metabolism into bioactive lipid mediators such as prostaglandins, leukotrienes, lipoxins, and resolvins. Research indicates that omega-3 PUFAs have mitigating effects on various inflammation-related diseases. They achieve this by attenuating the activation of MAPK, NF-κB, activator protein-1, and oxidative stress pathways, or by enhancing the activation of PPARγ or GPR120, thus diminishing inflammation .

Omega-3 PUFAs can act as alternative substrates for cyclooxygenase or lipoxygenase, impeding the conversion of arachidonic acid into pro-inflammatory eicosanoids, thereby reducing the production of inflammatory mediators. The reduction in IL-6 levels associated with oil fish intake in our findings further underscores this perspective. Research suggests that omega-3 PUFAs from fish oil can reduce the abundance of Firmicutes in animal models, corroborating our results . In terms of maintaining intestinal epithelial integrity, PUFAs influence gut inflammation either as precursors for anti-inflammatory eicosanoids synthesis or by regulating tight junction functionality to enhance gut integrity . When intestinal barrier function is compromised, resulting in heightened gut permeability and increased lipopolysaccharide translocation, the subsequent postprandial endotoxemia results in mild systemic inflammation.

Omega-3 PUFAs can alter the microbiota composition by inducing the production and secretion of intestinal alkaline phosphatase, thereby reducing the abundance of LPS-producing bacteria. The decrease in CRP in our results also supports this notion.

In addition to being rich in saturated fatty acids, cheese, as a dairy product, contains a diverse array of constituents including vitamins D, calcium, protein,

probiotics, and bioactive peptides. Research indicates a close correlation between the Dietary Approaches to Stop Hypertension diet pattern (inclusive of dairy products) and reduced inflammation. Studies in mouse models have revealed that dietary calcium, found in cheese, reduces the expression of inflammatory cytokines in adipocytes by inhibiting the formation of calcitriol. Dairy-derived proteins such as lactoferrin may also exhibit anti-inflammatory effects through the regulation of cytokine release, immune cell recruitment, and activation. Thus, while saturated fats might potentially elevate inflammation levels, dairy protein could exert a neutral or even beneficial impact on inflammation. Moreover, dairy products can modulate immune function within the gastrointestinal tract by interacting with the mucosal layer, enhancing gut barrier function, and stimulating immune cells. Cheese, categorized as a probiotic food, harbors an abundance of live microorganisms. Existing research indicates that certain specific bacterial strains possess the ability to inhibit inflammatory responses within the body.

Clinical trials have demonstrated the beneficial effects of probiotic members, such as Lactobacillus and Bifidobacterium, which are prominent within cheese, on immunity and inflammation. Adequate cheese consumption can regulate the gut microbiota and confer beneficial effects on the host. A study has shown that supplementation with probiotics in cheese alleviates symptoms and prevents recurrent infections in UTI patients, aligning with our finding of a negative correlation between cheese intake and infection indicators , supporting the conclusion that cheese consumption has a protective effect against UTIs.

Dietary factors play a multifaceted role in influencing UTIs. Our focus has centered on investigating the impact of dietary factors on UTI biomarkers and the interplay between gut microbiota and UTIs. The human gut, harboring an estimated 100 trillion microorganisms, stands as one of the most densely colonized organs, endowing the host with a multitude of functionalities. Since the emergence of the gut-kidney axis concept proposed by Meijers in 2011, numerous clinical and animal model studies have affirmed the correlation between gut microbiota and various diseases. Within the framework of the gut-kidney axis theory, the bidirectional communication between gut microbiota and kidney

		health has been emphasized. It has been established that the gut microbiota influences renal function . An increasing body of research is focusing on this area, and prior observational studies have indeed confirmed a link between a high relative abundance of Romboutsia and a reduced risk of UTIs. Our groundbreaking Mendelian randomization study sheds light on the relationship between dietary factors, gut microbiota, and urinary tract infections. As we continue to refine our understanding of nutrient intake and the gut microbiota, the concept of the gut-kidney axis can better help us comprehend its significance in the context of UTIs. Effectively reducing the incidence and clinical burden of UTIs through dietary modifications offers promising avenues, particularly in the prevention of infections in individuals with complex UTIs, such as those with congenital anomalies, diabetes-related UTIs, and, notably, kidney transplant recipients.
		c) **Clinical relevance: Discuss whether the results have clinical or public policy relevance, and to what extent they inform effect sizes of possible interventions** P11-12 Dietary factors play a multifaceted role in influencing UTIs. Our focus has centered on investigating the impact of dietary factors on UTI biomarkers and the interplay between gut microbiota and UTIs. The human gut, harboring an estimated 100 trillion microorganisms, stands as one of the most densely colonized organs, endowing the host with a multitude of functionalities . Since the emergence of the gut-kidney axis concept proposed by Meijers in 2011, numerous clinical and animal model studies have affirmed the correlation between gut microbiota and various diseases. Within the framework of the gut-kidney axis theory, the bidirectional communication between gut microbiota and kidney health has been emphasized. It has been established that the gut microbiota influences renal function. An increasing body of research is focusing on this area, and prior observational studies have indeed confirmed a link between a high relative abundance of Romboutsia and a reduced risk of UTIs. Our groundbreaking Mendelian randomization study sheds light on the relationship between dietary factors, gut microbiota, and urinary tract infections. As we continue to refine our understanding of nutrient intake and the gut microbiota, the concept of the gut-kidney axis can better help us comprehend its significance in the context of UTIs. Effectively reducing the incidence and clinical burden of UTIs through dietary modifications offers promising avenues, particularly in the prevention of infections in individuals with complex UTIs, such as those with congenital anomalies, diabetes-related UTIs, and, notably, kidney transplant recipients.
17	**Generalizability**	**Discuss the generalizability of the study results (a) to other populations, (b) across other exposure periods/timings, and (c) across other levels of exposure** P12 Several limitations must be acknowledged in the context of our study. Firstly, our investigation was confined to the European population, thereby potentially constraining the external validity of our findings to other ethnic cohorts. Secondly, the classification of gut

		microbiota at the genus level limited our ability to elucidate the causal relationship between specific microbial species and urinary tract infections. Additionally, the intricate interplay of dietary factors in the context of urinary tract infections is manifold. Despite our emphasis on biomarkers and gut microbiota, these facets may not comprehensively encapsulate the entirety of potential effects. And, real-world research is susceptible to unidentified or unmeasured confounding factors. While these factors could potentially introduce variability in the results, the absence of pertinent data precludes their thorough control. Lastly, the primary objective of this article may not be to immediately advocate for UTI patients to consume specific foods. Rather, it aims to provide a novel perspective on the treatment of complex urinary tract infection cases, offering insights that could serve as precise references and guidance for future research endeavors.
	OTHER INFORMATION	
18	**Funding**	**Describe sources of funding and the role of funders in the present study and, if applicable, sources of funding for the databases and original study or studies on which the present study is based** P12 This work was supported by the National Natural Science Foundation of China (No. 81972373), and the Medical Leading Talents of Xiamen City.
19	**Data and data sharing**	**Provide the data used to perform all analyses or report where and how the data can be accessed, and reference these sources in the article. Provide the statistical code needed to reproduce the results in the article, or report whether the code is publicly accessible and if so, where** p13 The data used in this study are all derived from publicly available academic journals, books, or other published sources. All cited data and references can be accessed through their respective publications or databases.
20	**Conflicts of Interest**	**All authors should declare all potential conflicts of interest** P13 The authors declare no conflict of interest.

This checklist is copyrighted by the Equator Network under the Creative Commons Attribution 3.0 Unported (CC BY 3.0) license.

1. Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the Reporting of Observational Studies in Epidemiology using Mendelian Randomisation (STROBE-MR): Explanation and Elaboration. BMJ. 2021;375:n2233.