1. Study Design Considerations

Simon R Carding, Falk Hildebrand & George M Savva

To tackle health monitoring and its modulation via the gut microbiota effectively, and obtain accurate and meaningful data, a meticulous study design is required. Here we provide advice on key aspects of the design of research studies relating to the microbiome. Additional information can be found in several recent excellent review articles [1–6]. While here we set out some of the specific challenges related to microbiome research, you must study the general literature on study design that is specific for the type of study that you are conducting, whether an epidemiological (population based) study, clinical trial involving human participants, or other *in vivo* or *in vitro* work. Many books cover research design issues for laboratory biologists and those working with animals [7–9], for those doing human clinical trials [10] and for epidemiologists [11, 12].

1.1 Why do a microbiome study?

The microbiome has the potential to be of use as a:

- Diagnostic adjunct to traditional clinical and laboratory measures
- Determinant of responses to treatment
- Window into the side-effects of exposure to antibiotics
- Baseline measurement prior to the initiation of therapy
- Signature of immune processes such as inflammation
- Guide to nutrition and preventive interventions
- Monitoring to measure severity, progression or recovery from disease
- Identify small-molecule drugs and mechanisms of pathogenesis for targeted therapeutic interventions
- Investigating the adaptation of bacteria to their human host
- Identify pathobionts and their responses to all sorts of interventions

- Track microbial / viral strains across different sites, individuals and ecosystems

Each of these potential uses raises research questions that need different study designs to answer. In every case before designing a study it is essential that you consider clearly what your research questions are, and how the research findings are likely to be used, either for further research or application. This will inform the nature of the outputs you need from your study and hence its design.

Be clear about what the goal of your study is; that is, what question(s) must your study be able to answer? It is helpful if this question can be operationalised as one or more specific testable hypotheses or quantities to be estimated, and that it is as specific as possible. Even if the goals are mainly exploratory or hypothesis generating, it should be possible to state the question to be answered by the study in such a way that it will be clear when the answer has been attained and hence the study has been successful. Without these goals in mind and agreed within a research team it is near impossible to design a study.

It is important to note that the gut microbiome is highly individualized; even within the same individual, the gut microbiome can be highly dynamic, even over the short term. As a consequence, a good definition of the 'healthy microbiome' is still missing, and the same is true for a 'core microbiome' of microbial species present in the majority of human hosts. This hinders translational and interventional research since it is unclear what represents a 'good' outcome, or what might be considered a 'clinically significant' change.

The aim of research study design is to identify and remove (as far as possible) potential sources of bias and potential sources of variation from the experimental process, thereby improving accuracy and precision of research findings. Research must also be *ethical*, *acceptable* to researchers and participants, produce *useful* outputs, and be conducted within resource constraints. Ultimately research design should enable a statistical analysis that answers the research question of interest with optimal precision, so it is important that the research is planned up to and including this analysis.

What 'determines' the experimental design is clarity in the controls and replication required to enable rigorous statistical analysis of repeated measurements – but is affected by all the factors discussed below which must be considered and addressed in the design so that overall data collection is still sufficient for rigorous analysis and interpretation.

1.2. Basic study types and the evidence they provide

Research studies can be broadly classified into two groups, (a) those in which the investigators manipulate experimental conditions to observe the outcome (called experimental or interventional studies), and (b) those in which conclusions are drawn through observation (non-interventional or observational studies).

Randomised controlled trials are the most commonly applied form of interventional study. Generally speaking, randomised controlled trials provide the most powerful evidence of causal relationships between interventions and outcomes. However, they are difficult and expensive to conduct and rely on having clearly defined interventions and expected outcomes. They can be used to answer mechanistic questions and to test the efficacy or effectiveness of interventions. Sometimes researchers conduct 'uncontrolled' pre/post experiments, whereby inferences are drawn by comparing outcomes before vs after an intervention. This design is rarely appropriate - for two main reasons. First, pre/post designs often give false positives because of a statistical phenomenon called 'regression to the mean', whereby units selected into a study will naturally revert toward their population mean just through natural variation. Second, in a pre/post design the process of conducting the study becomes confounded with the intervention under investigation. Therefore, inferences made by comparing units before vs after an intervention are not valid, unless this is explicitly contrasted with units before vs after a simultaneous comparator condition, usually a control.

Observational studies, (sometimes called natural experiments) are more common and can answer a much wider range of research questions. They are also simpler and usually cheaper to conduct than trials, although they can be more difficult to analyse and draw conclusions from because of the potential for biases due to selection and confounding. Observational studies are usually classified based on how the participants are sampled as follows:

- Cross-sectional studies rely on recruiting a representative cohort of participants and measuring them once. These can be used for describing the distribution of microbiome features in a population and for estimating associations between microbiomes and other factors, although it is difficult to identify causal relationships using this design without making extra assumption or using additional sources of information that would allow you to rule out the possibility of reverse causation.
- Case-control studies. Here, instead of a single representative cohort, associations are estimated by comparing 'cases' (selected from individuals with a specific outcome) to 'controls', a group used to represent the rest of the population without that outcome. The

main advantage over a cross-sectional study is that outcomes that would be rare in a representative population cross-section can be selected for and so more easily studied. For microbiome research, an important limitation of both cross-sectional and cohort studies is that there is no way to 'recall' a previous microbiome from a participant, and so the conventional epidemiological use of case-control studies in estimating past factors contributing to current disease incidence is difficult if not impossible. Despite this, case-control studies are common in microbiome research, though possible causal inferences from them are severely limited.

- Longitudinal (prospective) studies. Here the issue of recall is avoided by following participants prospectively, usually to estimate a rate of change or to monitor the influence of factors measured at baseline on subsequent outcomes. However, the disadvantage of this design is the need to retain participants for the often long time period needed for outcomes to emerge (for example years or decades for many diseases), and the large sample sizes needed if the conditions under investigation are rare. If the anticipated changes occur more quickly, e.g., disease progression or treatment response, then this design is more feasible. Biomarkers can also be used as proxies for disease status.

1.3 Elements of experimental design

Once a research question and a basic suitable design are identified, the common elements that should be considered when developing any research design are:

- The target and sample populations (who should the results apply to, and which groups should be sampled from; defined using inclusion/exclusion criteria)
- How experimental units should be sampled and recruited from the sample population.
- The selection and recruitment of controls or comparator populations
- The optimal number of replications at different levels (see Section 2: Sample Size for a Microbiome Study)
- The time points at which samples are taken.
- How outcomes should be measured
- How information will be stored
- How samples will be processed and how research data will be analysed

Additionally for interventional studies:

- How treatments are designed (e.g., doses)
- How treatments are allocated to experimental units

Additionally for observational studies:

- How selection bias might affect estimates or generalisability
- How to identify and measure meta-data or covariates that will be needed for analysis and interpretation.
- Whether and how causal relationships can be identified

1.3.1 Use all available tools to improve reproducibility and integrity

Steps should be taken throughout the process of design, conduct, analysis and reporting of a microbiome study to improve its reproducibility and robustness. Follow a checklist such as that provided by Schloss [13] to evaluate and improve your practice ahead of the design and conduct of any microbiome study. If you are conducting an animal study, use the available study design and reporting resources available from the NC3RS [14].

1.3.2 Choice of controls and accounting for covariate-microbiome associations in observational studies

Case-control studies that aim to estimate the association between microbiomes and disease outcomes rely heavily on the selection of appropriate controls. Controls are groups of participants that do not experience the outcome under investigation while having an exposure (in this case microbiome) that is representative of the general population from which the cases were sampled. By comparing the controls with the cases, the features of the microbiome that are more over-represented in the cases compared with the wider population, can be identified. However, to attribute the disease status to differences in the microbiome between the cases and the general population, other possible factors that might explain this difference in risk must also be accounted for. These are potential *confounding* factors. This can be done either by measuring and adjusting for these factors in analysis, or by matching the control cohort to the cases so that the covariates are approximately balanced. Matching can be done at the group level, or at the individual level by selecting one or more controls to match each individual case.

Note a statistical adjustment for the confounding factors must always be included in the analysis of a matched case control study, even though they are matched, either by controlling for the confounders using multivariable methods or by using a ‘matched’ analysis such as a mixed model, conditional regression, paired test or repeated measured analysis.

With respect to matching criteria, many parameters have been reported to influence gut microbiome composition [1, 15], but the number is still expanding and the relative importance of these and other unknown factors remains unclear. In deciding on controls or on covariates to measure it is therefore important to consider and account for as many of these factors as possible including those listed here:

- Sex
- Age
- Ethnicity
- Geographical region
- Medication (especially antibiotics)
- Diet (e.g., fibre intake, bread preference, fruit consumption)/ Dietary supplements (e.g., probiotics, synbiotics)
- Alcohol use
- Adiposity (e.g., via body mass index or hip circumference)
- Blood chemistry (e.g., oxygen uptake capacity)
- Lifestyle factors (housing, pets)

It is also important to ensure that cases and controls are treated in the same way from a methodological perspective; for example, by ensuring the following are held constant for all study participants, or at least are not confounded with key exposure or outcome variables:

- Season/ diurnal cycle at sampling
- Methods of sample collection
- Sources and modes of collection of meta-data

In small scale observational or analytical studies, a good control group is one that can account for as many of the environmental and host (e.g., genetic) covariates as possible, such as healthy individuals related to and living in the same household as the patient and who are exposed to the same living conditions, food / diets etc. This may, however, raise logistical issues for large population-based studies that include hundreds of participants from diverse demographics and geographical areas. Larger, representative population cohorts can be effective at identifying relevant microbiome signals beyond multiple expected confounding factors and producing clearer discriminatory microbiome signals associated with health versus disease states.

When the microbiome is used for monitoring the severity or progression of a disease, or the efficacy of a treatment process, patient subsets with either less severe disease or lower responses to treatment, could be used as controls.

For an intervention, having control samples from before, (during), and after the trial alongside the 'no intervention' controls, is a good option.

Also take into account the nature of the host and specifics of their lifestyle. For example, mice are coprophagic and therefore exchange their microbiome with each other when held in the same cage (so called cage effect [16]), therefore case and control groups should be intermixed and distributed among several cages prior to the start of a study.

1.3.3 Measurement of covariates

Questionnaires can help capture information about participant demographics, lifestyle, hygiene and other behaviours known to contribute to individual variations in gut microbe populations. *The Microbiome Health Questionnaires* in use at the Quadram Institute have been used in longitudinal sampling studies and have helped demonstrate that the composition of an individual's gut microbiome is highly personalised and reflects their lifestyle and behaviour [17]. In new studies, collecting data using these questionnaires (or adaptations of them) can help interpret the associated analysis of gut microbe populations, and also identify covariates that help explain differences within and between individuals over time.

1.4. Defining the sample (inclusion and exclusion criteria)

Inclusion and exclusion criteria define the population and hence the individuals (sample population) that are eligible to participate in the study. Inclusion criteria should describe the population that you are interested in studying. For a randomised controlled trial this might be anybody who might be eligible for a treatment. For an observational study, inclusion criteria broadly define the sample population.

Exclusion criteria then specify which, if any, subgroups in that target population that you do not wish to include in your study. Participants might be excluded for practical reasons (e.g., inability to consent or difficulty collecting samples), because they are at high risk of adverse effects from interventions, or because they have some other feature that would strongly influence outcome measures or influence the association that the study is designed to test, thereby causing a bias or a loss of precision. However, such features might instead be dealt

with using control matching or by statistical analysis and, having too many exclusion criteria, could limit the generalisability of a study and potentially prevent adequate recruitment, particularly amongst groups such as older people who are likely to be using lots of medications and have a lot of comorbidities. Setting exclusion criteria too strictly also limits the applicability of your research to some demographics of the population and so this should be avoided. All exclusion criteria should be explicitly justified.

In addition, in observational research, selecting only specific subgroups of populations can lead to selection biases that affect estimates of associations. So, exclusion criteria should be limited to those features that would prevent successful participation in the study, would place the participant at excess risk, or might affect the measurement of outcomes or exposures in a way that cannot be handled in analysis.

Considering that >90% of all medications are taken orally and are potential substrates for the microbiome (xenometabolism), individuals who have undergone medication that can introduce huge variation in the microbiome, e.g., antimicrobials, in the preceding 3 months should be excluded from a microbiome study. Exclusion criteria should be considered on a study-by-study basis, but typical exclusion criteria to consider might be:

- Taken more than a daily dose of probiotics
- A long-standing gastrointestinal (GI) or liver function abnormality requiring on-going medical management or medication
- A history of cancer within the last 5 years (exceptions are squamous or basal cell carcinomas of the skin that have been medically managed by local excision)
- Unstable dietary history as defined by major changes in diet during the previous month, i.e. where a major food group in the diet has either been stopped or significantly increased, for example becoming vegetarian, vegan or no longer eating red meat
- A history of alcohol, drug or substance abuse
- A history of Hepatitis B or Hepatitis C
- Had major surgery of the GI tract in the past five years, with the exception of gall bladder removal
- Had any major bowel resection at any time.
- A history of IBS, ulcerative colitis, Crohn's disease or diverticulitis
- Persistent, infectious gastroenteritis; colitis or gastritis; persistent or chronic diarrhoea of unknown cause; *Clostridium difficile* infection (recurrent); or *Helicobacter pylori* infection (untreated).
- Constipation

- Regularly uses laxatives

Depending on the nature of the study and the site(s) of sampling, other exclusion criteria may be appropriate. For example, in studies investigating the relationship between gut microbes and the brain (the gut-(microbiota)-brain axis), participants with a pre-existing or current condition affecting their brain and mental health (dementias, bipolar disorder, schizophrenia, clinical depression, brain injury, stroke and epilepsy) might be excluded.

1.4.1 Sampling and recruitment

Once you have defined inclusion and exclusion criteria, it is important to have a clear plan for recruiting participants.

For inferences from an observational study to be valid, random sampling should be used. Typically this proceeds by identifying a sampling frame (all those in your study population who are accessible and who might be eligible) and then inviting all or a random subset of these to be screened for participation. The convenience of unsolicited volunteer samples should be avoided in observational studies, as inferences from these samples may be biased by factors related to their propensity to join the study, and it can be difficult to define what population such samples actually represent. In any case it is important that the process of selection and recruitment of individual participants is completely documented, including those who are not recruited, so that you and others can understand how generalisable your findings are to the target population and beyond.

Randomised controlled trials use randomised allocation of participants to treatments to ensure the validity of their inferences, and so it is less important that individual samples are representative of the sample population; convenience and unsolicited volunteer samples are acceptable in this case.

1.5. Sampling times and frequencies

An important aspect of study design is determining the timing and frequency of sample collection from the study population. The sampling intervals should be appropriate to the time period over which you might expect to see an effect, the natural rate of change of the microbiome and outcomes of interest. The number of intervals should also be determined by your research objectives.

To identify microbial features associated with disease incidence via a case-control study, participants should be recruited, and samples collected prior to incidence or as close as possible afterward. Controls should be selected that match cases at the time of incidence and should be sampled as close to the time of their corresponding case as possible.

To understand the impact of a treatment on the microbiome in an interventional study, then the primary comparison will be between patients on active treatment vs controls. Whether having pre-treatment data provides any additional information depends on whether the correlation between pre-treatment and post-treatment samples is high enough. In a parallel groups study collecting pre-treatment data will usually be helpful in explaining post-treatment differences. However in a cross-over study collecting samples ahead of each treatment period is unlikely to be useful, since each patient already has 'control' data. The only situation where this could be useful is when the study periods are a very long time apart compared with the treatment duration.

In practice the frequency of sample collection in longitudinal study designs is limited by logistical factors such as the cost of sample collection and storage, invasiveness of the sampling procedure, subject compliance and, in the case of retrospective studies, availability of samples from a pre-existing biorepository. In using samples retrieved from a biorepository it is important to note that they may not be uniformly separated in time, and suitable methods for analysis should be used.

It is important to be realistic about the precision that your experiment is likely to achieve and whether your main research questions are likely to be answered reliably enough. If practical considerations make the number of samples required technically difficult or impossible to achieve then the research question or design should be adapted to achieve the required rigour – if the study won't mean anything then it's not worth doing!

1.6. Sample types

If a specific aim is to enable cross study data comparisons then, first and foremost, materials and methods should be standardised wherever possible to ensure experimental consistency and eliminate methodological variation. Elsewhere in this collection of protocols advice is provided on the inclusion of controls and reference reagents to ensure standardisation and reproducibility from sample collection and processing to sequence and data analysis.

Sampling will be informed by the aims of the study as there is a broad range of samples that can be collected to study the microbiome: faecal, skin, oral, urine, vaginal and other body sites. Apart from its use in studies of GI conditions and liver disease, faeces (stool) is also an

amenable sample type for longitudinal studies. Compared with more invasive sample types (e.g. tissue obtained through colonoscopy), it can be collected easily at multiple time points from the same individuals with minimal invasiveness. On the other hand, faecal samples will be poorly representative of the upper GI tract, which requires more invasive and costly sampling and can vary a lot depending on food passage time; this can be partly approximated using the Bristol Stool Scale.

Positive and negative controls that account for variation in parameters that can influence the results of microbiome research, should be included in all microbiome studies. These include the type of DNA extraction kit ('kitomes'), sampling methods, likely contaminations, and sequencing methods [18]. This is particularly important in those studies using low-biomass samples (blood, amniotic fluid, cerebrospinal fluid, synovial fluid, placenta) [18]. Ultimately logistic and financial constraints will impose limits on the scale and nature of the sampling and control measures.

1.7 Reducing bias through randomisation and blinding

Randomisation should be used at all stages of an experiment where a systematic bias might occur. This includes the allocation of treatments to experimental units in interventional studies, but also to aspects of the study such as sequencing, batches and plate location for high throughput assays or the order in which samples are processed or measured. If these aspects are not randomised then it is possible for systematic differences to affect some subgroups of a study but not others, making the inferences of interest difficult or impossible. Randomisation ensures that all such factors are equally likely to affect each observation, reducing this risk.

Blinding refers to the masking of group allocations to researchers and participants, and like randomisation this prevents human factors from influencing measurements and results. In randomised controlled trials participants should be blinded to their allocations as far as is possible, in *all* research studies researchers should be blinded to group allocations or disease status when conducting analyses, taking measurements or handling samples.

1.8. Study and sample size, design effects

Optimal sample size is influenced by a number of factors that have been highlighted earlier. Studies designed to determine whether there is an association or a causative link between an intervention and changes in the microbiome, while accounting for as many covariates as possible (see 1.2 here and Section 2: Sample Size for a Microbiome Experiment), may require a large number of samples (>500 participants sampled) [10].

Where treatments or subject groups are compared, having sufficient replicates is critical for achieving the statistical power needed to detect differences. Related previous studies may suggest an appropriate study size although most microbiome studies lack sufficient power to identify microbiome signatures of health or disease. In many cases, it may not be possible to know *a priori* how many samples will be necessary, and the decision will often have to be based on the best estimate and the availability of resources. In these cases avoiding a 'hypothesis testing' approach to research is advisable, since the type 2 error is impossible to calculate [19]. Methods for sample size determinations are discussed in Section 2: Sample Size for a Microbiome Experiment.

Where subjects are not independent of each other, then the effective sample size of a study might be smaller than the actual number of experimental units used [20]. In human clinical trials this is most commonly caused by cluster-randomised designs, whereby for practical reasons groups are randomised to treatments instead of individual participants. In *in vivo* studies it is common for animals to be randomised to treatments at the cage level, in which case corrections need to be made to the sample size calculations and the statistical analysis to account for this. See Section 2: Sample Size for a Microbiome Experiment, for more details.

1.9. Ethics

Research ethics requires that the risk of harm associated with research participation must be balanced against the anticipated social benefits. Research on the human microbiome is fraught with numerous unanswered microbiological, clinical and social questions, so balancing possible risks and benefits is sometimes very difficult.

Particular caution must be taken in the context of clinical applications of microbial research. Because the risks associated with most human microbiome research and using samples provided by biobanks are often negligible, they are defined as a 'minimal risk' to participants' health and wellbeing in many regulations. However, collecting the microbiome from the gut by invasive sampling using endoscopy may pose some minimal additional risk. Institutional Review Board (IRB) and/or Health Research Authority (HRA) approvals are a common step in study planning as well. See guidance from the HRA if you are planning a study involving humans; this describes the ethical considerations there are likely to be and what approvals will be required for your specific study [20]. If your study is international or involves foreign funding, also make sure you comply with the relevant country's requirements.

For a human study, key regulatory considerations are:

- **Informed consent**

Obtaining informed consent is a well-established moral and legal obligation and requirement. Consent is considered fully informed when a competent patient or research subject to whom full disclosures have been made and who understands fully all that has been disclosed, voluntarily consents to treatment or participation on this basis. A subject's decision-making capacity, voluntariness, and vulnerability also demand serious consideration, especially in clinical trials research and innovative therapies.

- **Confidentiality and anonymity - Personal data storage and security**

Considerations about how personal data will be stored, and the infrastructure required to protect patient privacy are essential.

Due to the sensitive nature of microbial data (both clinical and genetic), any privacy breaches arising as a result of participation in microbial research could have serious negative consequences for participants, for your institute and for your research project. When microbial, genomic and medical information are linked to an individual they represent an unprecedented amount of personally-revealing information. Not only can this relate to personal disease susceptibility, but also travel experience, sexual practice, and consumption of alcohol and drugs.

Researchers should be aware of the legal requirements and institutional policies around handling such sensitive health data; all research studies should have an associated data management plan ahead of any data collection, and information governance specialists should be consulted in case of any uncertainty.

- **Data sharing**

Consider: who should be responsible for informing and explaining to participants (e.g., general practitioners [GPs]); to what extent the results and information should be disclosed; and what criteria and means should be adopted for returning research results.

When publishing a study, data should be made available using the FAIR (https://www.go-fair.org/fair-principles/) criteria. This also includes publishing sequencing data and metadata. However, metadata needs to be properly anonymized, ensuring that individuals cannot be identified, and human genome data can under no circumstances be shared in publicly available sequencing archives, such as ENA (https://www.ebi.ac.uk/ena/browser/home) or SRA (https://www.ncbi.nlm.nih.gov/sra). Therefore, ensure that human reads are removed from metagenomes before uploading these.

When preparing ethics applications keep in mind the possible tension between data confidentiality and the increasingly common requirement to publish research data. While it is good practice for anonymised research datasets to be published alongside research outputs, researchers should be mindful of the risk of patients becoming identifiable through having unusual combinations of covariates, and should ensure that the data released does not allow this to happen. Do not write statements into your ethics applications, consent forms or protocols that would prevent you from reasonably sharing research data.

1.10 Other considerations

These include control for experimental variation, particularly for multi-site or longitudinal studies, and sequencing technology (what type of technology is most appropriate to yield sufficient data to address the study objectives). See SOPs in this collection for specific guidance on some of these issues.

References

1. Falony G, Vandeputte D, Caenepeel C, et al (2019) The human microbiome in health and disease: hype or hope. Acta Clin Belg 74:53–64. https://doi.org/10.1080/17843286.2019.1583782

2. Kim D, Hofstaedter CE, Zhao C, et al (2017) Optimizing methods and dodging pitfalls in microbiome research. Microbiome 5:52. https://doi.org/10.1186/s40168-017-0267-5

3. Knight R, Vrbanac A, Taylor BC, et al (2018) Best practices for analysing microbiomes. Nat Rev Microbiol 16:410–422. https://doi.org/10.1038/s41579-018-0029-9

4. Qian X-B, Chen T, Xu Y-P, et al (2020) A guide to human microbiome research: study design, sample collection, and bioinformatics analysis. Chin Med J (Engl) 133:1844–1855. https://doi.org/10.1097/CM9.0000000000000871

5. Li D, Gao C, Zhang F, et al (2020) Seven facts and five initiatives for gut microbiome research. Protein Cell 11:391–400. https://doi.org/10.1007/s13238-020-00697-8

6. Johnson AJ, Zheng JJ, Kang JW, et al (2020) A Guide to Diet-Microbiome Study Design. Front Nutr 7:. https://doi.org/10.3389/fnut.2020.00079

7. Lazic SE (2016) Experimental Design for Laboratory Biologists: Maximising Information and Improving Reproducibility. In: High. Educ. Camb. Univ. Press.

https://www.cambridge.org/highereducation/books/experimental-design-for-laboratory-biologists/31C1A347D0ADB25226D7220A99C0EF56. Accessed 24 May 2021

8. Glass DJ (2014) Experimental Design for Biologists. Cold Spring Harbor Laboratory Press

9. Ruxton GD, Colegrave N (2016) Experimental Design for the Life Sciences, 4th edition. OUP Oxford, New York, NY

10. Senn S (2008) Statistical Issues in Drug Development, 2nd edition. Wiley

11. Celentano D, Szklo M (2019) Gordis Epidemiology, 6th ed. Elsevier Health

12. Rothman KJ, Greenland S, Lash TL (2008) Modern Epidemiology. Lippincott Williams & Wilkins

13. Schloss PD (2018) Identifying and Overcoming Threats to Reproducibility, Replicability, Robustness, and Generalizability in Microbiome Research. mBio 9:. https://doi.org/10.1128/mBio.00525-18

14. Percie du Sert N, Bamsey I, Bate ST, et al (2017) The Experimental Design Assistant. PLOS Biol 15:e2003779. https://doi.org/10.1371/journal.pbio.2003779

15. Vujkovic-Cvijin I, Sklar J, Jiang L, et al (2020) Host variables confound gut microbiota studies of human disease. Nature 587:448–454. https://doi.org/10.1038/s41586-020-2881-9

16. Hildebrand F, Nguyen TLA, Brinkman B, et al (2013) Inflammation-associated enterotypes, host genotype, cage and inter-individual effects drive gut microbiota variation in common laboratory mice. Genome Biol 14:R4. https://doi.org/10.1186/gb-2013-14-1-r4

17. Flores GE, Caporaso JG, Henley JB, et al (2014) Temporal variability is a personalized feature of the human microbiome. Genome Biol 15:. https://doi.org/10.1186/s13059-014-0531-y

18. Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt, MF, Turner P, Parkhill J, Loman NJ, Walker AW (2014) Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol 12: 87: http://www.biomedcentral.com/1741-7007/12/87

19. Amrhein V, Greenland S, McShane B (2019) Scientists rise up against statistical significance. Nature 567:305–307. https://doi.org/10.1038/d41586-019-00857-9

20. Basson AR, LaSalla A, Lam G, et al (2020) Artificial microbiome heterogeneity spurs six practical action themes and examples to increase study power-driven reproducibility. Sci Rep 10:5039. https://doi.org/10.1038/s41598-020-60900-y

21. Health Research Authority. https://www.hra.nhs.uk. Accessed 24 May 2021