HYPOTHESIS-DRIVEN MEDIATION ANALYSIS FOR COMPOSITIONAL DATA: AN APPLICATION TO GUT MICROBIOME

Noora Kartiosuo[1,2,3], **Jaakko Nevalainen**[4], **Olli Raitakari**[2,3,5], **Katja Pahkala**[2,3,6], **Kari Auranen**[1,7]

1 Department of Mathematics and Statistics, University of Turku, Turku, Finland
2 Research Centre of Applied and Preventive Cardiovascular Medicine, University of Turku, Turku, Finland
3 Centre for Population Health Research, University of Turku, Turku, Finland
4 Health Sciences Unit, Faculty of Social Sciences, Tampere University, Tampere, Finland
5 Department of Clinical Physiology and Nuclear Medicine, University of Turku and Turku University Hospital, Turku, Finland
6 Paavo Nurmi Centre & Unit for Health and Physical Activity, University of Turku, Turku, Finland
7 Department of Clinical Medicine, University of Turku, Turku, Finland

Corresponding author: Noora Kartiosuo[1,2,3]. e-mail: nsekar@utu.fi. Address: 20014 University of Turku, Turku, Finland.

ABSTRACT

Sequencing read-count data often exhibit sparsity (zero-count inflation) and overdispersion. As most sequencing techniques provide an arbitrary total count, taxon-specific counts should be treated under the compositional data-analytic framework. There is increasing interest in the role of gut microbiome composition in mediating the effects of exposures on health. Previous approaches to compositional mediation have focused on identifying mediating taxa among a number of candidates. We here consider compositional causal mediation when *a priori* knowledge is available about the hierarchy for a restricted number of taxa, building on a single hypothesis structured as contrasts between appropriate sub-compositions. Based on the assumed causal graph and the theory of multiple contemporaneous mediators, we define non-parametric estimands for the overall and coordinate-wise mediation effects and show how these indirect effects are estimated based on simple parametric linear models. The mediators have straightforward and coherent interpretations, related to causal questions about interrelationships between the sub-compositions. We perform a simulation study focusing on the impact of sparsity on estimation. While unbiased, the precision of estimators depends on sparsity and the relative magnitudes of exposure-to-mediator and mediator-to-outcome effects in a complex manner. In the empirical application we find an inverse association of fibre intake on insulin level, mainly attributable to the direct effects.

***K*eywords** Causal inference; Compositional data analysis; Isometric logratio transformation; Microbiota; Multiple mediators

1 Introduction

With improved sequencing methods, biological measurements consisting of read counts, such as those recording the bacterial makeup of the gut flora, have become increasingly common. Such data consist of counts of multiple and sometimes thousands of different taxa and often exhibit sparsity (excess of zero counts) as well as overdispersion (extra-Poisson variability). Most of the current sequencing techniques provide data with an arbitrary total count. As all information carried by the taxon-specific counts relates to their proportion of the total, the observed data should be

ideally interpreted as a composition where a change in the abundance of one part inevitably changes the abundances of the other parts of the composition [1].

Approaches based on compositional data analysis have been endorsed for analysing microbiome data [1,2]. In this framework, the compositionality of observations is accounted for by scaling taxon-specific read counts by their total into proportions of the whole [3,4]. Due to the resulting unit-sum constraint, data based on proportions cannot be validly analysed with standard statistical methods. Their use is enabled, however, by transforming count-based data into logratios, which contrast distinct parts of the composition against each other [4]. A common choice is the isometric logratio transformation (ilr), where the ratios are based on pre-specified contrasts [4]. Unlike proportions, the resulting ilr coordinates can be presented in the Euclidean space and have been shown to be asymptotically multivariate normally distributed in the case of multinomial counts [5]. It has been suggested that appropppriate contrasts in the analysis of microbiome data can be based on prior knowledge of the bacterial taxonomy or phylogenetic systematics [2], results from data-driven approaches [6], or a specific interest in a particular taxon [7]. Such *a priori* knowledge may also be employed to restrict the analysis to concern only a smaller-dimensional subset of the taxa.

There is an increasing interest in the role of the gut microbiome composition in mediating the effects of different exposures on health outcomes. Mediation analysis in general aims at discovering and quantifying causal pathways between an exposure (X) and a response (Y), and the extent to which their association is mediated by an intermediate variable (mediator M) [8,9]. Figure 1 (a) presents a directed acyclic graph (DAG) depicting a simple model with one mediator. The exposure affects the response both directly via pathway (i) and indirectly through the mediator via pathways (ii) and (iii). Figure 1 (b) presents a DAG with a compositional mediator, such as the microbiome, where two different types of indirect effects may be of interest: the mediating effect through the entire composition (light arrows), and the mediating effects through the individual components (e.g. arrows (ii.1) and (iii.1)).

Different data-driven methods have been proposed to assess the mediating role of the gut microbiome and to discover specific mediating taxa [7,10,11,12]. Sohn and Li [10] derived expressions for taxon-specific as well as microbiome-wise indirect effects. Their approach was based on contrasting individual taxa against a chosen reference taxon and compositional algebra constructed in the simplex space, thus precluding the use of orthonormal (Euclidean) coordinates. Wang et al. [11] developed a similar yet regularised model in the simplex space, further accounting for the sparsity of the read counts using Dirichlet regression. Zhang et al. used the ilr transformation with contrasts of a single component against the other parts of the composition [7]. While the ilr transformation allows interpreting the coordinates as log-ratios of the distinct parts of the composition, the authors only considered the mediating effect of each component at a time, while other components were considered as nuisance and their effects were regularised using a penalisation criterion. In a somewhat similar approach, Fu et al. [12] utilised the ilr transformation within a joint model for the response and the microbial composition to find mediating taxa and to estimate indirect effects mediated through their relative abundances.

Although data-driven methods are useful when the aim is to search candidates for particular mediating taxa, incorporation of available biological knowledge in the assessment of mediators has been advocated, especially if the aim is to interpret mediation through a causal lens [13]. Instead of applying a sequence of large number of different and often regularised models, the analysis should then be based on a pre-specified causal structure. The presence of multiple, possibly correlated mediators then poses the key analytical problem. In our application to the microbiome, these mediators will be ilr coordinates.

In the presence of dependencies between multiple mediators, special attention is required when defining and estimating mediator-specific effects. Methods to deal with multiple, causally ordered mediators [14,15,16] or correlated yet not causally dependent mediators [17,18,19] have been presented before. VanderWeele and Vansteelandt demonstrated for correlated mediators that the indirect effects are different when estimated based on one multivariable mediation analysis compared to applying mediation analysis to single mediator at a time [14]. With a focus on causally dependent mediators, these authors suggested an incremental approach to define and estimate the additional contribution of each added mediator [14]. However, this approach does not enable obtaining mediator-specific indirect effects based on a single model of mediation.

Focusing on a binary response variable, Wang et al. derived a non-parametric decomposition of the overall mediated effect into coordinate-specific effects mediated via individual coordinates of a vector-valued mediator [17]. Their decomposition, based on a single causal model, allows interpreting the mediating effect of each coordinate in a coherent manner. In particular, the coordinate-wise indirect effects sum up to the overall indirect effect. Kim et al. presented an alternative approach, considering contributions attributable to each of the mediators as well as their combinations [18]. However, under their non-parametric formulation, the coherence between the overall and mediator-specific indirect effects does not hold.

In this article, we consider causal inference about the mediating role of the microbiome composition when sufficient knowledge is available to form appropriate constrasts between a small number of subcompositions. We consider the taxonomic structure of the microbiome as such *a priori* information although any other knowledge could be used to built the structure of causal influences. Similarly to Zhang et al. [7] and Fu et al. [12], we base our approach on Euclidean (ilr) coordinates but formulate the mediated effects in the counterfactual framework. Following the non-parametric definitions of indirect effects for multiple contemporaneous mediators as presented by Wang et al [17], we consider the overall as well as mediator-specific indirect effects for each of the ilr coordinates. Based on a single causal *a priori* hypothesis coded as a DAG, where each coordinate pertains to comparing the relative proportions of two subcompositions of interest, these effects are straightforward to interpret. This hypothesis-driven approach enables estimating all mediated effects in a single mediation analysis scheme, in contrast to investigating multiple mediation analyses with altering contrast matrices [7]. We then build a parametric model of mediation and specify conditions that allow estimating the mediated effects from empirical observations in terms of standard linear regression models. As opposed to previous research on microbial mediation, in our simulation study we demonstrate the estimation of the indirect effects in presence of sparsity and overdispersion, often present in data on microbiome, investigating their effect on the precision and power of estimation.

The structure of the paper is as follows. In Section 2, we review the construction of ilr coordinates based on a hypothesis-driven approach and then define the non-parametric estimands of overall and coordinate-wise indirect effects, specify the necessary conditions for their identifiability and present the parametric forms of the estimands. In Section 3, we first build a simulation study to illustrate the proposed method of mediation analysis and investigate the effect of certain features of the data, such as total count, sparsity and overdispersion. We then apply the proposed method to a question regarding the mediating role of the gut microbiome between fibre intake and insulin level in the Special Turku Coronary Risk Factor Intervention Project (STRIP) study. The discussion in Section 4 includes a comparison of the proposed approach with other analysis methods in compositional mediation regarding the microbiome.

2 Materials and methods

2.1 Compositional counts and ilr coordinates

Consider an observation of counts $K_j, j = 1, \ldots, J+1$, with $J+1$ distinct classes. Denote the sum of these counts, i.e., the total count, as K. As we are interested of the composition of the counts rather than the counts themselves, we scale the counts by their total into proportions $p_j = K_j/K, j = 1, \ldots, J+1$. The class-specific proportions are considered as parts (i.e. components) of a composition with the unit-sum constraint $\sum_{j=1}^{J+1} p_j = 1$. To be able to use standard statistical models, the proportions need to be transformed from the $(J+1)$-dimensional simplex to a J-dimensional Euclidean space. We accomplish this using the isometric logratio (ilr) transformation.

The first step in transforming proportions into ilr coordinates is to define an appropriate set of contrasts between the $J+1$ components. We here build on the idea of deriving the contrasts from prior biological knowledge and review how the ilr coordinates are calculated in practice [3,4].

The three-level dendrogram of Figure 2 (a) presents an example of an *a priori* known hierarchy, such as taxonomic knowledge, that can be used as the basis for defining contrasts between the five taxa using sequential binary partition (SBP) [3]. For example, we may build an $(J+1) \times J$ SBP matrix for the 5 classes using the following *balances* [20]

$$\mathbf{\Psi} = (\psi_{jk}) = \begin{bmatrix} +1 & 0 & 0 & 0 \\ -1 & +1 & +1 & 0 \\ -1 & +1 & -1 & 0 \\ -1 & -1 & 0 & +1 \\ -1 & -1 & 0 & -1 \end{bmatrix}, \tag{1}$$

where each column defines how the components indicated by +1 are contrasted against those indicated by -1. The first contrast compares taxon A against all other taxa. Taxonomic groups $\{B, C\}$ and $\{D, E\}$ are contrasted against each other (second column), whereas the taxa within both groups are contrasted against each other (B against C, third column; D against E, fourth column). Figure 2 (b) describes the hierarchy that is encoded by matrix 1. Whereas Figure 2 (a) describes the taxonomic hierarchy, Figure 2 (b) shows the hierarchy that is hypothesised to carry causally mediated effects. For example, even though the groups A, $\{B, C\}$ and $\{D, E\}$ are on the same taxonomic level, the first contrast in the matrix entails the assumption that the relative proportion of taxon A against the other groups carries the mediating effect.

Let n_k^+ and n_k^- be the numbers of cells with +1 and -1, respectively, in the kth column of the SBP matrix $\mathbf{\Psi}$, and let gm() denote the geometric mean. The ilr transformation of proportions $\boldsymbol{p}$ is defined as follows [3]:

$$\boldsymbol{M} = \text{ilr}(\boldsymbol{p}) = \mathbf{V}^T \log(\boldsymbol{p}), \tag{2}$$

where the elements of a $(J+1) \times J$ contrast matrix $\mathbf{V}$ are:

$$v_{jk} = \sqrt{\frac{n_k^+ n_k^-}{n_k^+ + n_k^-}} \left[\frac{1}{n_k^+}\right]^{\mathbf{I}[\text{sign}(\Psi_{jk})=+]} \left[\frac{1}{n_k^-}\right]^{\mathbf{I}[\text{sign}(\Psi_{jk})=-]}.$$

Thus, the ilr coordinates are

$$M_k = \sqrt{\frac{n_k^+ n_k^-}{n_k^+ + n_k^-}} \log \frac{\text{gm}(p_k^+)}{\text{gm}(p_k^-)}, k = 1, \ldots, J. \tag{3}$$

In summary, the ilr coordinates presented here correspond to a scaled log-ratio of geometric means of proportions indicated by the chosen contrast matrix. Whereas the original proportions p_j lie in the simplex, the ilr-transformed coordinates reside in the Euclidean space. Each ilr coordinate pertains to a log-ratio between two sets of components, entailing a specific hypothesis and carrying thus a straightforward interpretation. Depending on the empirical research question and other a priori knowledge at hand, the matrix could be built in a different manner as well. However, choosing another matrix with different contrasts would entail a different hypothesis with a different interpretation.

A special way of building balances is contrasting each of the components against the remaining components. Here, we call this approach *pivotal* [20]. The pivotal SBP matrix is

$$\mathbf{\Psi} = \begin{bmatrix} +1 & 0 & 0 & \ldots & 0 \\ -1 & +1 & 0 & \ldots & 0 \\ -1 & -1 & +1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & -1 & -1 & \vdots & +1 \\ -1 & -1 & -1 & \vdots & -1 \end{bmatrix}. \tag{4}$$

It follows from Equation (3) that this choice of contrasts leads to the following ilr coordinates [21]:

$$M_k = \sqrt{\frac{(J+1)-k}{(J+1)-k+1}} \log \frac{p_k}{\sqrt[(J+1)-k]{\prod_{h=k+1}^{(J+1)} p_h}}, k = 1, \ldots, J.$$

Having fixed a specific order of the $J+1$ classes, the pivotal SBP matrix $\mathbf{\Psi}$ leads to contrasting the first component against the remainder, then the second component against the remaining components excluding the first one, and so forth. Each class is thus used as a pivot against the remaining classes of the composition.

Some previous analyses of compositional mediation have applied pivot coordinates without a strict causal *a priori* structure. In practice, each of the potentially mediating taxa is then used as a first pivot in turn [7]. This means that a set of ilr coordinates is built altogether $J+1$ times, alternating the order of the components at each time. In any of the $J+1$ sets of coordinates, the mediating effect of only the first coordinate is of interest, with an interpretation as the relative proportion of a single taxon contrasted to the other taxa. Unlike the balances based on a clear SBP matrix derived based on *a priori* knowledge, pivot coordinates do not necessarily convey other meaningful hypotheses than contrasting the pivot element to the rest. To distinguish this approach from ours, we will refer to it as *alternating pivot* coordinates. If a single balance matrix can be built as as sequence of pivots, following the structure of matrix (4), we refer the analysis as based on pivot SBP coordinates.

2.2 Compositional mediation analysis

In this section, we first introduce the non-parametric causal quantities of interest. We then define the parametric models for the exposure, compositional mediator (i.e., the ilr coordinates) and the response and present the identifiability conditions. This leads to parametric expressions for the direct as well as overall and coordinate-wise indirect effects.

2.2.1 Causal effects with a compositional mediator

We treat the total, direct and indirect effects of exposure within the potential outcomes framework [8]. Denote the levels of the binary exposure X as x and x^* and let Y be a continuous response variable. Let $Y(x', \boldsymbol{m})$ denote the response if exposure was set to level x' and a vector-valued mediator $\boldsymbol{M}$ to value $\boldsymbol{m}$. In our application, the mediator will be the J-vector of the ilr coordinates. Then, $Y(x', \boldsymbol{M}(x''))$ is the response if X was set to level x' and the mediator $\boldsymbol{M}$ would obtain the value it would naturally have when exposure is set at x''. Here $x', x'' \in \{x, x^*\}$.

To define coordinate-wise effects, we also need more detailed manipulations of the vector-valued mediator $\boldsymbol{M}$. Specifically, for any $k = 1, \ldots, J$, let $Y(x_0, \boldsymbol{M}_{k-}(x_1), M_k(x_2), \boldsymbol{M}_{k+}(x_3))$ denote the response if exposure is set at x_0, and $(M_1, \ldots, M_{k-1})$, M_k and $(M_{k+1}, \ldots, M_J)$ would obtain values they would have when exposure is set at x_1, x_2, and x_3, respectively. Here $x_0, x_1, x_2, x_3 \in \{x, x^*\}$. We assume that for each coordinate such an assignment under treatment X can be done independently of the other coordinates. In our application, the plausibility of this assumption is based on the nature of a sequential binary partition, in which one balance can be set without affecting the rest of the balances. Note that with $x_0 = x'$ and $x_1 = x_2 = x_3 = x''$ (and for any k) we obtain as a special case that $Y(x', \boldsymbol{M}_{k-}(x''), M_k(x''), M_{k+}(x'')) = Y(x', \boldsymbol{M}(x''))$.

Following the standard approach to causal mediation analysis [8,9], we define the natural direct effect (NDE), the natural overall indirect effect (OIE) and the total effect (TE) as follows:

$$\begin{aligned} \text{NDE} &= \text{E}[Y(x, \boldsymbol{M}(x^*))] - \text{E}[Y(x^*, \boldsymbol{M}(x^*))], \\ \text{OIE} &= \text{E}[Y(x, \boldsymbol{M}(x))] - \text{E}[Y(x, \boldsymbol{M}(x^*))], \\ \text{TE} &= \text{E}[Y(x, \boldsymbol{M}(x))] - \text{E}[Y(x^*, \boldsymbol{M}(x^*))]. \end{aligned} \tag{5}$$

As usual, the above effects are average direct and indirect natural effects where the average is taken with respect to the population of units (individuals) [22]. The total effect equals the sum of the natural direct and natural indirect effects, i.e., TE = NDE + OIE. The indirect effect above refers to the effect on exposure mediated through the entire vector $\boldsymbol{M}$. NDE is the expected change in the response Y when the value of exposure changes from x^* to x while the mediator is set to the value it would have under $X = x^*$. OIE is the expected change in the response Y when the mediator is changed from the level it would have at $X = x^*$ to the level it would have at $X = x$ while holding the level of exposure at $X = x$.

Based on Wang et al. [17], we define the coordinate-wise indirect effects sequentially, concerning one coordinate at a time:

$$\begin{aligned} \text{CIE}_k = &\text{E}[Y(x, \boldsymbol{M}_{k-}(x), M_k(x), \boldsymbol{M}_{k+}(x^*)] - \\ &\text{E}[Y(x, \boldsymbol{M}_{k-}(x), M_k(x^*), \boldsymbol{M}_{k+}(x^*)],\ k = 1, ..., J. \end{aligned} \tag{6}$$

An alternative definition considers each coordinate with all remaining coordinates set at their reference levels [18]:

$$\begin{aligned} \overline{\text{CIE}}_k = &\text{E}[Y(x, \boldsymbol{M}_{k-}(x), M_k(x), \boldsymbol{M}_{k+}(x))] - \\ &\text{E}[Y(x, \boldsymbol{M}_{k-}(x), M_k(x^*), \boldsymbol{M}_{k+}(x)],\ k = 1, ..., J. \end{aligned} \tag{7}$$

In definition (6), the order of coordinates is important when interpreting the coordinate-wise indirect effects. Each CIE_k corresponds to the expected change in response Y when an individual coordinate M_k is changed from the level it would have at $X = x^*$ to the level it would have at $X = x$, keeping the perceding coordinates at the exposed level (x) and the rest at the base level (x^*). Of note, this sequential approach complies with the sequential binary partition used in building the ilr coordinates. Based on definition (6), the coordinate-wise indirect effects sum up to the overall indirect effect, i.e., $\text{OIE} = \sum_{k=1}^{J} \text{CIE}_k$ (see Supplementary material A.1).

In definition (7), $\overline{\text{CIE}}_k$ is the expected change in the response Y when an individual component M_k of the mediator is changed from the level it would have at $X = x^*$ to the level it would have at $X = x$, while all other mediators are set to the same level $X = x^*$. Thus, the order of the mediators is not of importance. Under this definition, the coordinate-wise effects do not sum up to the overall indirect effect. Nevertheless, under the linear dependence of the response on the mediators parameterised in Section 2.2.2 both of these non-parametric definitions lead to the same parametric expression.

It is important to note that changing the level of exposure x inevitably changes the entire composition due to the unit-sum constraint. This means that indirect effects on the *component-wise* level cannot be defined in a counterfactual manner.

However, it is possible, and even meaningful, to consider the coordinate-wise indirect effects with this counterfactual interpretation. In particular, each of the CIE coordinates, as coded on the basis of a hierarchical hypothesis-driven contrast matrix, pertains to a specific causal question on the mediating role of the partition into two subcompositions, with the relative proportions of all other parts remaining intact.

2.2.2 Parametric models

We here define parametric models for the dependencies between the exposure, mediator and response (Figure 1 (b)). In what follows, we denote the individual explicitly. For example, for individual i, the count observation is $\boldsymbol{K}^{(i)} = (K_1^{(i)}, ..., K_{J+1}^{(i)})$ and the mediator $\boldsymbol{M}^{(i)} = (M_1^{(i)}, ..., M_J^{(i)})$. There are three sets of variables that confound causal influences between $X^{(i)}$, $Y^{(i)}$ and $\boldsymbol{M}^{(i)}$, namely exposure-outcome confounders $C^{XY(i)}$, exposure-mediator confounders $C^{XM(i)}$ and mediator-outcome confounders $C^{MY(i)}$. We here consider these confounders as categorical so that the regression models can be stratified by their levels. Let H denote the set of strata pertaining to the variables $C = \{C^{XM}, C^{XY}, C^{MY}\}$ and $P(h)$ be the probability of a unit belonging to stratum $h \in H$. The number of exposed and unexposed participants in each stratum h may vary but are assumed to be at least one. The set of an individual's levels of confounders is denoted as $C^{(i)}$ and the stratum based on the confounders as $h^{(i)}$.

Parametric models for the compositional mediator and response As the read counts often exhibit overdispersion in the total count as well as in the class-specific counts, we define the distribution of the counts for individual i $(K_{i1}, \ldots, K_{i,J+1})$ hierarchically as follows:

$$\begin{aligned} &K^{(i)} \sim \text{NegativeBinomial}(\mu, \theta), \\ &(\pi_1^{(i)}, \ldots, \pi_{J+1}^{(i)}) \sim \text{Dirichlet}(\alpha_1, \ldots, \alpha_{J+1}), \\ &(K_1^{(i)}, \ldots, K_{J+1}^{(i)} | K^{(i)}; \pi_1^{(i)}, \ldots, \pi_{J+1}^{(i)}) \sim \text{Multinomial}(K^{(i)}, \pi_1^{(i)}, \ldots, \pi_{J+1}^{(i)}). \end{aligned} \tag{8}$$

The sum $\alpha_S = \sum_{j=1}^{J+1} \alpha_j$ controls the sparsity of counts. In particular, with small values of α_S the individual's class-specific counts are sparse in the sense that they tend to concentrate in a single class while the other classes may have very small or even zero counts. The dominating class then also varies across individuals. When $\alpha_S \to \infty$, the taxon-specific counts approach the multinomial model. Parameter θ controls the variation of the total count across individuals. When $\theta = 0$, the distribution of the total count is Poisson whereas large values of θ correspond to overdispersion with respect to the Poisson model. Conditionally on the total count $K^{(i)} = \sum_{j=1}^{J+1} K_j^{(i)}$, the distribution of $(K_1^{(i)}, \ldots, K_{J+1}^{(i)})$ is Dirichlet-multinomial.

The distribution of the counts depends on the individuals' exposure and confounders. Specifically, we let the parameters of the Dirichlet distribution in (8) depend on exposure $X^{(i)}$ as follows:

$$\alpha_j^{(i)} | X^{(i)}, C^{(i)} = a_{h_0 j}^{(i)} + a_{1j} X^{(i)}, j = 1, \ldots, J+1. \tag{9}$$

The ilr coordinates, i.e., the mediators, are taken to depend on the exposure as follows:

$$M_k^{(i)} | X^{(i)}, C^{(i)} = \beta_{h^{(i)}0k} + \beta_{h^{(i)}1k} X^{(i)} + \varepsilon_k^{(i)}, k = 1, \ldots, J, \tag{10}$$

where the stratum-specific regression parameters $\beta_{h^{(i)}0k}$ and $\beta_{h^{(i)}1k}$, pertaining to the expected values of the ilr coordinates, follow from models (8) and (9) through transformation (3). However, they lack analytical expressions under sparsity and overdispersion. Importantly, the variation in the class probabilities and the counts across individuals is reflected in the variation of the mediator. In addition, for any one individual, the error terms $\varepsilon_k^{(i)}$ for $k = 1, \ldots, J$, are correlated, reflecting the dependence between the ilr coordinates. This dependence is based on the choice of the contrast matrix and the correlation between the proportions. We note that the linear dependence of the Dirichlet parameters on the confounders in equation (9) does not necessarily induce a linear model for the ilr coordinates in equation (10). Therefore, for convenience, we have assumed that all confounders are categorical or can be categorized so that the linear dependence on them remains trivially in equation (6). We return to discuss the implications of this assumption in the Discussion. In addition, we explore the practical estimability of mediated effects under continuous covariates in Section 3.1.3. In our model, the exposure and confounders are not assumed to affect the total count $K^{(i)}$. Instead, they only affect the individual-specific proportions and, subsequently, the ilr coordinates.

Lastly, the exposure and mediators are assumed to affect the response in a linear way:

$$Y^{(i)} | X^{(i)}, \boldsymbol{M}^{(i)}, C^{(i)} = \gamma_{h^{(i)}0} + \sum_{k=1}^{J} \gamma_{1k} M_k^{(i)} + \gamma_2 X^{(i)} + e^{(i)}, \tag{11}$$

where the error term $e^{(i)}$ is assumed to be normally distributed with variance σ_2^2. Of note, in this model we assume no exposure-mediator interaction.

Identifiability To enable the estimation of the natural direct and indirect effect from empirical observations, a set of assumptions is needed [8]. Here, we require the existence of set of variables C such that the following sequentical ignorability assumptions hold [17]:

$$\text{(i) } \{Y^{(i)}(x, \mathbf{m}), \boldsymbol{M}^{(i)}\} \perp\!\!\!\perp X^{(i)} | C^{(i)},$$
$$\text{(ii) } Y^{(i)}(x', \boldsymbol{m}) \perp\!\!\!\perp M_k^{(i)}(x') | X^{(i)} = x', C^{(i)};\ \ k = 1, \ldots, J; x' \in \{x, x^*\}.$$

The identifiability assumptions are discussed in more detail in Supplementary material A.1.

Parametric expressions for the mediated effects Based on the linear models (10) and (11), the non-parametric definitions (5) and (6), and the identifiability conditions, the parametric expressions of the direct and indirect effects are found to be:

$$\begin{aligned} \text{NDE} &= \gamma_2(x - x^*), \\ \text{OIE} &= \sum_{k=1}^{J} \gamma_{1k} \sum_{h \in H} \beta_{h1k} \text{P}(h), \\ \text{CIE}_k &= \gamma_{1k} \sum_{h \in H} \beta_{h1k} \text{P}(h). \end{aligned} \tag{12}$$

The equations leading to these parametric forms are derived in Supplementary material A.1. With the above definitions, the coordinate-wise indirect effects sum up to the overall indirect effect, i.e., $\text{OIE} = \sum_{k=1}^{J} \text{CIE}_k$. While the OIE describes the overall influence transmitted through the composition (arrows (ii) and (iii) in Figure 1 (b)), each CIE_k describes the effect of a single coordinate (e.g. arrows ii.1 and iii.1 in Figure 1 (b)). As the coordinate-wise effects may counteract, it is important to consider both the OIE and the CIE_k's. This also means that individual CIE_k's may be larger than the OIE. Note that with the chosen linear parametric model for the dependence of the response on the mediators, both definitions (6) and (7) lead to the same parametric expressions and thus, in this parametric model, the additivity holds regardless of which of the two non-parametric definitions of the coordinate-wise indirect effects is used.

In practice, the total effect is estimated through linear regression of the exposure on the response. The direct effect is estimated as the effect of exposure on the response controlling for the mediators based on model (11). The indirect effects are estimated based on the multivariate linear model (10) of the dependence of the ilr coordinates on the exposure and model (11) of the dependence of the response on ilr coordinates. The standard errors of the total, direct, and indirect effects are based on the delta method (Supplementary material A.2).

A remark on a single vs. several contrast matrices. In this article, we address overall and coordinate-wise indirect effects as based on a single causal DAG or, equivalently, a single SBP matrix. It is important to note, that the relationship between the coordinate-wise and overall effects only holds in this case under the parametrisations as defined above. If alternating pivot coordinates are built, so that each of the coordinates is used as the pivot in turn, relying each time on a different SBP matrix, the sum of individual indirect pivot effects would not sum up to the overall indirect effect. Nevertheless, under linear models, regardless of how the contrast matrix is built the OIE remains the same, as choosing a different contrast matrix is equivalent to choosing another orthonormal basis for the simplex, and the information contained in all of the coordinates still captures the information contained in the composition. Hence, the coordinates $\boldsymbol{M}$ corresponding to a contrast matrix are rotations of those with another contrast matrix and can be obtained by multiplication by an orthogonal matrix $\mathbf{P}$. When regression equations (10) and (11) are formulated in terms of the new coordinates $\tilde{\boldsymbol{M}} = \mathbf{P}\boldsymbol{M}$, the regression parameters are transformed to $\tilde{\boldsymbol{\gamma}} = \mathbf{P}\boldsymbol{\gamma}$ and $\tilde{\boldsymbol{\beta}} = \mathbf{P}\boldsymbol{\beta}$, where $\boldsymbol{\gamma}$ and $\boldsymbol{\beta}$ are the vectors of parameters $\gamma_{11}, \ldots, \gamma_{1,J}$ and $\beta_{h11}, \ldots, \beta_{h1,J}$, respectively. It follows that $\tilde{\boldsymbol{\gamma}}^T\tilde{\boldsymbol{\beta}} = \boldsymbol{\gamma}^T\mathbf{P}^T\mathbf{P}\boldsymbol{\beta} = \boldsymbol{\gamma}^T\boldsymbol{\beta}$, i.e., the inner product of the two sets of coefficients is invariant under the orthogonal transformation of the basis. We demonstrate empirically that the individual coordinate-wise effects differ when using different matrices while the OIE remains the same in Section 3.1.3 (part "Misspecification of the contrast matrix"). While a similar invariance of the contrast matrix also holds for the overall indirect effects for e.g. binary responses under the log-odds scale, the identification of coordinate-wise effects would require further assumptions about a correlation structure of the component-wise mediators, not verifiable from empirical observations [17].

It is important to note that while the OIE remains the same, the interpretation of the individual coordinates changes. Using another contrast matrix corresponds to another hypothesised pathway that no longer corresponds to the same sequential definition of the coordinate-wise indirect effects as presented in Equation (6). If another causal structure is,

however, hypothesised and corresponding estimands are defined, a different set of identifiability assumptions is needed in order to the resulting coordinates to carry a causal interpretation.

3 Results

3.1 Simulation study

To demonstrate the estimation of effects mediated through a compositional entity we set up a simulation study. Specifically, we investigated the effect of overdispersion and sparsity of the underlying counts on the estimation.

3.1.1 Simulation set-up

We investigated the estimation of the OIE and the CIE's in different settings with varying extents of sparsity α_S and overdispersion θ. Table 1 presents the parameters in the simulation study. The sample size of each simulated dataset was $n = 1000$. We also performed two additional simulation studies for the one of the simulation scenarios with sample sizes 100 and 10000. The expected value μ of each individual's total count was fixed at 10000, a realistic level of the total in read-count data.

The amount of overdispersion, which reflects technical rather than biological variation across individuals, was controlled by θ, ranging from no overdispersion to approximately 1000-fold or 5000-fold excess variance of the total count relative to the standard Poisson distribution. Sparsity in turn follows from true biological variation and reflects actual heterogeneity in the microbiome across individuals. To gain insight on the effect of sparsity on the performance of estimation, we considered three distinct settings, controlled by parameter α_S: no sparsity, leading to fixed class-specific multinomial probabilities, intermediate sparsity (200-fold excess variance in the class-specific counts, $\alpha_S = 50$) and extreme sparsity (5000-fold excess variance, $\alpha_S = 1$).

Data were simulated following the DAG of Figure 1 (b) assuming 5 classes of underlying counts. We assumed two binary confounders, C_1 and C_2, that affect the exposure, mediator and response, and divided the data accordingly into four strata. We set the values of the class-specific probabilities, α_j/α_S, under three distinct scenarios (Table 1). In each scenario, the values were chosen so that the class-specific probabilities were different between the five classes so that not all simulated taxa have the same relative abundance. In scenarios 1 and 2, we assumed the underlying taxonomic hierarchy of Figure 2 (b). In scenario 1, the parameters were chosen so that differences in the expected class-specific proportions (π in equation (8)) between groups $X = 0$ and $X = 1$ were clear-cut, while in scenario 2 these differences were set to be moderate. Scenario 3 demonstrates the use of coordinates based on a single pivotal SBP matrix in which exposure X affects the proportion of the first component, while the remaining components compensate so that their relative probabilities remain the same.

Given μ, θ and α_S, read counts were simulated from model (8) and further scaled into ilr coordinates through Equation (2). For scenarios 1 and 2, the ilr coordinates were built following the contrasts of matrix (1) whereas in scenario 3 a single pivotal SBP matrix with the mediating part of the composition as the pivot was used. Given models (8) and (9), the compositions and ilr coordinates and, subsequently, the true values of the stratum-specific parameters $\beta_{h1k}, k = 1, \ldots, J$, of Equation (10) are defined. However, lacking an analytical expression, we approximated the true values of the β parameters as differences in the mean values of the ilr coordinates under $X = 1$ and $X = 0$, as realised by Monte Carlo simulation, separately for each scenario and each combination of θ and α_S.

For each scenario, we fixed OIE at 0.10 and the CIE's as presented in Table 1. The OIE of 10 % corresponds to 20 % of the total effect. In scenario 1, we fixed each of the coordinates to carry a positive indirect effect. In scenario 2, the magnitudes of the CIE's were the same as in scenario 1, but the β_{ik} parameters were smaller due to the smaller difference between the $X = 0$ and $X = 1$ groups. In scenario 3, OIE = CIE_1 = 0.10. Under each scenario and parameter setting, the γ_{1k} parameters were chosen so that the products $\gamma_{1k}\beta_{ik}$ were equal to the corresponding coordinate-wise CIE's. A weighted average value β_{1k} over all four strata of the β_{1kh} values was used (see Equation (12) and Table S1 in Supplementary material). Finally, continuous responses were simulated according to model (11).

We conducted an additional analysis to investigate how continuous confounders affect the estimation (Supplementary material B.2). This analysis also includes an exploration of the linearity of the relationship between a continuous confounder and the ilr coordinates.

3.1.2 Estimation of mediated effects

In scenarios 1 and 2, the indirect effects (CIE and OIE) were estimated using ilr coordinates that were based on the same contrast matrix used to simulate the data. In scenario 3, we applied two approaches for building the ilr coordinates when assessing mediation. First, we employed the correct contrast matrix with the first part of the composition as the

pivot (model 3A). Second, we used a misspecified contrast matrix by setting the pivot as the last part of the composition (model 3B). Figures 3 (a) and 3 (b) show the DAGs corresponding to these two assumed causal structures. When deriving the ilr coordinates, zero-counts were replaced by 0.5, which is a preprocessing step commonly used in the analysis of microbial count data [10,7].

We estimated the CIE's and OIE's in each stratum following models (10) and (11) and derived the effects adjusted for confounding by calculating weighted averages over the strata. The dependencies between the ilr coordinates were taken into account by assuming a multivariate linear model for the coordinates. The parameter estimation was conducted using `R` (version 4.1.1). For each parameter setting, we generated 1000 data sets. The performance of estimation was assessed by bias, average standard error (mean of the standard error estimates) and empirical standard error (standard deviation of the estimates). The ability to identify the mediated effects was assessed by the statistical power and coverage probability of the 90 % confidence interval (CI). The confidence interval was calculated based on the estimates of the mediated effects and their standard errors, derived using the delta method (Supplementary material A.2), as follows: $[\text{IE} - 1.645\text{s.e.(IE)}; \text{IE} + 1.645\text{s.e.(IE)}]$.

3.1.3 Performance of the estimation

Tables 2 and 3 present the results of the simulation study for scenarios 1 and 2. The results of analyses 3A and 3B under Scenario 3 are presented in Supplementary material (Tables S2 and S3, respectively). For further insights, Tables S4 and S5 present the expected values and variances of the ilr coordinates and Table S6 the average values of the estimates of the β_{1k} and γ_{1k} parameters and their variances.

Bias. The estimation of the coordinate-wise and overall indirect effects was generally unbiased. Under the multinomial counts, i.e., when $\alpha_S \to \infty$, the estimation appears occasionally slightly biased. This stems from the large variances of the estimators of the γ_{1k} parameters under multinomial counts and the relatively small sample size ($n = 1000$). The overdispersion term θ had practically no effect on the bias.

Standard errors. The average standard error and the standard deviation of the 1000 replicate estimates were essentially the same, even under extreme sparsity (Tables 2, 3, S2, S3). To understand how the standard error depends on the parameters, consider the asymptotic variance of the estimator of a single coordinate-wise indirect effect,

$$\sigma^2_{\gamma_k}\sigma^2_{\beta_k}\left[\left(\frac{\beta_{1k}}{\sigma_{\beta_k}}\right)^2 + \left(\frac{\gamma_{1k}}{\sigma_{\gamma_k}}\right)^2\right], \tag{13}$$

where $\sigma^2_{\gamma_k}$ and $\sigma^2_{\beta_k}$ are the variances of the estimators of β_{1k} and γ_{1k}, $k = 1, \ldots, J$ (see Supplementary material A2 and Equation S5 therein). Based on the simulations, we found that the product $\sigma^2_{\gamma_k}\sigma^2_{\beta_k}$ remains essentially constant. This is obviously due to the fact that when the variance of the ilr coordinates (mediators) is large, the β_k's are estimated less precisely but at the same time the large variability of the ilr coordinates leads to more precise estimation of the γ_k's. The variance (13) thus mainly depends on the sum of the squared ratios $(\beta_{1k}/\sigma_{\beta_k})^2$ and $(\gamma_{1k}/\sigma_{\gamma_k})^2$. Table 4 presents these ratios and their product for the first CIE under a number of different parameter settings.

Higher sparsity means larger between-individual heterogeneity of the ilr coordinates and, subsequently, larger β_{1k} parameters due to the larger difference between the mean ilr values in the unexposed and exposed groups (cf. Equation (7)). However, with increasing sparsity, the variance $\sigma^2_{\beta_k}$ increases faster than β^2_{1k}, and thus the ratio $(\beta_{1k}/\sigma_{\beta_k})^2$ decreases. The effect of sparsity on γ^2_{1k} and $\sigma^2_{\gamma_k}$ is reciprocal and thus the opposite occurs with the ratio $(\gamma_{1k}/\sigma_{\gamma_k})^2$ (Table 4).

The reciprocal behaviour of the two ratios in (13) appears to imply that the asymptotic variances of the estimators of the indirect effects are smallest at intermediate levels of sparsity. However, where the optimum lies also depends on the relative magnitudes of the two ratios and how fast they decrease/increase as sparsity increases. For example, the larger the contrast between the two exposure group ($X = 0$ vs. $X = 1$) is, the larger the value β_{1k} and the corresponding ratio and, due to the fixed value of CIE_k, the opposite is true for γ_{1k} and the other ratio. We found that under scenarios 1 and 3, where the contrasts between the exposure groups were large, the variances of the indirect effects were smallest under the highly sparse settings and largest in the multinomial settings (Tables 2 and S2)). By contrast, under scenario 2 with smaller contrasts between the exposure groups, the variances were smallest at intermediate levels of sparsity (Table 3) .

The impact of overdispersion on the standard errors of the indirect effects remained moderate. Higher overdispersion appears only slightly to increase the standard errors compared to the case with a Poisson variance in scenario 2. With intermediate sparsity, overdispersion did not seem to have a strong effect on the standard errors in either of the scenarios, whereas under multinomial counts overdispersion slightly decreased the standard errors.

Power. In scenarios 1 and 3, the best statistical power was obtained under extreme sparsity (Tables 2 and S2) whereas in scenario 2 intermediate sparsity yielded the best power (Table 3). These findings result from the differences in the

standard errors as described above. Despite the slight impact of the overdispersion parameter θ on the standard errors of the indirect effects, it only affected the statistical power in an inconsistent and marginal manner.

In scenario 1, most often the mediating effect via the first coordinate was estimated with the best statistical power while the power of estimation for the other coordinates was smaller. Often, the statistical power of the OIE was rather similar to the statistical power of the first CIE, stemming from its large magnitude and standard error. In simulation scenario 2, on the other hand, the CIE-specific statistical power varied depending on sparsity. While under extreme sparsity the 3rd and 4th CIE had the largest statistical power, under other levels of sparsity the 1st CIE was estimated with the best statistical power. This boils down to the interplay of the path-specific estimators and their variances as explained below Equation (13). For simulation scenario 3, the statistical power of the first CIE was large. For the other coordinates with a null effect, interpreting the statistical power is not reasonable. As only the first coordinate contributes to the OIE, its statistical power is very similar to the statistical power of the first coordinate.

Misspecification of the contrast matrix. In scenario 3, true data were generated so that the first ilr coordinate carries the entire mediated effect. We analysed the simulated data using the correct contrast matrix (model 3A, Table S2) and a misspecified matrix (model 3B, Table S3). Figure 3 describes the log-contrasts included in each of the coordinates under the correct contrast matrix (a) and the misspecified one (b). Under model 3A the first class is only included in the first coordinate (M_{1A}) as a "pivot" element and not contributing to the other coordinates, while under the misspecified contrast matrix (model 3B) this class is involved in each log-contrast. While the OIE remains the same in both models, in agreement with the argument given in Section 2, model 3B represents a different causal model, and thus the CIE's have a different interpretation. Under model 3A, the pivotal indirect effect of the first part of the composition were identified correctly. However, under model 3B each of the coordinates carries a part of the mediating effect of coordinate M_{1A} (Table S7). Of note, while model 3A offers the most parsimonious explanation of the causal pathways, under model 3B this simple causal explanation is masked. With misspecified contrasts, each coordinate appears to carry a mediating effect (Table S7). One can thus be led to infer that the contrasts of each element of the composition against p_1 and other remaining parts of the composition play a role in carrying the mediating effects. Importantly, if the pivotal contrast matrix was used so that each of the components were used as the first pivot in turn, the first coordinate-wise indirect effects for these alternations would be 0.10, 0.01, 0.01, 0.01 and 0 summing up to 0.13 (in an example scenario with $\theta = 0$ and $\alpha_S = 0$; data not shown). This demonstrates that the sum of individual first pivot effects does not equal the true overall indirect effect, and thus the alternating pivot approach cannot be used to estimate the OIE. Of note, these alternating approaches aiming at searching of signals of causal mediation are deliberately based on trying multiple (often a very large number of) missspecified pivots. It is evident that the estimated effects from a misspecified model would not correspond to the true effects based on the correctly specified model. Furthermore, defining the "true" values for the effects under the misspecified model would not be reasonable. For this reason, statistics relying on the true values (i.e., bias, power and coverage probability) have been omitted from Table S3.

Sample size. As an additional analysis under simulation scenario 1, we investigated the effect of sample size on the performance of estimation using sizes $n = 100$ and $n = 10000$. The detailed results of these simulations are summarised in the Supplementary Tables S8 and S9. Compared to the base case of $n = 1000$, the biases were larger with the smaller sample ($n = 100$) when the counts were multinomial but remained similar under extra-multinomial variation. The average standard errors were nearly five-fold. However, especially under large amounts of sparsity, the statistical power was poorer. When the sample size was large ($n = 10000$), the performance of the mediation analysis clearly improved. The bias was small in each scenario, and the standard errors were smaller. Also the statistical power was better, especially for scenarios with sparsity.

Continuous confounders. Another additional analysis under simulation scenario 1 was performed to study the effect of confounders that are continuous instead of categorical. The details of the simulation set-up and results are presented in Supplementary material B.2. Briefly, in most simulation settings, an approximate linearity in the association between the ilr coordinates and the confounders held (Figure S1). However, under multinomial counts with no excess sparsity, the relationship between the coordinates and confounders departed from linear under some settings, which was also reflected in the bias in the estimation of the indirect effects (Table S10). However, under all other settings the performance of estimation was satisfactory even with continuous confounders.

3.2 Estimation of the effect of fibre intake on insulin level via the gut microbiome

3.2.1 Background

We applied the methods of Section 2 to investigate the mediating role of the gut microbiome in the association between sufficient fibre intake and serum insulin level. Sufficient fibre intake is known to bring numerous health benefits, such as healthy and diverse gut microflora and metabolic health including insulin sensitivity [23]. Furthermore, based on a recent literature review, the gut microbiome plays a role in the pathophysiology of type 2 diabetes [24]. Thus, it is of

interest to investigate whether fibre intake affects the insulin level through the gut microbiome. The potential mediating pathway of fibre intake on the insulin level via the microbiome stems from the production of short-chain fatty acids (SCFA) by microbial fermentation of dietary fibre. SCFAs in turn may play a role in insulin sensitivity via the gut barrier function and reduced inflammation [23,25].

In addition to numerous specific genera, the *Actinobacteria* phylum has been suggested as an example of a taxonomically high level that might be associated with type 2 diabetes (e.g. [24]). Here, we investigate whether the relative proportion of the *Actinobacteria* compared to the other phyla mediates the effect of fibre intake on the insulin level. Furthermore, we investigate whether the genera within the *Actinobacteria* phylum mediate such effects.

3.2.2 Data description

We retrieved data from the most recent (26-year) follow-up visit conducted in the Special Turku Coronary Risk Factor Intervention Project (STRIP) (Supplementary material C.1 and [26,27]). Of the 546 participants who participated in the follow-up study, 325 provided data on diet, insulin levels and had successfully-sequenced faecal samples. We excluded participants with type 1 diabetes and those who had not fasted before blood sampling or had had antibiotic treatment during the past three months. In addition, we excluded obese participants (BMI $\geq 30\text{kg/m}^2$) to avoid a potential collider bias (see Supplementary material C.2 and Figure S2 therein for further details). The resulting sample size was $n = 264$.

The participants were initially recruited at age 6 months in well-baby clinics and randomised into intervention and control groups. The intervention group received dietary counselling on biannual visits throughout childhood and adolescence until 20 years of age, whereas the control group visited the research clinic twice a year until the age of seven and after that once a year but did not receive counselling. The intervention included promoting the use of unsaturated fats, whole-grain products, fruits and vegetables and lowering the intake of saturated fat and salt. The 26-year follow-up visit was carried out six years after the intervention period had ended.

The raw gut microbiome sequence data were processed into an amplicon sequence variant (ASV) table using the *DADA2* pipeline [28]. Prior to transforming the counts into ilr coordinates, the taxa that appeared in less than 10 $\%$ of the samples were filtered out. As the ilr transformation does not allow zero-counts, we replaced all remaining zero counts by the maximum rounding error of 0.5. The observed insulin level was transformed using the natural logarithm. The fibre intake was assessed based on food diary and quantified using the Micro Nutrica programme [29]. Fibre intake was treated as binary, and sufficient fibre intake was defined as ≥ 25 grams per day, as per Finnish dietary recommendations in 2014 [30]. Of the 264 participants, 63 (23.9%) achieved this target. Belonging to the intervention group and sex were considered as potential confounders, and the sample was stratified accordingly. The stratum-specific sample sizes were 68, 82, 59, and 55 for intervention females, control females, intervention males and control males, respectively.

3.2.3 Conduct of the mediation analysis and results

We investigated two mediation questions on different phylogenetic levels of the microbiome. The first question concerned the mediating role of the genera within the *Actinobacteria* phylum in the association between sufficient fibre intake and $\log$(insulin) level. The second question utilised a pivotal SBP matrix, focusing on the mediating effect of the relative proportion of the *Actinobacteria* in the association between sufficient fibre intake and $\log$(insulin) level.

Models (10) and (11) were fitted separately for each stratum defined by the confounders. We report estimates averaged over the strata (see Equation 12). The stratum-specific results are available in Supplementary material C.3.

The average total counts were approximately 4200 and 180 000 for the mediation problems within the *Actinobacteria* phylum and over the different phyla, respectively, whereas the amounts of excess variance in the total count were 12000-fold and 54000-fold, respectively. In terms of the simulation setting and model (8), these correspond to values 3 and 0.3 of parameter θ.

To gain insight on the sparsity present in our data, we calculated a rough moment-based estimate of α_S for the genera- and phyla-level analyses based on the observed counts. We first calculated the empirical variance for each of the taxon-specific counts, the average total count and the average taxon-specific proportion from the empirical data set. We then utilized the expression of the (marginal) variance of count K_j for the jth taxon under the assumed Dirichlet-multinomial model:

$$\text{Var}(K_j) = K(\alpha_j/\alpha_S)(1 - \alpha_j/\alpha_S)\frac{(\alpha_S + K)}{(\alpha_S + 1)}.$$

For each taxon separately, we obtained a rough estimate of α_S based on the empirical variance, the average total count (K) and the taxon-specific proportions in place of α_j/α_S. Finally, a rough estimate of α_S was determined as the median of these taxon-specific estimates. While the rough taxon-specific estimate for α_S varied between the taxa, the median α_S was 36 within the phylum-specific counts and 33 within the genus-specific counts.

Mediation through the genera within the *Actinobacteria*
After the preprocessing, the data on the *Actinobacteria* phylum included one class, which was divided into three orders and within these, four families. In the four families, there were altogether 11 genera (Figure 4). We first investigated the mediating effect of this subcomposition of the microbiome. Table 5 presents the corresponding counts and proportions in the two fibre groups by genera. Based on the taxonomy (Figure 4) we built the SBP matrix as follows:

$$\begin{array}{c} \\ Bifidobacterium \\ Actinomyces \\ Rothia \\ Enterorhabdus \\ Eggerthella \\ Collinsella \\ Adlercreutzia \\ Slackia \\ Senegalimassilia \\ Olsenella \\ Gordonibacter \end{array} \begin{bmatrix} -1 & 0 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ +1 & +1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ +1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & +1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & -1 & +1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & -1 & -1 & +1 & 0 & 0 & 0 & 0 \\ -1 & 0 & -1 & -1 & -1 & -1 & +1 & 0 & 0 & 0 \\ -1 & 0 & -1 & -1 & -1 & -1 & -1 & +1 & 0 & 0 \\ -1 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & +1 & 0 \\ -1 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & +1 \\ -1 & 0 & -1 & -1 & -1 & -1 & -1 & -1 & -1 & -1 \end{bmatrix}.$$

Each of the 10 columns thus corresponds to one ilr coordinate M_k, $k = 1, \ldots, 10$, describing a single contrast. For example, the fourth column contrasts the *Enterorhabdus* genus against the other genera within the *Coriobacteriales* order.

The total effect of fibre on log(insulin) was estimated at -0.106 (90% CI $[-0.202; -0.011]$) whereas the direct effect was $\hat{\gamma}_2 = -0.095$ (90% CI $[-0.202; 0.012]$). The sufficient fibre intake thus seems to reduce insulin levels mostly via the direct route. The parameters $\hat{\beta}_{1k}$ describing the effects of sufficient fibre intake on the ilr coordinates are presented in Table 6. For example, sufficient fibre intake reduced the second coordinate, which corresponds to the relationship between relative proportions of *Actinomyces* and *Rothia* ($\hat{\beta}_{12} = -0.255$, $[-0.501; -0.009]$). By contrast, sufficient fibre intake increased M_4, which corresponds to the relative proportion of *Enterorhabdus* compared to the other genera within the *Coriobacteriales* ($\hat{\beta}_{14} = 0.409$, $[0.060; 0.759]$).

The $\hat{\gamma}_{1k}$ parameters in Table 6 describe the effects of each ilr coordinate on the log(insulin) level in a multivariate linear regression. For example, a higher level of M_3 decreased the insulin level, i.e., the higher the relative proportion of *Bifidobacterium* contrasted to *Coriobacteriales*, the higher the insulin levels ($\hat{\gamma}_{13} = 0.025$, $[0.002; 0.048]$). While M_4 (relative proportion of *Enterorhabdus* compared to other genera within the *Coriobacteriales*) decreased the log(insulin) level ($\hat{\gamma}_{14} = -0.041$, $[-0.074; -0.008]$), M_5 (relative proportion of *Eggerthella* contrasted to remaining genera within the *Coriobacteriales*) increased the log(insulin) level ($\hat{\gamma}_{15} = 0.038$, $[0.015; 0.060]$).

The CIE's $\hat{\beta}_{1k}\hat{\gamma}_{1k}$ with their 90 % confidence intervals are presented in Table 6. The largest mediated effect, yet inconclusive, acted through coordinate M_4, as sufficient fibre intake increased the relative proportion of *Enterorhabdus* compared to other genera within the *Coriobacteriales*, and the relative proportion of *Enterorhabdus* had a decreasing effect on the log(insulin) level ($\hat{\beta}_{14}\hat{\gamma}_{14} = -0.017$, $[-0.036; 0.003]$). The OIE was $\sum_{k=1}^{10} \hat{\beta}_{1k}\hat{\gamma}_{1k} = -0.011$ (90 % CI $[-0.073; 0.051]$), corresponding to approximately 10% of the total effect. These results demonstrate our previous note that the coordinate-wise indirect effects may counteract, making some of the CIE's larger than OIE in absolute value.

Mediation through the *Actinobacteria* phylum
We investigated whether the entire *Actinobacteria* phylum in comparison to the other phyla has a mediating role using a pivotal SBP matrix (4). After the preprocessing stage altogether 11 phyla remained. The contrasts of interest were encoded through a pivotal SBP contrast matrix where the pivot coordinate describes the proportion of the *Actinobacteria* against the proportions of the other phyla. Those who had sufficient fibre intake had a smaller proportion of *Actinobacteria* (1.6%) as compared to those who had insufficient fibre intake (2.7%).

Based on model (10), the average effect of sufficient fibre intake on the first ilr coordinate contrasting the *Actinobacteria* phylum to the other phyla was $\hat{\beta}_{11} = -0.451$ (90 % CI $[-0.762; -0.140]$). Furthermore, the higher proportion of the *Actinobacteria* phylum against the other phyla may have a positive association with higher insulin level, as the effect of the pivot coordinate on log(insulin) was $\hat{\gamma}_{11} = 0.042$ (90 % CI $[-0.004; 0.088]$). The CIE of the coordinate of interest was thus $\hat{\beta}_{11}\hat{\gamma}_{11} = -0.019$ (90 % CI $[-0.043; 0.006]$). The total effect of sufficient fibre intake on log(insulin) level remained as -0.106 (90 % CI $[-0.202; -0.011]$) whereas the direct effect was -0.10 (90 % CI $[-0.206; 0.005]$). The OIE was -0.006, i.e., only 5% of the total effect.

3.2.4 Sensitivity analysis for unmeasured confounding

To explore the possibility that our findings of mediated effects are due to unmeasured or unadjusted confounding, we undertook a sensitivity analysis following the framework of VanderWeele [31,32]. Briefly, we examined the extent to which an unmeasured confounder would need to affect the response and be associated with the exposure for it to explain away the total, direct and overall indirect effects. We considered being physically active as a potential unmeasured binary confounder that could affect the fibre intake, gut microbiome and insulin levels. While it seems very unlikely that the true total and direct effects would be null if we had accounted for physical activity, it is possible that the overall indirect effects could be explained away by physical activity. A comprehensive account of the sensitivity analysis approach and its results is given in Supplementary material C.4.

3.2.5 Conclusion of empirical mediation analysis

In this example of estimating compositional mediation, we observed that the effect of fibre intake on insulin level primarily stems from the direct effects, with the mediated effect of the microbiome accounting only 5-10 % of the total effect, with considerable uncertainty. Nonetheless, our empirical data analysis should be considered as a demonstration rather than a comprehensive analysis. For example, we did not adjust for other confounders apart from sex and intervention status. However, based on the sensitivity analysis, the unadjusted confounders, such as physical activity, would need to have a large effect on the insulin levels and highly correlate with fibre intake to be able to explain away the observed total and direct effects, while it is possible that the overall indirect effect could be explained away. In addition, even though reverse causality is implausible in the light of current biological knowledge, the temporal ordering of the exposure (fibre intake), mediator (gut microbiome) and response (insulin level) remains obscure as each was measured at the same visit. Although the direct and overall effects were of similar magnitudes whether analysing mediation of *Actinobacteria* at a phylum-level contrasted to other phyla or of the genera within the *Actinobacteria*, the models reflect different levels of taxonomic hierarchy and thus potentially different biological mechanisms.

4 Discussion

In this study, we proposed and investigated a hypothesis-driven approach to mediation analysis with a compositional mediator based on sparse and overdispersed microbial count data. We specified non-parametric causal estimands for both coordinate-wise and overall indirect effects relying on an *a priori* defined causal structure and presented conditions under which they can be identified from empirical data based on simple parametric models. The method can be applied when some *a priori* knowledge about the structure between the parts of the composition is available and allows simultaneous estimation of coordinate-wise indirect effects (CIE's) and the overall indirect effect (OIE). The estimation of mediated effects was found to be generally unbiased, even under extreme amounts of sparsity and overdispersion in the underlying count data. However, sparsity and the magnitudes of the path coefficients were found to affect the precision of the estimation in a complex way. In addition, we demonstrated the estimation of compositional mediation with an empirical data set, investigating the potential indirect effect of the gut microbiome in the association between dietary fibre intake and insulin level. The total effect was found to be dominated by the direct effect, but we cannot rule out the possibility that certain parts of the gut microbiome, such as the *Enterorhabdus*, carry a small mediating role.

The use of ilr coordinates transforms proportions in the simplex into Euclidean coordinates. A standard statistical approach for mediation analysis can thus be applied to estimate indirect effects [9]. Importantly to the causal interpretation, the ilr coordinates have a one-to-one correspondence with a sequence of partitions of the entity of interest into subcompositions. Each single ilr coordinate corresponds to a partition of a specific subcomposition into two parts, while the relative proportions of all other parts remain intact, and the coordinate can thus be interpreted as representing a causal relationship encoded by that partition. Furthermore, with the assumed purely linear dependence of the response on the mediators and no exposure-mediator interaction, the overall effect is the sum of all coordinate-wise effects and quantifies mediation via the entire composition of interest.

The theory on multiple mediators mainly focuses on two scenarios, causally dependent mediators [14,15,16] and mediators that are correlated but do not affect each other causally [17,18,19]. Our treatment of mediation through ilr coordinates is an example of the latter. We considered two alternative ways to define causal effects non-parametrically. The definition of Wang et al. [17] leads to coherent estimands for the overall indirect effects and coordinate-wise indirect effects. For the counterfactual interpretation of the indirect effects, the order of the mediators is important, as the exposure levels of the other mediators are assigned sequentially. Of note, these non-parametric definitions correspond to the sequential build-up of the ilr coordinates. In contrast, in the alternative non-parametric definition of Kim et al. [18], the order of the mediators plays no role. While this formulation does not lead to a coherent decomposition of the overall indirect effect into the coordinate-wise indirect effects, under linear models and no interaction between the mediators

the parametric formulations are the same under both alternative non-parametric definitions, and the coordinate-wise indirect effects sum up to the overall indirect effect.

In Wang et al. [17], the parametric dependence of the response on the mediators was based on logistic regression. The parametric expressions of the indirect effects then depend on the joint distribution of the mediators. This in turn requires an *a priori* specification of their correlation structure. As our application here focuses on a linear dependence of the response on the mediators, the parametric expressions only depend on the estimable marginal distributions of the mediators.

Given that the aims and scope of our proposed method fundamentally differ from other studies that have addressed compositional mediation via the microbiome [7,10,11,12], a formal comparison of the methods' performance was not carried out. Instead, we here provide a comparison of the main characteristics of these approaches. Table 7 summarises some key features of these studies. The previous studies have focused on identifying signals of mediation through different taxa and rely on testing a large number of tentative (causal) associations between the components of the microbiome. As argued in the present study, many of such associations are potentially misleading if the aim is to answer to a structured causal question. In particular, when causal effects based on relevant biological mechanisms are of special interest, a more targeted analysis based on *a priori* knowledge may be called for in contrast to data-driven approaches that rather emphasise the predictive performance of the model. An analysis that is conditioned on a given set of contrasts can be thought as the next step on the pathway of building evidence about mediation.

All previously presented methods to analyse compositional mediation through the microbiome have been based on some log-ratio transformation. Sohn and Li [10] and Wang [11] used the additive logratio which differs from the isometric logratio in interpretation, whereas Zhang et al. [7] and Fu et al. [12] relied on the isometric logratio using alternating pivot coordinates. While we applied the isometric logratio, we deviate from these methods in three aspects. Firstly, instead of sieving for signals of mediation (i.e. testing for multiple null hypotheses), we aim at estimating mediated effects within a predefined *a priori* partition, which leads to investigating multiple contrasts of interest simultaneously without additional re-fitting of ilr transformations using alternating pivots. Secondly, we investigated how the sampling variation among the units, caused by sparsity and overdispersion, affects the estimation of mediated effects. Thirdly, our approach is confined to low-dimensional problems. Thus, regularisation or Bayesian model selection approaches are not needed. In addition, due to the lower dimension, our approach is free of multiple comparison problems of sieving approaches especially present in Zhang et al. [7], and as such allows unbiased estimation of the overall indirect effect. Lastly, the coordinate-wise effects defined in Fu et al. [12] were basically regression-based, i.e., they relied on a parametric formulation and thus lacked proper definition in terms of causal contrasts. Consequently, the parametric forms of the indirect effects are not similar to those of Wang et al. [17] or those considered in our paper.

Our approach is based on a pre-specified contrast matrix that contains the *a priori* causal assumptions encoded as a DAG. Thus, any conclusions that are drawn will depend on the specification of the contrast matrix, the analysis of mediation throughof a coordinate-wise indirect effects is conditional on the assumed contrasts. The idea of utilising knowledge about taxonomic distances in building ilr coordinates has been presented earlier, although in a data-driven manner [2]. Other possible sources of taxonomic knowledge include biological functionalities of different taxa, or hypotheses based on previous (data-driven) studies. The pivotal approach of Zhang et al. [7], where the coordinates are alternated $J + 1$ times to identify the mediating role of each part of the composition, corresponds to testing $J + 1$ different hypotheses, each concerning a role of a specific part of the composition.

Furthermore, as opposed to some previous research on compositional mediation where mediators were simulated directly as logratio coordinates based on a multivariate normal distribution [7,10], we simulated data at the count level to investigate the effect of sparsity and overdispersion on the performance of the estimation. Larger sparsity increases the variation of the ilr coordinate. Consequently, with any given level of CIE, we found that larger sparsity increased the uncertainty in the estimation of the β parameters and vice versa in the estimation of the γ parameters (see the discussion following equation (13)). The impact of overdispersion, on the other hand, was mostly removed when treating the counts as compositions. However, in some scenarios higher overdispersion slightly increased the variability of the ilr coordinates. Reassuringly, the levels of sparsity (α_S) and overdispersion (θ) in our empirical data analysis indicate that our choices of these parameters in the simulation study, as well as the chosen magnitudes of the indirect effects, were reasonable.

In terms of the precision of estimated mediation effects, it would be ideal to conduct the analysis under intermediate levels of sparsity. When designing a study, however, the components affecting the variance of the CIE's – that is, the magnitudes of the pathway coefficients (β_{1k} and γ_{1k}) and their variability – cannot be controlled in advance. The amount of sparsity varies in microbiome data and also depends on the taxonomic level of interest. At deeper taxonomic levels, such as genus level, the number of taxa is larger and data are more likely sparse. However, if the mediated effects of interest are considered at a higher level, each taxon can be considered as an amalgamation of lower-level taxa within that taxon, which potentially reduces sparsity. Thus, while the sparsity present in the data cannot be controlled within

any chosen level of taxonomy, restricting the analysis to a higher taxonomic level may result in less sparse counts. Additionally, the most extremely sparse counts are often removed while preprocessing the data. In general, one should pay specific attention to the sparsity of the counts and be cautious of its implications to the estimation of indirect effects.

There are some limitations to our approach. Firstly, as in any mediation analysis based on DAGs, it is possible to misspecify the causal structure e.g. by misspecifying the contrast matrix or not accounting for relevant covariates. As shown by the simulation example, a misspecified contrast matrix may result in estimates that do not reflect the true causal effects and lead to erroneous inferences on the mediating roles of the parts of the composition. This emphasises the importance of utilising expert knowledge or *a priori* information when coding the hypotheses into the contrast matrix. The *a priori* knowledge can be based on e.g. sieving studies, such as those suggested by e.g. Zhang et al., Sohn and Li or Wang et al. [7,10,11], or on biological knowledge of the roles of different parts of the composition, such as the relationship between *Bacteroidetes* and *Firmicutes* in human gut microbiome. It is of note that the contrast matrix depends on the question in hand, and thus no general guidelines on building the contrasts in a hypothesis-driven manner can be given. In addition, if there are unobserved or unadjusted confounders, the causal effects may not be estimated correctly.

Secondly, while we investigated the impact of sampling variation of the class probabilities and counts on the estimation of mediated effects in the simulation study, in the empirical analyses we simply assumed that the ilr coordinates are normally distributed. However, the normality may not always hold, as the variation in the class probabilities and counts is reflected in the variation of the ilr coordinates. Regardless, we showed that the estimated standard errors corresponded on average to the true ones as found by simulation.

Thirdly, we formulated the approach on the presupposition that confounders are categorical or can be categorised. Without this assumption, the linear dependence of the Dirichlet parameters on the confounders would not necessarily induce a linear dependence of the ilr coordinates on the confounders (see Equations (9) and (10)). The assumption of categorical confounders was used in the simulation study when assessing the performance of estimating mediated effects. Nevertheless, the general idea of basing causal mediation analysis on *a priori* defined hierarchy coded through a single SBP matrix remains valid, and there is nothing in principle that would preclude using continuous covariates in model (10).

Our additional analysis with continuous confounders showed that when an approximate linearity of association between the ilr coordinates and confounders holds, the performance of estimation of mediated effects remains similar. Previous studies on compositional mediation analysis in the context of microbiome data have simulated data directly at the log-ratio level assuming a multivariate normal distribution and linear dependence between coordinates and covariates [10,7]. Such approaches fail to account for the variation across individuals in the actual counts. In this respect, our analysis can be viewed as more comprehensive despite the assumption of categorical covariates.

In summary, we here have combined the framework of compositional data analysis with causal mediation analysis with specific focus on theory on multiple mediators and the causal interpretation of the coordinates based on the microbiome composition. The contribution of our work is threefold. First, we suggested incorporating *a priori* or expert knowledge in microbial mediation and presented microbial mediation analysis with specific focus on the contrasts between sub-compositions of interest rather than individual taxa. Second, we have combined the sequential approach of defining mediator-specific indirect effects [17] with the hypothesis-driven sequential binary partition in the context of microbial phylogenetics and thus defined coordinate-wise and overall indirect effects that are coherent in their non-parametric definitions. We also showed that while the coordinate-wise indirect effects depend on the chosen partition matrix, the overall indirect effect is invariant of this choice. Furthermore, these effects are based on a single causal mediation model and their empirical derivation does not require multiple testing, regularisation or model selection. Third, unlike in other simulation studies on microbial mediation, we investigated the effect of sparsity and overdispersion on the estimation of the mediated effects.

In addition to the analysis of the microbiome, the approach adopted in this study is applicable for a wide range of research questions where the mediator can be considered compositional. For example, the method could be utilised in other types of biological read-count data, such as epigenetic RNA counts. Although RNA molecules do not follow similar taxonomic phylogenies as the microbiome, they can be classified e.g. based on their type [33]. Such classifications can be utilised as *a priori* knowledge to formulate meaningful contrasts. Compositional data analysis, and the proposed approach, also provide opportunities within the context of e.g. dietary composition, blood lipid composition or composition of the daily activity levels, and using *a priori* knowledge of the hierarchies between the parts of composition may also provide interesting insights into the mediating role of these compositions. However, the applicability of compositional transformations under different sources of variation and different types of data still needs further research.

Acknowledgments

NK has been financially supported by Emil Aaltonen Foundation and the MATTI programme in The University of Turku Graduate School (UTUGS). The STRIP study has been financially supported by the Academy of Finland (grants 206374, 294834, 251360, 275595, 307996, and 322112), the Juho Vainio Foundation, the Finnish Foundation for Cardiovascular Research, the Finnish Ministry of Education and Culture, the Finnish Cultural Foundation, the Sigrid Jusélius Foundation, Special Governmental grants for Health Sciences Research (Turku University Hospital), the Yrjö Jahnsson Foundation, the Finnish Medical Foundation, and the Turku University Foundation.

Declaration of interest

The authors declare no potential conflict of interests.

Data availability statement

The details of the simulation study, including the extensive simulation codes, are available upon request from the corresponding author. The empirical dataset comprises health related participant data and their use is therefore restricted under the regulations on professional secrecy (Act on the Openness of Government Activities, 612/1999) and on sensitive personal data (Personal Data Act, 523/1999, implementing the EU data protection directive 95/46/EC). Due to these legal restrictions, the data from this study (can not be stored in public repositories or otherwise made publicly available. However, data access may be permitted on a case by case basis upon request only. Data sharing requires a data-sharing agreement. Investigators can submit an expression of interest to the corresponding author.

Financial disclosure

There are no financial conflicts of interest to disclose.

Ethical approval and patient consent

The empirical study was approved by the associated university and hospital district ethical authorities. Written informed consent was obtained from the participants.

References

[1] G. B. Gloor, J. M. Macklaim, V. Pawlowsky-Glahn, and J. J. Egozcue, "Microbiome datasets are compositional: And this is not optional," *Frontiers in Microbiology*, vol. 8, no. NOV, pp. 1–6, 2017.

[2] J. D. Silverman, A. D. Washburne, S. Mukherjee, and L. A. David, "A phylogenetic transform enhances analysis of compositional microbiota data," *eLife*, vol. 6, pp. 1–20, 2017.

[3] V. Pawlowsky-Glahn, J. J. Egozcue, and R. Tolosana-Delgado, *Modeling and Analysis of Compositional data*. John Wiley & Sons, Ltd, 2015.

[4] V. Pawlowsky-Glahn and A. Buccianti, *Compositional data analysis - Theory and Applications*. 2011.

[5] J. Graffelman, "A parametric test for HWE based on isometric logratio coordi- nates," pp. 1–5, 2011.

[6] E. Gordon-Rodriguez, T. P. Quinn, and J. P. Cunningham, "Learning Sparse Log-Ratios for High-Throughput Sequencing Data," *bioRxiv*, p. 2021.02.11.430695, 2021.

[7] H. Zhang, J. Chen, Z. Li, and L. Liu, "Testing for Mediation Effect with Application to Human Microbiome Data," *Statistics in Biosciences*, 2019.

[8] J. Pearl, "Interpretation and identification of causal mediation," *Psychological Methods*, vol. 19, no. 4, pp. 459–481, 2014.

[9] T. J. VanderWeele, "Mediation Analysis: A Practitioner's Guide," *Annual Review of Public Health*, vol. 37, no. 1, pp. 17–32, 2016.

[10] M. B. Sohn and H. Li, "Compositional mediation analysis for microbiome studies," *Annals of Applied Statistics*, vol. 13, no. 1, pp. 661–681, 2019.

[11] C. Wang, J. Hu, M. J. Blaser, H. Li, and I. Birol, "Estimating and testing the microbial causal mediation effect with high-dimensional and compositional microbiome data," *Bioinformatics*, vol. 36, no. 2, pp. 347–355, 2020.

[12] J. Fu, M. Koslovsky, A. Neophytou, and M. Vannucci, "A Bayesian joint model for compositional mediation effect selection in microbiome data.," *Statistics in Medicine*, vol. 42, no. 17, pp. 2999–3015, 2023.

[13] M. G. Blum, L. Valeri, O. François, S. Cadiou, V. Siroux, J. Lepeule, and R. Slama, "Challenges raised by mediation analysis in a high-dimension setting," *Environmental Health Perspectives*, vol. 128, no. 5, pp. 1–8, 2020.

[14] T. J. VanderWeele and S. Vansteelandt, "Mediation Analysis with Multiple Mediators," *Epidemiologic Methods*, vol. 2, no. 1, pp. 95–115, 2014.

[15] R. Daniel, B. De Stavola, S. Cousens, and S. Vansteelandt, "Causal mediation analysis with multiple mediators.," *Biometrics*, vol. 71, no. 1, p. 1:14, 2015.

[16] K. Imai and T. Yamamoto, "Identification and sensitivity analysis for multiple causal mechanisms: revisiting evidence from framing experiments," *Political Analysis*, vol. 21, no. 2, pp. 141–161, 2013.

[17] W. Wang, S. Nelson, and J. M. Albert, "Estimation of Causal Mediation Effects for a Dichotomous Outcome in Multiple-Mediator Models using the Mediation Formula," *Statistics in Medicine*, vol. 32, no. 24, pp. 4211–4228, 2013.

[18] C. Kim, M. Daniels, J. Hogan, C. Choirat, and C. Zigler, "Bayesian methods for multiple mediators: relating principal stratification and causal mediation in the analysis of power plant emission controls," *Annals of Applied Statistics*, vol. 13, no. 3, pp. 1927–1956, 2019.

[19] M. Taguri, J. Featherstone, and J. Cheng, "Causal mediation analysis with multiple causally non-ordered mediators," *Statistical Methods in Medical Research*, vol. 27, no. 1, pp. 3–19, 2018.

[20] T. M. Filzmoser P, Hron K, *Applied Compositional Data Analysis.* Springer Nature Switzerland AG, 2018.

[21] L. Barabesi and P. Sartorelli, *Compositional data analysis*. 2020.

[22] J. Pearl, "Direct and Indirect Effects," *Proceedings of the Seventeenth Conference on Uncertainy in Artificial Intelligence*, pp. 411–20, 2001.

[23] T. M. Barber, S. Kabisch, A. F. Pfeiffer, and M. O. Weickert, "The health benefits of dietary fibre," *Nutrients*, vol. 12, no. 10, pp. 1–17, 2020.

[24] M. Gurung, Z. Li, H. You, R. Rodrigues, D. B. Jump, A. Morgun, and N. Shulzhenko, "Role of gut microbiota in type 2 diabetes pathophysiology," 2020.

[25] E. S. Chambers, T. Preston, G. Frost, and D. J. Morrison, "Role of Gut Microbiota-Generated Short-Chain Fatty Acids in Metabolic and Cardiovascular Health," *Current Nutrition Reports*, vol. 7, no. 4, pp. 198–206, 2018.

[26] K. Pahkala, T. T. Laitinen, H. Niinikoski, N. Kartiosuo, S. P. Rovio, H. Lagström, B. M. Loo, P. Salo, E. Jokinen, C. G. Magnussen, M. Juonala, O. Simell, A. Jula, T. Rönnemaa, J. Viikari, and O. T. Raitakari, "Effects of 20-year infancy-onset dietary counselling on cardiometabolic risk factors in the Special Turku Coronary Risk Factor Intervention Project (STRIP): 6-year post-intervention follow-up," *The Lancet Child and Adolescent Health*, vol. 4, no. 5, pp. 359–369, 2020.

[27] T. Ronnemaa, H. Lagstrom, A. Jula, P. Hakala, J. Viikari, O. Simell, M. Laurinen, H. Niinikoski, I. Valimaki, O. T. Raitakari, E. Jokinen, and M. Aromaa, "Cohort Profile: The STRIP Study (Special Turku Coronary Risk Factor Intervention Project), an Infancy-onset Dietary and Life-style Intervention Trial," *International Journal of Epidemiology*, vol. 38, no. 3, pp. 650–655, 2008.

[28] B. J. Callahan, P. J. McMurdie, M. J. Rosen, A. W. Han, A. J. A. Johnson, and S. P. Holmes, "DADA2: High-resolution sample inference from Illumina amplicon data," *Nature Methods*, vol. 13, no. 7, pp. 581–583, 2016.

[29] P. Hakala, J. Marniemi, L. R. Knuts, J. Kumpulainen, R. Tahvonen, and S. Plaami, "Calculated vs analysed nutrient composition of weight reduction diets," *Food Chemistry*, vol. 57, no. 1, pp. 71–75, 1996.

[30] V. Ravitsemusneuvottelukunta, *Terveyttä ruoasta - Suomalaiset ravitsemussuositukset 2014 (Finnish Nutrition Recommendations 2014) Valtion ravitsemusneuvottelukunta: Tampere.* 2014.

[31] T. VanderWeele, "Bias formulas for sensitivity analysis for direct and indirect effects.," *Epidemiology*, vol. 21, no. 4, pp. 540–51, 2010.

[32] T. VanderWeele and O. Arah, "Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders.," *Epidemiology*, vol. 22, no. 1, pp. 42–52, 2011.

[33] K. M. Hombach S, "Non-coding RNAs: Classification, Biology and Functioning.," *Adv Exp Med Biol.*, no. 937, pp. 3–17, 2016.

Tables

Table 1: Parameters of the simulation study.

Parameter	Value	Note
Distribution of the counts		
E(K_i)= μ	10000.	
Dispersion θ	$\to$ 0; 0.1; 0.5	
Sparsity α_S	1; 50; $\alpha_S \to \infty$	
Binary confounders		
$p(C_1 = 1)$	0.50	
$p(C_2 = 1)$	0.50	
Binary exposure		
$p(X = 1\|C_1, C_2)$	$0.25 + 0.05C_1 + 0.05C_2$	
Mediator		
	(0.02, 0.01, -0.01, -0.01, -0.01)	Effect of C_1 on $\boldsymbol{\alpha}$
	(-0.01, -0.01, 0.02, 0.00, 0.00)	Effect of C_2 on $\boldsymbol{\alpha}$
$(\alpha_1/\alpha_S, \ldots, \alpha_5/\alpha_S)\|X = 0$		
	(0.60, [0.15, 0.10], [0.10, 0.05])	scenario 1
	(0.60, [0.15, 0.10], [0.08, 0.07])	scenario 2
	(0.60, 0.10, 0.10, 0.10, 0.10)	scenario 3
$(\alpha_1/\alpha_S, \ldots, \alpha_5/\alpha_S)\|X = 1$		
	(0.50, [0.20, 0.05], [0.10, 0.15])	scenario 1
	(0.58, [0.13, 0.13], [0.10, 0.06])	scenario 2
	(0.50, 0.125, 0.125, 0.125, 0.125)	scenario 3
Response		
$(\gamma_0, \gamma_2, \overline{\gamma_{C_1}, \gamma_{C_1}})$	(2.0, 0.40, 0.05, -0.05)	
Indirect effects		
OIE; CIE	0.10; (0.04, 0.01, 0.03, 0.02)	scenario 1
	0.10; (0.04, 0.01, 0.03, 0.02)	scenario 2
	0.10; (0.10, 0.00, 0.00, 0.00)	scenario 3

Table 2: Simulation results for scenario 1. The true and estimated indirect effects, bias, average standard errors ($\widehat{\text{SE}}$), standard deviation of the estimates ($\text{SE}_{\text{est.}}$), and power and coverage probability of the 90 % confidence intervals of the coordinate-wise (CIE) and overall (OIE) indirect effects with varying sparsity α_S and overdispersion θ.

α_S	θ	**Eff.**	**True**	**Est.**	**Bias**	$\widehat{\text{SE}}$	$\text{SE}_{\text{est.}}$	**Power**	**Coverage**
1	$\to 0$	CIE_1	0.04	0.04	0.00	0.015	0.015	0.91	0.91
1	$\to 0$	CIE_2	0.01	0.01	0.00	0.017	0.016	0.12	0.93
1	$\to 0$	CIE_3	0.03	0.03	0.00	0.017	0.016	0.56	0.93
1	$\to 0$	CIE_4	0.02	0.02	0.00	0.020	0.020	0.27	0.91
1	$\to 0$	OIE	0.10	0.10	0.00	0.036	0.035	0.90	0.92
50	$\to 0$	CIE_1	0.04	0.04	0.00	0.044	0.044	0.24	0.91
50	$\to 0$	CIE_2	0.01	0.01	0.00	0.061	0.060	0.09	0.91
50	$\to 0$	CIE_3	0.03	0.03	0.00	0.057	0.058	0.16	0.89
50	$\to 0$	CIE_4	0.02	0.02	0.00	0.055	0.054	0.12	0.91
50	$\to 0$	OIE	0.10	0.10	0.00	0.089	0.091	0.30	0.88
$\to \infty$	$\to 0$	CIE_1	0.04	0.05	-0.01	0.592	0.595	0.11	0.89
$\to \infty$	$\to 0$	CIE_2	0.01	0.00	0.01	0.774	0.801	0.11	0.89
$\to \infty$	$\to 0$	CIE_3	0.03	0.03	0.00	0.742	0.737	0.11	0.90
$\to \infty$	$\to 0$	CIE_4	0.02	0.03	-0.01	0.728	0.700	0.09	0.91
$\to \infty$	$\to 0$	OIE	0.10	0.11	-0.01	1.212	1.202	0.10	0.91
1	0.1	CIE_1	0.04	0.04	0.00	0.015	0.015	0.89	0.90
1	0.1	CIE_2	0.01	0.01	0.00	0.017	0.016	0.13	0.92
1	0.1	CIE_3	0.03	0.03	0.00	0.017	0.017	0.55	0.91
1	0.1	CIE_4	0.02	0.02	0.00	0.020	0.019	0.27	0.91
1	0.1	OIE	0.10	0.10	0.00	0.036	0.035	0.88	0.91
50	0.1	CIE_1	0.04	0.04	0.00	0.043	0.043	0.22	0.92
50	0.1	CIE_2	0.01	0.01	0.00	0.061	0.059	0.09	0.91
50	0.1	CIE_3	0.03	0.03	0.00	0.057	0.056	0.13	0.90
50	0.1	CIE_4	0.02	0.02	0.00	0.055	0.053	0.12	0.92
50	0.1	OIE	0.10	0.10	0.00	0.090	0.090	0.29	0.89
$\to \infty$	0.1	CIE_1	0.04	0.05	-0.01	0.563	0.566	0.10	0.90
$\to \infty$	0.1	CIE_2	0.01	0.00	0.01	0.736	0.746	0.11	0.89
$\to \infty$	0.1	CIE_3	0.03	0.04	-0.01	0.707	0.726	0.11	0.89
$\to \infty$	0.1	CIE_4	0.02	0.00	0.02	0.692	0.683	0.09	0.91
$\to \infty$	0.1	OIE	0.10	0.09	0.01	1.152	1.189	0.11	0.89
1	0.5	CIE_1	0.04	0.04	0.00	0.015	0.015	0.90	0.90
1	0.5	CIE_2	0.01	0.01	0.00	0.017	0.017	0.14	0.91
1	0.5	CIE_3	0.03	0.03	0.00	0.017	0.017	0.55	0.91
1	0.5	CIE_4	0.02	0.02	0.00	0.020	0.020	0.26	0.89
1	0.5	OIE	0.10	0.10	0.00	0.036	0.036	0.89	0.88
50	0.5	CIE_1	0.04	0.04	0.00	0.043	0.043	0.22	0.91
50	0.5	CIE_2	0.01	0.01	0.00	0.061	0.060	0.09	0.91
50	0.5	CIE_3	0.03	0.03	0.00	0.057	0.057	0.13	0.90
50	0.5	CIE_4	0.02	0.02	0.00	0.055	0.054	0.11	0.91
50	0.5	OIE	0.10	0.09	0.01	0.089	0.087	0.27	0.91
$\to \infty$	0.5	CIE_1	0.04	0.05	-0.01	0.431	0.427	0.09	0.91
$\to \infty$	0.5	CIE_2	0.01	0.01	0.00	0.562	0.578	0.12	0.89
$\to \infty$	0.5	CIE_3	0.03	0.03	0.00	0.541	0.565	0.11	0.89
$\to \infty$	0.5	CIE_4	0.02	0.03	-0.01	0.528	0.510	0.08	0.92
$\to \infty$	0.5	OIE	0.10	0.11	-0.01	0.879	0.868	0.10	0.90

Table 3: Simulation results for scenario 2. The true and estimated indirect effects, bias, average standard errors ($\widehat{\text{SE}}$), standard deviation of the estimates ($\text{SE}_{\text{est.}}$), and power and coverage probability of the 90 % confidence intervals of the coordinate-wise (CIE) and overall (OIE) indirect effects with varying sparsity α_S and overdispersion θ.

α_S	θ	**Eff.**	**True**	**Est.**	**Bias**	$\widehat{\text{SE}}$	$\text{SE}_{\text{est.}}$	**Power**	**Coverage**
1	$\to 0$	CIE_1	0.04	0.04	0.00	0.034	0.034	0.33	0.91
1	$\to 0$	CIE_2	0.01	0.00	0.01	0.102	0.102	0.10	0.90
1	$\to 0$	CIE_3	0.03	0.03	0.00	0.013	0.012	0.81	0.90
1	$\to 0$	CIE_4	0.02	0.02	0.00	0.011	0.010	0.57	0.93
1	$\to 0$	OIE	0.10	0.09	0.01	0.110	0.109	0.21	0.91
50	$\to 0$	CIE_1	0.04	0.04	0.00	0.015	0.015	0.93	0.91
50	$\to 0$	CIE_2	0.01	0.01	0.00	0.015	0.014	0.14	0.92
50	$\to 0$	CIE_3	0.03	0.03	0.00	0.024	0.024	0.33	0.91
50	$\to 0$	CIE_4	0.02	0.02	0.00	0.019	0.018	0.26	0.92
50	$\to 0$	OIE	0.10	0.10	0.00	0.037	0.034	0.89	0.92
$\to \infty$	$\to 0$	CIE_1	0.04	0.05	-0.01	0.123	0.122	0.12	0.91
$\to \infty$	$\to 0$	CIE_2	0.01	0.01	0.00	0.026	0.026	0.14	0.90
$\to \infty$	$\to 0$	CIE_3	0.03	0.03	0.00	0.320	0.328	0.11	0.89
$\to \infty$	$\to 0$	CIE_4	0.02	0.02	0.00	0.244	0.238	0.09	0.91
$\to \infty$	$\to 0$	OIE	0.10	0.11	-0.01	0.417	0.419	0.12	0.90
1	0.1	CIE_1	0.04	0.04	0.00	0.035	0.037	0.34	0.88
1	0.1	CIE_2	0.01	0.02	-0.01	0.106	0.105	0.10	0.90
1	0.1	CIE_3	0.03	0.03	0.00	0.013	0.012	0.83	0.92
1	0.1	CIE_4	0.02	0.02	0.00	0.011	0.010	0.57	0.92
1	0.1	OIE	0.10	0.11	-0.01	0.114	0.113	0.23	0.90
50	0.1	CIE_1	0.04	0.04	0.00	0.015	0.015	0.91	0.89
50	0.1	CIE_2	0.01	0.01	0.00	0.014	0.014	0.15	0.91
50	0.1	CIE_3	0.03	0.03	0.00	0.025	0.024	0.36	0.90
50	0.1	CIE_4	0.02	0.02	0.00	0.020	0.019	0.25	0.92
50	0.1	OIE	0.10	0.10	0.00	0.037	0.036	0.85	0.90
$\to \infty$	0.1	CIE_1	0.04	0.04	0.00	0.117	0.113	0.11	0.91
$\to \infty$	0.1	CIE_2	0.01	0.01	0.00	0.025	0.025	0.12	0.90
$\to \infty$	0.1	CIE_3	0.03	0.02	0.01	0.305	0.307	0.10	0.90
$\to \infty$	0.1	CIE_4	0.02	0.02	0.00	0.232	0.230	0.10	0.90
$\to \infty$	0.1	OIE	0.10	0.09	0.01	0.397	0.391	0.10	0.90
1	0.5	CIE_1	0.04	0.04	0.00	0.035	0.034	0.33	0.92
1	0.5	CIE_2	0.01	0.01	0.00	0.128	0.128	0.10	0.91
1	0.5	CIE_3	0.03	0.03	0.00	0.013	0.012	0.84	0.93
1	0.5	CIE_4	0.02	0.02	0.00	0.011	0.011	0.56	0.91
1	0.5	OIE	0.10	0.10	0.00	0.136	0.134	0.20	0.90
50	0.5	CIE_1	0.04	0.04	0.00	0.015	0.015	0.91	0.90
50	0.5	CIE_2	0.01	0.01	0.00	0.015	0.014	0.14	0.93
50	0.5	CIE_3	0.03	0.03	0.00	0.025	0.024	0.34	0.92
50	0.5	CIE_4	0.02	0.02	0.00	0.019	0.020	0.28	0.91
50	0.5	OIE	0.10	0.10	0.00	0.037	0.036	0.88	0.91
$\to \infty$	0.5	CIE_1	0.04	0.04	0.00	0.090	0.088	0.12	0.92
$\to \infty$	0.5	CIE_2	0.01	0.01	0.00	0.020	0.018	0.11	0.92
$\to \infty$	0.5	CIE_3	0.03	0.04	-0.01	0.234	0.229	0.10	0.91
$\to \infty$	0.5	CIE_4	0.02	0.02	0.00	0.177	0.178	0.12	0.89
$\to \infty$	0.5	OIE	0.10	0.11	-0.01	0.304	0.299	0.11	0.90

Table 4: Empirical variances of the estimators of β_{11} and γ_{11} and their product under different scenarios and varying degrees of overdispersion (θ) and sparsity (α_S). The last two columns present the ratios between the squared estimators and the variances for for the first coordinate wise indirect effect (CIE_1). For each setting, the values are based on 1000 replications of simulated data.

Scenario	θ	α_S	β_{11}	γ_{11}	$\hat{\sigma}^2_{\beta_1}$	$\hat{\sigma}^2_{\gamma_1}$	$\hat{\sigma}^2_{\beta_1}\hat{\sigma}^2_{\gamma_1}$	$\beta^2_{11}/\hat{\sigma}^2_{\beta_1}$	$\gamma^2_{11}/\hat{\sigma}^2_{\gamma_1}$	var(IE)
1	$\to 0$	1	-0.87	-0.05	0.0263	0.0002	0.000005	28.72	11.46	0.0002
1	$\to 0$	50	-0.36	-0.11	0.0004	0.0139	0.000006	323.99	0.86	0.0018
1	$\to 0$	$\to \infty$	-0.34	-0.15	0.0000	3.0048	0.000005	64007.84	0.01	0.3494
1	0.1	1	-0.86	-0.05	0.0262	0.0002	0.000005	28.24	11.31	0.0002
1	0.1	50	-0.36	-0.10	0.0004	0.0139	0.000006	323.92	0.75	0.0018
1	0.1	$\to \infty$	-0.34	-0.15	0.0000	2.7107	0.000005	57500.11	0.01	0.3152
1	0.5	1	-0.85	-0.05	0.0258	0.0002	0.000005	28.21	11.33	0.0002
1	0.5	50	-0.36	-0.10	0.0004	0.0138	0.000006	319.80	0.79	0.0018
1	0.5	$\to \infty$	-0.34	-0.14	0.0000	1.5932	0.000006	31618.72	0.01	0.1853
2	$\to 0$	1	-0.19	-0.21	0.0252	0.0002	0.000005	1.46	226.34	0.0011
2	$\to 0$	50	-0.07	-0.55	0.0004	0.0142	0.000005	14.31	21.43	0.0002
2	$\to 0$	$\to \infty$	-0.07	-0.64	0.0000	3.0710	0.000006	2603.26	0.13	0.0151
2	0.1	1	-0.19	-0.22	0.0290	0.0002	0.000006	1.27	234.20	0.0014
2	0.1	50	-0.07	-0.54	0.0004	0.0143	0.000005	13.90	20.49	0.0002
2	0.1	$\to \infty$	-0.07	-0.53	0.0000	2.7673	0.000006	2459.66	0.10	0.0136
2	0.5	1	-0.19	-0.22	0.0245	0.0002	0.000005	1.49	233.19	0.0011
2	0.5	50	-0.07	-0.54	0.0004	0.0142	0.000005	14.39	20.89	0.0002
2	0.5	$\to \infty$	-0.07	-0.51	0.0000	1.6191	0.000006	1418.24	0.16	0.0080
3	$\to 0$	1	-0.95	-0.11	0.0274	0.0002	0.000005	32.96	61.93	0.0005
3	$\to 0$	50	-0.39	-0.26	0.0003	0.0145	0.000005	435.50	4.82	0.0022
3	$\to 0$	$\to \infty$	-0.36	-0.27	0.0000	3.1057	0.000005	82809.50	0.02	0.4133
3	0.1	1	-0.94	-0.11	0.0272	0.0002	0.000005	32.45	62.23	0.0005
3	0.1	50	-0.39	-0.26	0.0003	0.0145	0.000005	429.88	4.61	0.0022
3	0.1	$\to \infty$	-0.36	-0.28	0.0000	2.8001	0.000005	74406.43	0.03	0.3724
3	0.5	1	-0.93	-0.11	0.0270	0.0002	0.000005	32.00	63.13	0.0005
3	0.5	50	-0.39	-0.25	0.0003	0.0144	0.000005	432.93	4.46	0.0022
3	0.5	$\to \infty$	-0.37	-0.29	0.0000	1.6401	0.000005	40849.87	0.05	0.2185

Table 5: Average counts and proportions of the 11 genera within the phylum *Actinobacteria* by fibre intake (low, i.e., < 25 grams per day vs. high, i.e. $\geq$ 25 grams per day). The data are based on the STRIP study [26].

	Count		**Proportion**	
Genus	**Fibre** < 25 g/d	**Fibre** $\geq$ 25 g/d	**Fibre**< 25 g/d	**Fibre**$\geq$ 25 g/d
Bifidobacterium	3780.22	1865.53	66.27	62.27
Actinomyces	3.30	1.48	0.18	0.16
Rothia	2.04	1.44	0.15	0. 15
Enterorhabdus	8.56	4.56	0. 41	0.56
Eggerthella	52.95	29.96	3.00	2.60
Collinsella	697.60	433.02	20.95	25.04
Adlercreutzia	95.59	83.05	5.22	6.14
Slackia	22.42	21.24	1.35	1.29
Senegalimassilia	31.26	7.61	1.29	0.88
Olsenella	14.34	7.88	0.79	0.55
Gordonibacter	7.32	4.95	0.39	0.36
Total	4715.61	2460.73	100	100

Table 6: Coefficients for the effect of fibre on the ilr coordinates and coefficients for the effects of the ilr coordinates on the response, and the coordinate-wise indirect effects (CIE's) $\beta_{1k}\gamma_{1k}$ for each coordinate in the analyses for mediation through the genera within the Actinobacteria (section 3.2.3).

	$\hat{\beta}_{1k}$	(90% CI)	$\hat{\gamma}_{1k}$	(90% CI)	$\hat{\beta}_{1k}\hat{\gamma}_{1k}$ (s.e.)	(90% CI)
M_1	-0.025	[-0.341; 0.291]	0.031	[-0.006; 0.067]	-0.001 (0.006)	[-0.010; 0.009]
M_2	-0.255	[-0.501; -0.009]	-0.031	[-0.078; 0.016]	0.008 (0.009)	[-0.006; 0.022]
M_3	-0.269	[-0.740; 0.201]	0.025	[0.002; 0.048]	-0.007 (0.008)	[-0.020; 0.007]
M_4	0.409	[0.060; 0.759]	-0.041	[-0.074; -0.008]	-0.017 (0.012)	[-0.036; 0.003]
M_5	-0.176	[-0.712; 0.359]	0.038	[0.015; 0.060]	-0.007 (0.012)	[-0.027; 0.014]
M_6	0.264	[-0.220; 0.749]	0.011	[-0.013; 0.035]	0.003 (0.005)	[-0.005; 0.011]
M_7	0.071	[-0.414; 0.557]	0.001	[-0.024; 0.027]	0.000 (0.001)	[-0.002; 0.002]
M_8	0.012	[-0.434; 0.458]	0.014	[-0.010; 0.038]	0.000 (0.004)	[-0.006; 0.006]
M_9	-0.165	[-0.589; 0.259]	0.017	[-0.010; 0.044]	-0.003 (0.005)	[-0.011; 0.006]
M_{10}	-0.151	[-0.596; 0.295]	0.012	[-0.015; 0.039]	-0.002 (0.004)	[-0.009; 0.005]

Table 7: Comparison of different approaches to mediation analysis with compositional mediator in [7], [10], [11] and [12]. Abbreviations: SEM = Structural Equation Model; ilr = isometric logratio transformation; alr = additive logratio transformation; LM = linear regression model. By sieving we refer to testing for statistically significant signals in a large set of potential mediators. †: by overdispersion and sparsity, we refer to investigating the effect of these features on the estimation and performance in the simulation study.

	This study	**Zhang et al.**	**Sohn and Li**	**Wang**	**Fu**
Framework	Counterfactual	SEM	Counterfactual	Counterfactual	Counterfactual
Mediator	Ilr	Ilr	Alr	Alr	Taxa
Overall IE	Yes	No	Yes	Yes	Yes
Mediator-wise IE	Yes	Yes	Yes	Yes	Yes
Zero counts	Replace with 0.5	Replace with 0.5	Replace with 0.5	Dirichlet regression	Replace with 0.5
Overdispersion†	Yes	No	No	No	No
Sparsity†	Yes	No	No	No	No
Dimension	Low	High	High	High	High
Regularization	No	de-biased Lasso	No	L1 penalty	Bayesian priors
Space	Euclidean	Euclidean	Simplex	Not specified	Not specified
Taxonomic hierarchy	Yes	No	No	Possible	No
Model for mediator	LM	LM	Compositional algebra	Dirichlet regression	Log-linear regression
Model for response	LM	LM	Linear log-contrast	Linear log-contrast	LM
Aim	Estimation of causal effects	Sieving	Estimation, sieving	Sieving	Estimation, sieving

Figures

List of figure captions

Figure 1 Directed acyclic graphs (DAGs) depicting mediation between exposure X and response Y through (a) a single mediator, and (b) a compositional mediator, such as the gut microbiome.

Figure 2 An example of a taxonomic hierarchy among five classes, i.e., five distinct taxa. (a) A taxonomy describing a known relational structure between five taxa. (b) An example of a hierarchy based on taxonomy (a) and encoded by the sequential binary partition matrix 1 of the main text.

Figure 3 DAGs describing the contrasted parts of the composition in models 3A and 3B. The correct causal DAG used in simulating data is presented in (a) and the DAG encoded by the misspecified contrast matrix is presented in (b).

Figure 4 Taxonomic structure within the *Actinobacteria* phylum. Within the phylum, there is one class, which was further divided into three orders consisting of eleven genera within four families.

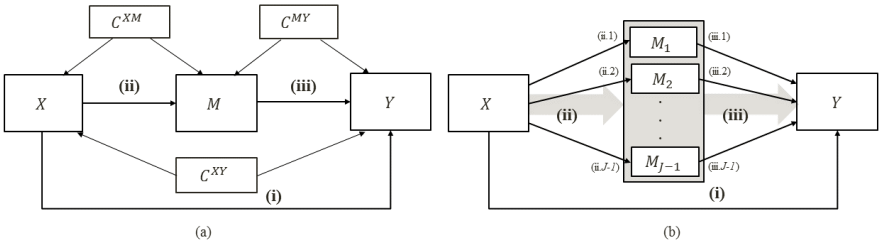

Figure 1: Directed acyclic graphs (DAGs) depicting mediation between exposure X and response Y through (a) a single mediator, and (b) a compositional mediator, such as the gut microbiome. Potential confounders are denoted as C^{XY}, C^{XM} and C^{MY} in DAG (a) and omitted for clarity in DAG (b).

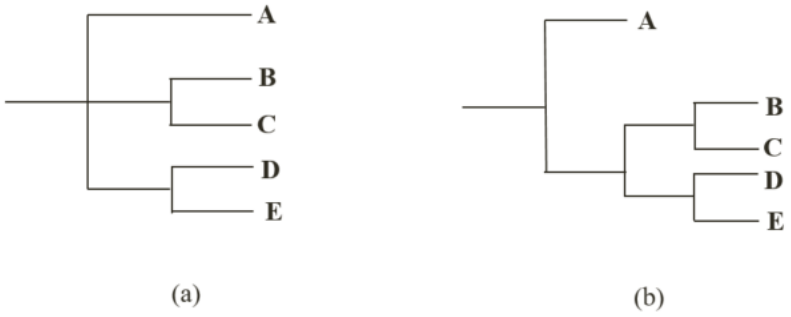

Figure 2: An example of a taxonomic hierarchy among five classes, i.e., five distinct taxa. (a) A taxonomy describing a known relational structure between five taxa. (b) An example of a hierarchy based on taxonomy (a) and encoded by the sequential binary partition matrix 1 of the main text.

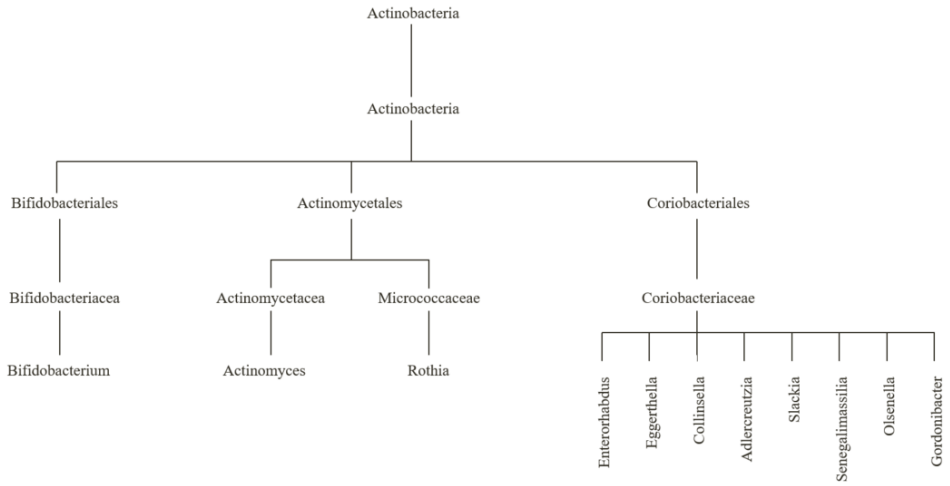

Figure 4: Taxonomic structure within the *Actinobacteria* phylum. Within the phylum, there is one class, which was further divided into three orders consisting of eleven genera within four families.

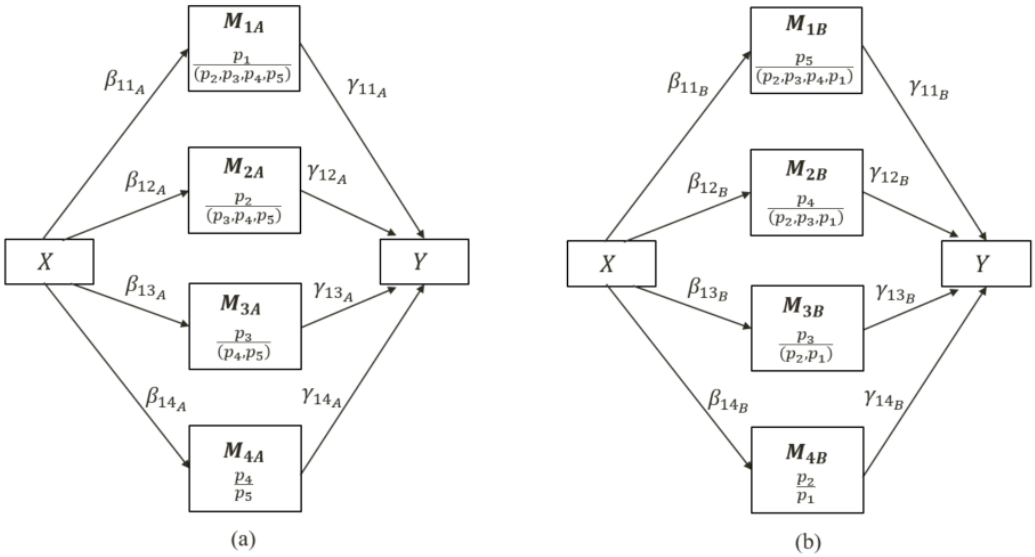

Figure 3: DAGs describing the contrasted parts of the composition in models 3A and 3B. The correct causal DAG used in simulating data is presented in (a) and the DAG encoded by the misspecified contrast matrix is presented in (b).