

**Novel study designs and modelling approaches in nutrition research: application of advanced statistical modelling approaches and N-of-1 study designs**

Edward Payne<sup>1</sup>, Rute Vieira<sup>2</sup> and Baukje de Roos<sup>1</sup>

<sup>1</sup>The Rowett Institute, University of Aberdeen, Foresterhill, Aberdeen AB21 2ZD, United Kingdom

<sup>2</sup>Institute of Applied Health Sciences, Foresterhill, Aberdeen AB21 2ZD, United Kingdom

**Short title:** Novel study designs in nutrition research



This is an Accepted Manuscript for Proceedings of the Nutrition Society. This peer-reviewed article has been accepted for publication but not yet copyedited or typeset, and so may be subject to change during the production process. The article is considered published and may be cited using its

DOI 10.1017/S0029665124007602

Proceedings of the Nutrition Society is published by Cambridge University Press on behalf of The Nutrition Society.

## Abstract

Randomised controlled trials are the ‘gold standard’ approach in nutrition research to show a causal relationship between a dietary intervention and clinically relevant outcomes at the population level. Here we review why different study designs are needed to establish the efficacy of dietary interventions at the individual level, and to better account for relevant factors that can also influence the outcomes. Over the past decade, precision nutrition approaches have been developed as a new way to measure the effectiveness of dietary interventions at the individual and population level. Precision nutrition aims to determine the individual factors that are associated with differences in responses to dietary interventions. This is complex, typically needing studies with a large number of participants, and using advanced statistical approaches and machine-learning algorithms to identify predictors that can explain why individuals do or do not respond to consuming specific foods, meals, or diets, for a given outcome. N-of-1 study designs, which are new to nutrition science, offer a robust alternative approach to assess how an intervention, and/or everyday behaviours, affect individual health outcomes. They utilise repeated measures within individuals, rather than baseline and end measures in a larger number of participants, to provide the statistical power required to determine an individual's responsiveness to an intervention. The adoption of new study designs and modelling approaches, particularly the N-of-1 approach, to examine responses to interventions within individuals, will help to further the understanding of the relationships between diet and health within individuals more effectively and accurately.

**Key words:** randomised controlled trials, N-of-1 trials, variable selection methods, predictors of response

Randomised controlled trials (RCTs) are the ‘gold standard’ approach to identify associations between a dietary intervention and one or more health or disease outcomes, on the population level. In principle, the randomisation of participants between the intervention and control groups evenly distributes the characteristics of the participants, and helps to reduce bias and confounding introduced by other known and unknown factors. Any observed differences in the outcome(s) between the groups, should, in theory, reflect the effects of the interventions delivered to each group. However, in practice, all randomised trials of dietary interventions face some degree of bias, including different types of selection or sample bias, or measurement bias, because most are conducted in the real world, with a limited sample size<sup>(1)</sup>. A common source of bias is the participants’ baseline characteristics (e.g., age and health status), which are often poorly distributed between the intervention and control groups, independently of the randomisation technique used. Although more complex randomisation techniques (e.g. restricted or stratified) can be used, not all characteristics will have been measured, and there is a limit to how many variables the sample can be stratified to, inevitably leading to unbalanced groups with regard to one or more potentially relevant characteristics<sup>(2)</sup>.

Complex confounding factors that describe the environment, (epi)genetics, nutritional status, food behaviours, gut microbiota, and the metabolome, are also often left unmeasured. Not only can these be responsible for significant variations in responses to interventions between individuals at a specific point in time<sup>(3)</sup>, they may also cause fluctuations in the outcome (e.g. plasma triglycerides, blood pressure, and heart rate) in each participant over time. For example, when studying the effect of a dietary intervention on blood pressure and heart rate using an RCT design, and assessing these outcomes at the start and end of the study period for both groups, if the participants show elevated blood pressure and/or elevated heart rate, compared to baseline, this change could have resulted from a particularly salty dinner they ate the night before, being stressed, having just finished a strenuous bout of exercise, or even coping with a sudden change in outside temperature, rather than the effect of the dietary intervention (Figure 1). These unacknowledged variables, that could also affect the outcome, can lead to variability in responses unrelated to the interventions, potentially causing spurious associations. These same factors that cause variation in individual responses can, however, be exploited to assess, or predict, how individuals respond to dietary interventions using more fit-for-purpose study designs.

The field of precision nutrition ultimately aims to develop more comprehensive and dynamic nutritional recommendations based on individual variables, including genetics, the microbiome, metabolic profile, health status, physical activity, dietary patterns, food environment, and socioeconomic and psychosocial characteristics <sup>(4)</sup>. Such ambitions can be realised by exploring existing large datasets and by conducting new intervention studies using appropriate study designs. Here we review where these methods have been applied so far.

### **Modelling data to identify variables that can explain responses**

Statistical models quantify the strength of an association between several independent variables, and an outcome, providing an insight on how multiple drivers, such as baseline variables, work together to cause a change in an outcome. Modelling is an important tool, however, when a study takes measurements of, or records, numerous independent variables, it can be difficult to determine which should be included in the model. Variable selection refers to the process of choosing the most relevant independent variables to include in a regression model, which will help to improve model performance, and avoid over fitting by reducing the number of independent variables included <sup>(5)</sup>. For a statistical model to work, both the input data and the choice of model need to be considered; inaccurate or incomplete data, or a non-optimal method, will result in inaccurate results and erroneous associations.

Independent variables could be selected *a priori* based on existing evidence of potential associations with an outcome. However, if researchers are faced with an overwhelming number of variables, or if there is uncertainty about where associations may exist, statistical methods exist to assist with variable selection. The importance of variable selection was highlighted in a secondary data analysis study that investigated changes in concentrations of plasma triglycerides after fish oil supplementation <sup>(6)</sup>, using original data from the FINGEN study <sup>(7)</sup>. The FINGEN study had previously shown that males, and those with the apolipoprotein E4 genotype, had the highest triglyceride-lowering response to fish oil intervention. However, the authors of the FINGEN study also specifically reported on substantial response heterogeneity between participants <sup>(7,8)</sup>. In the secondary data analysis <sup>(6)</sup>, four variable selection and analysis methods – forward stepwise selection, backward stepwise selection, least absolute shrinkage and selection operator (LASSO), and the Boruta algorithm – were applied to five datasets imputed by multiple imputation, using a pooled regression on

a per-outcome basis. The final model, and, therefore, the optimal variable selection and analysis method for that outcome, was chosen based on the lowest validation set root mean squared error (RMSE). Different variable selection methods were found optimal for different outcomes, highlighting the need to consider the selection method. For example, models generated by LASSO identified higher baseline plasma insulin concentrations and lower pre-intervention triglyceride concentrations as the strongest predictors of plasma triglyceride change, after fish oil intervention. On the other hand, backward stepwise selection identified being older, and female, and having lower baseline levels of plasma EPA and DHA, as the strongest predictors of plasma EPA and DHA change, after fish oil intervention <sup>(6)</sup>.

To explore how certain individuals might respond to a given intervention, researchers are using increasingly advanced statistical approaches, and machine-learning algorithms. These are mostly based on the measurement of multiple outcomes and independent variables, often including extensive questionnaires and the application of omics technologies, in order to identify the strongest predictors of change in clinically relevant outcomes on an individual level. Thus far, most studies have focussed on postprandial blood glucose levels due to the robust existing technologies available to measure blood glucose in a continuous fashion for longer periods of time, which are relevant, for example, to manage glycaemic control in diabetes <sup>(9)</sup>. To ensure sufficient statistical power, such studies typically involve large cohorts containing hundreds of participants, and require advanced statistical approaches and machine-learning algorithms to be applied to datasets in which the number of participants is greater than the number of predictors. Pioneering work by Zeevi and colleagues <sup>(10)</sup> allowed the development and validation of personalised algorithms to reduce individual postprandial glycaemic responses. But is this personalised/precision concept working in terms of increasing the efficacy of diets to improve health outcomes? A subsequent large-scale intervention study (the Personal Diet study) applied these algorithms to provide a personalised diet targeting the postprandial glucose response in adults with abnormal glucose metabolism and obesity. However, they found that the personalised diets resulted in no more weight loss, and no greater reduction in glycaemic variability and HbA1c levels, compared with a generic low-fat diet, after 6 months <sup>(11)</sup>. A second study by Berry and colleagues, the PREDICT-1 study, identified the main predictors of individual glucose responses, including meal composition, genetics, meal content, serum glucose markers and the microbiome <sup>(12)</sup>. The results of an efficacy study of a personalised nutrition programme, based on the

algorithms of the PREDICT-1 study, was recently published <sup>(13)</sup>. There were some positive but modest results: an average weight loss of 2.5 kg over four months, but no changes in various other biomarkers, including blood pressure, insulin, glucose, and postprandial triglycerides. The control group was also perhaps not a control group in the strictest sense of the term - they were simply given standard dietary advice and a helpline to call. It should be noted, however, that weight loss is a very ambitious goal, with the outcome also influenced by a significant number of behavioural factors, which are rarely taken into account in prediction models. Also, it has been questioned whether postprandial glucose is a relevant health outcome, especially in those who do not have diabetes, and whether healthy participants would benefit from flattening their blood glucose curves <sup>(14,15)</sup>. Recently, a study showed that individual postprandial continuous glucose monitoring responses to duplicate meals were unreliable in adults without diabetes <sup>(16)</sup>.

### **The application of novel precision nutrition study designs**

Over the past decade, we have seen the development of novel precision nutrition approaches as an innovative way to measure the efficacy of dietary interventions to improve *individual* health outcomes, rather than, or in addition to, *population* health. It is important to realise that these novel precision nutrition approaches answer different research questions, and provide different levels of evidence, compared to RCTs. Precision nutrition approaches target the personal aspect of nutrition information and aim to give people a better idea of whether a dietary intervention or change will work for them. The higher effectiveness of a precision nutrition approach to achieve this is based on an expectation that certain individuals may particularly benefit from certain dietary interventions, depending on factors such as their environment, (epi)genetics, nutritional status, food behaviours, gut microbiota, and the metabolome <sup>(17)</sup>. But how good are we at investigating what influences the individual responses to personalised diets?

Within the precision nutrition context, the repeated collection of outcome and potential explanatory variable measurements enhances the precision of association estimates, and is key to understanding which factors affect a participant's response to an intervention <sup>(18)</sup>. For example, the PRECISE study assessed whether supplementation with bilberry and grape seed extract for 12 weeks improved cardiometabolic outcomes in individuals at risk of developing

Type 2 Diabetes Mellitus (T2DM), with 14 participants acting as their own controls using a placebo-controlled cross-over randomised design <sup>(19)</sup>. Over the course of 12 weeks, multiple measurements of glycated haemoglobin (HbA1c), 2-hour oral glucose tolerance tests (OGTTs), total, low- (LDL) and high-density lipoprotein (HDL) cholesterol were taken, and continuous blood glucose levels were assessed. No significant changes in any of these outcomes were observed between groups, but it was found that bilberry and grape seed extract significantly reduced ambulatory blood pressure over 24 hrs, to a level which is comparable with the effect of anti-hypertensive drug treatments. The significant reduction in blood pressure may have been an early marker of a beneficial effect on glucose metabolism – a recent individual participant meta-analysis assessing the effect of blood pressure lowering and risk of new-onset type 2 diabetes found that a systolic blood pressure reduction by 5 mm Hg reduced the risk of T2DM by 11% <sup>(20)</sup>. It was also found that eight of the fourteen participants were identified as blood pressure ‘responders’ to bilberry and grape seed extract. These responders had significantly higher levels of phenylpropionic and phenyllactic acids in their faecal samples, and a higher proportional abundance of Fusicatenibacter-related bacteria in their baseline stool samples, which may provide an explanation for the blood pressure response <sup>(19)</sup>. However, these results at the individual-level will need to be tested in further studies, as they were obtained using simple statistical methodological approaches (e.g., ANOVA) which do not account for the fact that multiple measurements and observations of the same individual over time are not independent and are highly correlated.

### **Novel study designs for nutrition research – the potential of N-of-1**

An N-of-1 study is the opposite of a large-scale study aimed to make inferences about a population. Instead, it is a study specifically designed to generate statistically robust results using data collected from an individual, with the goal of drawing conclusions that apply only to that person. There are two main types of N-of-1 studies: observational N-of-1 studies which simply monitor a participant over time without introducing an intervention, and interventional N-of-1 studies that compare one intervention with a baseline period, or compare two or more interventions, and test hypotheses; the focus of this review is on interventional studies. Typically, an interventional N-of-1 study takes repeated measurements and/or recordings, of both the dependent variable (outcome) and independent variables of interest, in real time, and in the context of day-to-day activities, to determine associations

between the independent variables and the outcome(s). There are different ways an interventional N-of-1 study can be designed, depending on the aims of the study and the reversibility of the intervention, with the possibility of using a design with a randomised sequence of treatments where the individual is their own control <sup>(21)</sup>.

### Comparing interventional N-of-1 designs and RCTs

A key distinction between an RCT, or other common group-level interventional designs such as cross-over designs, and an interventional N-of-1 study, is the research question one intends to answer. An RCT aims to understand the effect(s) of an intervention(s) within the recruited sample that could be generalised to the population represented by the sample. Conversely, the primary aim of an interventional N-of-1 study is to understand how an individual responds to different interventions, to understand whether personal behavioural and lifestyle factors also influence variation in the outcome(s), and potentially explore any interactions between these factors and the intervention(s).

The focus of an N-of-1 study is on the individual, and whilst studies have been published on single participants, series of N-of-1 studies are more common <sup>(22,23)</sup>. Generally, N-of-1 studies have recruited less than 10 participants, whereas current protocols are aiming to recruit around 20 participants <sup>(22)</sup>. Data from multiple N-of-1 studies can be aggregated in a meta-analysis to estimate the effect of the intervention(s) in a population assumed to be represented by the sample of individuals studied <sup>(24,25)</sup>, much like analysing independent RCTs in a meta-analysis. Based on various assumptions including a statistical significance level of 0.05, and a power of 80%, it has been suggested that data from up to approximately 45 individual N-of-1 studies might be required for such an analysis <sup>(26)</sup>. Data from N-of-1 trials can also be used to help predict the effect of a treatment in a future person, which would provide improved methods to personalise treatment for individual patients. However, this type of analysis needs to be able to accurately estimate variation in between-patient differences, and to do this may require up to 100 individuals to be studied in the first place <sup>(26)</sup>. In a medical context, it may not be possible to recruit so many patients <sup>(26)</sup>, however, in nutrition, where the general population may often be the target, inclusion and exclusion criteria might be less constrained, and, while undeniably challenging, a study of this size might be more achievable for the purpose of developing personalised nutrition strategies.



The flexibility offered by a series of N-of-1 interventional studies – their ability to answer research questions pertaining to the individual and the population – could make them more desirable than an RCT in certain contexts. By choosing appropriate control or placebo interventions, applying randomisation, counterbalancing, and blinding, as has been suggested in guidelines for designing N-of-1 studies <sup>(27)</sup>, interventional N-of-1 studies will be able to conclude whether any of the independent variables describing time, the intervention, the environment, or an individuals' behaviour over time, are associated with variation in the outcome(s). This is a distinct advantage of N-of-1 studies over RCTs.

While N-of-1 studies and RCTs have obvious differences, it is important to understand that certain methods and principles relevant to RCTs also apply to N-of-1 studies. The concept of statistical power still applies and is derived from the number of repeated measurements taken in an individual (per intervention if the study is an intervention study) rather than the number of people recruited. A minimum of 50 repeated measures per intervention has been suggested, and the potential for missing data should be considered <sup>(21)</sup>. Randomisation can be used to generate intervention delivery sequences, and to allocate participants to these sequences when multiple sequences are involved, with counterbalancing being an alternative or complementary method. Exerting more control over the order the interventions are delivered, could also control for potential confounding introduced through order effects – the effect of one treatment on the effect of the following treatment. Similarly, the implementation of washout periods between treatments should also be implemented to reduce or eliminate the risk of carryover effects, the residual effects of an intervention beyond its assigned duration, as in a crossover trial, although analytical solutions exist when it is not possible or convenient to include washout periods <sup>(28)</sup>. Other publications (see <sup>(27,29)</sup>) have described and discussed these points in a more comprehensive manner with regards to the underlying theory and practical applications.

### Designing an N-of-1 study

Most discussion surrounding intervention order in an N-of-1 study focusses on the comparison between several interventions: how long each intervention is delivered for within a block (also referred to as a cycle), how they are ordered within a block, and how multiple blocks are arranged across the course of a study. It is also apparent that a lot of published N-

of-1 studies compare only two treatments (see <sup>(22,30)</sup>), although comparisons of three treatments do exist <sup>(22,31)</sup>. In nutrition research, the need to compare more than two interventions may be required – for example, when researchers wish to study the effects of individual components of a diet, or components of a single dietary item. Based on the assumption that an appropriate length of time to deliver a specific intervention is one week, a doubly counterbalanced design for three interventions, that seems to instinctively follow on from the design AB-BA-BA-AB (a common example in which each letter denotes a unique intervention, and a hyphen denotes the transition between blocks), could be ABC-CBA-CBA-ABC. However, in this design, intervention B is never delivered consecutively, so to appropriately account for order effects, one might need to add two more blocks, e.g., ABC-CBA-CBA-ABC-ACB-BCA. This study is now at least 18 weeks long and given washout periods are yet to be incorporated, and that participants would have to complete more than one data point per day to achieve at least 50 data points per intervention, the burden of this study on the participant may be deemed too great. An alternative approach would be to combine randomisation and counterbalancing; a random selection of four (out of a possible six) orders of three treatments could be selected, and then counterbalanced such that no treatment is delivered in consecutive weeks, e.g., ABC-BCA-CBA-BAC. A participant would still have to complete at least two measurements per day to achieve at least 50 per intervention, however, as a result the study is, favourably, reduced to 12 weeks.

#### *Imputation of missing data in N-of-1 studies*

N-of-1 studies often aim to collect a large amount of data from many time points and many different variables from a single participant using ecological momentary assessments (EMA) <sup>(27,29)</sup>. EMA involves repeated sampling of current behaviours and experiences of a participant in real time, and in a natural environment, aiming to minimise recall bias and maximise ecological validity <sup>(32)</sup>. It is entirely possible that a final dataset may have rows of missing data associated with a point in time when a participant was unable to return the information requested. This may depend on numerous factors such as the methods of collecting data (remotely or at a site visit), the frequency of data collection, a participant's reduced engagement with the study at times, or a random occurrence that prohibited data collection. However, to be able to appropriately apply statistical methodologies to an N-of-1 dataset, that is inherently time series data, one requires a complete set of data, to ensure proper modelling

of the time structure and autocorrelation patterns. Therefore, these missing data points require imputation which could be performed in various ways. One could impute the missing data based solely on the principle of autocorrelation, the relationship between any given data point and those that preceded it, on a variable-by-variable basis; within R, a programming language for statistical computing and data visualisation, the addition of a package such as `imputeTS` could be used to achieve this <sup>(33)</sup>. However, imputation in this manner would assume that variables are independent of one another, an assumption that is often unrealistic, especially in an N-of-1 study that is collecting data on numerous aspects of an individual's life. There is also the risk that basing the imputation of missing values on those immediately prior could disproportionately emphasise the effect of autocorrelation during the analysis. An alternative approach would be to use multiple imputation. In R there are various packages available for carrying out multiple imputation such as `Amelia II` <sup>(34)</sup> and `MICE` <sup>(35)</sup>. Multiple imputation usually generates multiple complete datasets that can be analysed together using functions provided by the imputation package, with the estimates from each dataset being pooled to provide an overall effect <sup>(36)</sup>. With time series data, however, it may be necessary to build statistical models for each imputed dataset separately before deciding on a final model to apply to all imputed datasets. This might be required if the variables that are statistically significant are particularly sensitive to the imputation, and to permit appropriate investigation of the existing autocorrelation patterns within an independent variable or outcome, per imputed dataset.

### Model variable selection

To determine the variables to be included in any final statistical model after multiple imputation, a decision would need to be made *a priori*, or in an exploratory analysis, based on the independent variables that are statistically significant in the models derived from each imputed dataset. It is also critical that autocorrelation is appropriately accounted for through the inclusion of lagged variables associated with both the independent and dependent variables. A lagged variable represents the value that specific variable had at a previous time point (e.g. previous day, two-hours ago, or seven days ago), and by being included in the model, allows the modelling of any existing autocorrelation within the variable's consecutive data points. Construction of lagged variables has been described in detail elsewhere <sup>(37)</sup>.

*Strengths and limitations of N-of-1 studies*

N-of-1 studies are relatively new to the field of human nutrition but have previously been applied in the fields of medicine, health psychology, education research, and experimental economics for some time <sup>(22,23,30,37)</sup>. The value of N-of-1 trials is in the opportunity they offer to include behavioural factors, which are important when considering the acceptability of, or compliance to, dietary interventions, especially in the context of health policies. For example, an N-of-1 design was used to assess the between-person variability in the psychological and social factors associated with daily alcohol consumption, in 25 adults with a history of alcohol dependence, using EMA <sup>(38)</sup>. This study, by fitting multi-level models to the longitudinal data, was able to identify individual factors that contributed to alcohol consumption changes before and after minimum unit pricing implementation in Scotland, including psychosocial factors such as temptation. This was despite the fact that each participant provided only 27 responses on average (out of a possible 84 questionnaires that were sent daily for 12 weeks), that the average time in study was only 64 out of 84 days, that the total response rate was only 48%, and only 15 out of 25 participants provided sufficient data for analysis <sup>(38)</sup>. Non-adherence to data collection, possibly including lower response rates near the end of longitudinal data collection, resulting in missing data, may be due to the ‘burden’ of daily data collection being too high for some participants. This has been identified as one of the challenges of N-of-1 studies, although overall, adherence to EMA has been considered high <sup>(37)</sup>. On the other hand, one main advantage of N-of-1 studies is that they offer the opportunity to tailor the intervention and data collection to an individual, increasingly involving wearables, smartphone sensors, and app or web-based data collection, which is believed to increase engagement and adherence to intervention <sup>(37)</sup>.

**Conclusion**

The adoption of new study designs and modelling approaches, including N-of-1 designs and statistical approaches to examine an individual’s response to interventions, will help to make predictions, and better understand associations, with regard to dietary interventions and health outcomes for individuals, or individuals that share certain characteristics, more effectively and accurately in the future. An advantage of these study designs is that they can include time-varying factors, such as the environment, epigenetics, nutritional status, food

behaviours, physical activity, gut microbiota, and the metabolome, that cause variation in individual responses to dietary interventions, and use this to assess, or predict, how individuals respond to dietary interventions. N-of-1 studies are especially suitable for clinically relevant outcomes that vary over time, due to the precision obtained from repeated measurements. However, N-of-1 studies need rigorous consideration of specific elements, including randomisation and blinding, the number of measurements and intervention cycles, and choosing appropriate outcomes that can be measured regularly, or continuously using wearables and other remote data collection methods.

**Funding:** The research of BdR is funded through the Scottish Government Rural and Environment Science and Analytical Services Division (RESAS), Grant RI-B5-06. The PhD project of EP is funded by Lipton Teas and Infusions.

## References

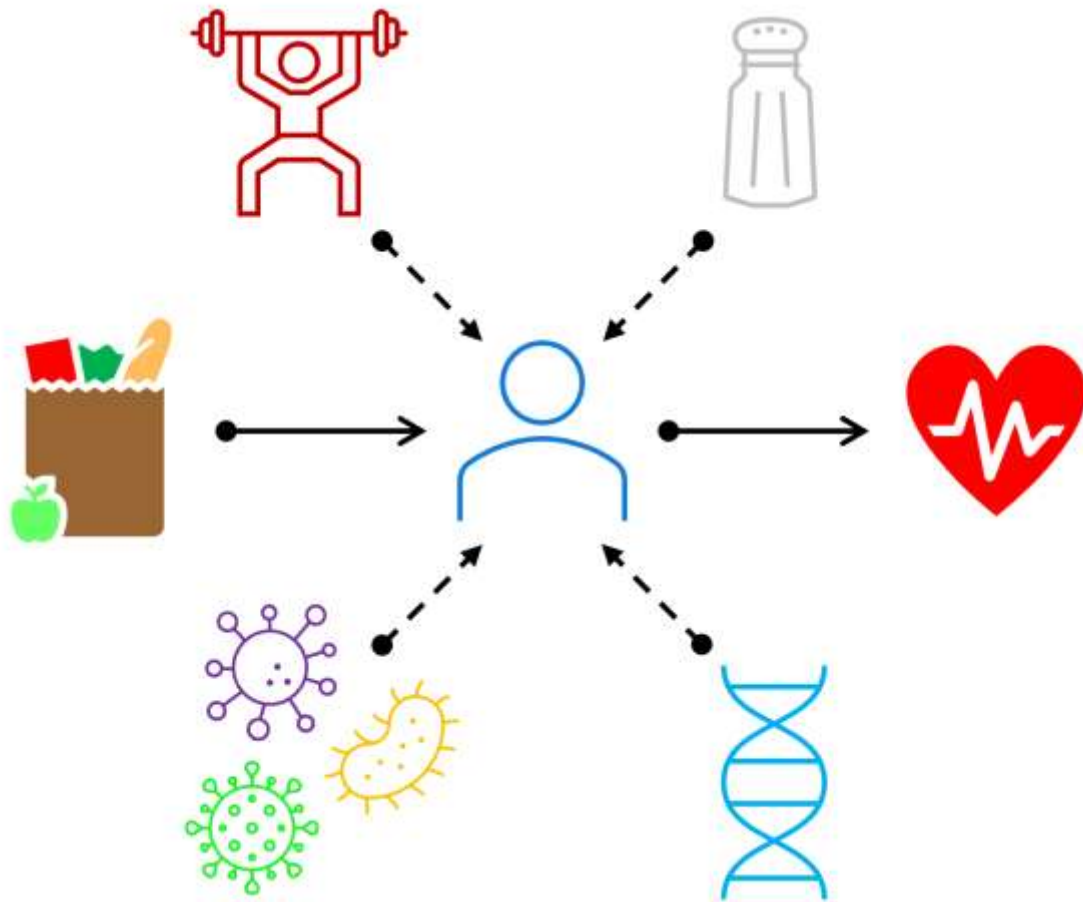
1. Krauss A (2018) Why all randomised controlled trials produce biased results. *Ann Med* 50, 312-322.
2. Hoare Z (2010) Randomisation: What, Why and How? *Significance* 7, 136–138.
3. Brennan L & de Roos B (2021) Nutrigenomics: lessons learned and future perspectives. *Am J Clin Nutr* 113, 503-516.
4. Berciano S, Figueiredo J, Brisbois TD *et al.* (2022) Precision nutrition: Maintaining scientific integrity while realizing market potential. *Front Nutr* 9, 979665.
5. Heinze G, Wallisch C & Dunkler D (2018) Variable selection - a review and recommendations for the practising statistician. *Biometrical Journal* 60, 431-449.
6. Potter TIT, Horgan GW, Wanders AJ *et al* (2022). Models predict change in plasma triglyceride concentrations and long-chain n-3 polyunsaturated fatty acid proportions in healthy participants after fish oil intervention. *Front Nutr* 9, 989716.
7. Caslake MJ, Miles EA, Kofler BM *et al* (2008) Effect of sex and genotype on cardiovascular biomarker response to fish oils: the FINGEN Study. *Am J Clin Nutr* 88, 618-29.

8. Madden J, Williams CM, Calder PC *et al* (2011) The impact of common gene variants on the response of biomarkers of cardiovascular disease (CVD) risk to increased fish oil fatty acids intakes. *Annu Rev Nutr* 31, 203-34.
9. Maiorino MI, Signoriello S, Maio A *et al* (2020) Effects of Continuous Glucose Monitoring on Metrics of Glycemic Control in Diabetes: A Systematic Review With Meta-analysis of Randomized Controlled Trials. *Diabetes Care* 43, 1146-1156.
10. Zeevi D, Korem T, Zmora N *et al* (2015) Personalized Nutrition by Prediction of Glycemic Responses. *Cell* 163, 1079-1094.
11. Popp CJ, Hu L, Kharmats AY *et al* (2022) Effect of a Personalized Diet to Reduce Postprandial Glycemic Response vs a Low-fat Diet on Weight Loss in Adults With Abnormal Glucose Metabolism and Obesity: A Randomized Clinical Trial. *JAMA Netw Open* 5, e2233760.
12. Berry SE, Valdes AM, Drew DA *et al* (2020) Human postprandial responses to food and potential for precision nutrition. *Nat Med* 26, 964-973.
13. Bermingham KM, Linenberg I, Polidori L *et al* (2024) Effects of a personalized nutrition program on cardiometabolic health: a randomized controlled trial. *Nat Med* 30, 1888-1897.
14. Blaak EE, Antoine JM, Benton D *et al* (2012) Impact of postprandial glycaemia on health and prevention of disease. *Obesity Reviews* 13,923-984.
15. Jarvis PRE, Cardin JL, Nisevich-Bede PM *et al* (2023) Continuous glucose monitoring in a healthy population: understanding the post-prandial glycemic response in individuals without diabetes mellitus. *Metabolism* 146, 155640.
16. Howard R, Guo J & Hall KD (2020) Imprecision nutrition? Different simultaneous continuous glucose monitors provide discordant meal rankings for incremental postprandial glucose in subjects without diabetes. *Am J Clin Nutr* 112, 1114-1119.
17. Brennan L & de Roos B (2023). Role of metabolomics in the delivery of precision nutrition. *Redox Biol* 65, 102808.
18. de Roos B (2024) How good are we at predicting the individual response to personalized diets? *Am J Clin Nutr* 120, 3-4.
19. Grohmann T, Walker AW, Russell WR *et al* (2023) A grape seed and bilberry extract reduces blood pressure in individuals at risk of developing type 2 diabetes: the PRECISE study, a double-blind placebo-controlled cross-over intervention study. *Front Nutr* 10, 1139880.

20. Nazarzadeh M, Bidel Z, Canoy D *et al* (2021) Blood pressure lowering and risk of new-onset type 2 diabetes: an individual participant data meta-analysis. *The Lancet* 398, 1803-1810.
21. Vieira R, McDonald S, Araújo-Soares V *et al* (2017) Dynamic modelling of n-of-1 data: powerful and flexible data analytics applied to individualised studies. *Health Psychology Review* 11, 222–234.
22. Hawksworth O, Chatters R, Steven Julious *et al* (2024) A methodological review of randomised n-of-1 trials. *Trials* 25, 263.
23. Gabler NB, Duan N, Vohra S *et al.* (2011) N-of-1 Trials in the Medical Literature: A Systematic Review. *Med Care* 49, 761-768.
24. Zucker DR, Ruthazer R & Schmid CH (2010) Individual (N-of-1) trials can be combined to give population comparative treatment effect estimates: methodologic considerations. *J Clin Epidemiology* 63, 1312–1323.
25. Senn S (2024) The analysis of continuous data from n-of-1 trials using paired cycles: a simple tutorial. *Trials* 25, 128.
26. Senn S (2019) Sample size considerations for n-of-1 trials. *Statistical Methods in Medical Research* 28, 372-383.
27. Kravitz RL, Duan N, eds, and the DEcIDE Methods Center N-of-1 Guidance Panel (Duan N, Eslick I, Gabler NB, *et al*). Design and Implementation of N-of-1 Trials: A User's Guide. AHRQ Publication No. 13(14)-EHC122-EF. Rockville, MD: Agency for Healthcare Research and Quality; February 2014.
28. Liao Z, Qian M, Kronish IM *et al* (2023) Analysis of N-of-1 trials using Bayesian distributed lag model with autocorrelated errors. *Stat Med* 42, 2044-2060.
29. Potter T, Vieira R, & de Roos B (2021) Perspective: Application of N-of-1 Methods in Personalized Nutrition Research. *Adv Nutr* 12, 579-589.
30. Lillie EO, Patay B, Diamant J *et al* (2011) The n-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Personalized medicine* 8, 161–173.
31. Nurmi J, Knittle K, Naughton F *et al.* (2023) Biofeedback and Digitalized Motivational Interviewing to Increase Daily Physical Activity: Series of Factorial N-of-1 Randomized Controlled Trials Piloting the Precious App. *JMIR Form Res* 7, e34232.
32. Shiffman S, Stone AA & Hufford MR (2008) Ecological momentary assessment. *Annu Rev Clin Psychol* 4, 1-32.

33. Moritz S & Bartz-Beielstein T (2017) imputeTS: Time Series Missing Value Imputation in R. *The R Journal* 9, 207–218.
34. Honaker J, King G, & Blackwell M (2011) Amelia II: A Program for Missing Data. *Journal of Statistical Software* 45, 1–47.
35. van Buuren S & Groothuis-Oudshoorn K (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of Statistical Software*, 45, 1–67.
36. McDonald S, Vieira R, & Johnston DW (2020) Analysing N-of-1 observational data in health psychology and behavioural medicine: a 10-step SPSS tutorial for beginners. *Health Psychology and Behavioral Medicine* 8, 32–54.
37. Kwasnicka D, Inauen J, Nieuwenboom W *et al* (2019). Challenges and solutions for N-of-1 design studies in health psychology. *Health Psychol Rev* 13,163-178.
38. Kwasnicka D, Boroujerdi M, O’Gorman A *et al* (2021) An N-of-1 study of daily alcohol consumption following minimum unit pricing implementation in Scotland. *Addiction* 116:1725–1733.





**Figure 1:** Endogenous and exogenous factors that can confound the association between a dietary intervention and an outcome, including environment, (epi)genetics, nutritional status, food behaviours, gut microbiota and the metabolome. These factors can also be exploited in precision nutrition studies to identify predictors of a response.