

# 5. Guidance

## In this guide

### [In this guide](#)

1. [1. Acknowledgements](#)
2. [2. Lay Summary](#)
3. [3. Executive Summary](#)
4. [4. Introduction](#)
5. [5. Guidance](#)
6. [Checklist](#)
7. [6. Case Studies](#)
8. [7. References](#)

## 5.1 Producing, assessing, and procuring research

### 5.1.1 Identifying the research topic

The first step in the research process is to identify the broader research topic or problem area (for example, “Technological advances that help improve food safety”) and what the added value of a new study is. New research should not be undertaken without first having identified a clear evidence gap or research need. It is important to anticipate future policy issues as well as addressing current ones ([Government Social Research Profession, 2018](#)).

### 5.1.2 Scoping the literature

A scoping or rapid review should be undertaken to identify evidence gaps and research needs. Scoping reviews aim to identify knowledge gaps and map out the key types of available evidence ([Arksey & O’Malley, 2005](#)). Rapid reviews aim to synthesise knowledge in a given area but in contrast to systematic reviews, several practical steps are simplified or omitted to produce information in a timely

manner ([Tricco et al., 2015](#)). For example, rapid reviews tend to limit the literature search to one or a few databases, have one person screen and/or extract the data with a second person to verify, and present the results as a narrative summary. See [https://www.betterevaluation.org/en/evaluation-options/rapid\\_evidence\\_assessment](https://www.betterevaluation.org/en/evaluation-options/rapid_evidence_assessment) for further details.

Details should be given in the form of succinct instructions and with sufficient detail to permit their replication by an independent group; well-known operations need not be described in detail. Equipment and materials (over and above basic, standard laboratory ware) should provide sufficient detail to allow replication of work and avoid confusion with similar but unsuitable or untried equipment and materials.

### **5.1.3 Designing the research question**

Research questions are a useful starting point for most forms of knowledge generation. A research question must be 'researchable' or 'testable' and should not be too broad or imprecise. A research question should generate knowledge that makes a significant contribution to the existing knowledgebase. Good example research questions include: "What are people's views on genetically modified foods?"; "Is poor sleep quality associated with sugary beverage consumption?"; and "Did adults in the United Kingdom consume fewer vegetables during the first acute phase of the COVID-19 pandemic (vs. before the pandemic)?". When specifying the research question, it is important to think and plan ahead - to promote the production of valuable knowledge, there needs to be a clear match between the research question, the research design, the research method, and the analytical approach (see 'Selecting a research design'). When time and resources are limited, it is particularly important to ensure that the research question can feasibly be tested with a suitable research design to avoid research waste.

#### **5.1.3.1. The role of social and health theory**

Social and health theory aims to explain how things work and how to bring about change. Where possible, theory should be used to design the research questions, select the research instruments and guide the interpretation of the results. For example, the Health Action Process Approach (HAPA) proposes that health behaviours (for example, physical activity, smoking) are influenced by (1) a pre-intentional motivation process (for example, outcome expectancies, self-efficacy) that leads to a behavioural intention and (2) a post-intentional volition process

(for example, action planning, coping planning) that facilitates the adoption and maintenance of health behaviours ([Schwarzer & Luszczynska, 2008](#)). HAPA can, for example, be used to design the research questions – for example, “Is there a positive association between HAPA-related variables (for example, outcome expectancies, self-efficacy, intention, action planning, and coping planning) and adherence to physical distancing measures in adults in Belgium during the COVID-19 pandemic?” ([Beeckman et al., 2020](#)).

#### **5.1.4 Selecting a research design**

A research design provides a framework for the collection and analysis of data. A choice of research design reflects decisions about the priority being given to different dimensions of the research process. These include the importance attached to expressing causal connections between variables, generalising to larger groups of individuals, understanding the meaning of a given phenomenon in a specific context, or having a temporal understanding of the relationships between phenomena ([Bryman, 2016](#)). To promote the production of valuable knowledge, there needs to be a clear match between the research question, the research design, the research method, and the analytical approach.

##### **5.1.4.1 Cross-sectional designs**

Cross-sectional designs typically aim to i) estimate the prevalence and/or characteristics of people, places or events, or ii) examine the relationship between one (or more) independent variable and one (or more) dependent variable. Cross-sectional designs are most useful for the first aim but may be biased due to the selection of the study population (see ‘Specifying the sampling method’). Relationships identified in cross-sectional designs must be interpreted with caution. It is difficult to establish what is cause and effect in cross-sectional designs because the independent and dependent variables are measured at the same time. Cross-sectional designs can still be useful when time and resources are limited. It should also be noted, however, that qualitative research often entails a form of cross-sectional design, such as when the researcher employs semi-structured interviewing with a number of people ([Bryman, 2016](#)).

##### **5.1.4.2 Longitudinal designs**

Longitudinal designs involve multiple measurements from the same individuals over time (e.g., weeks, months, years). Longitudinal designs are useful when the aim of the research is to gain a temporal understanding of the relationship

between phenomena. However, longitudinal designs can be costly and are often subject to drop out – for example, individuals may stop responding to additional surveys or interview requests after a period of time ([Bryman, 2016](#)). Qualitative longitudinal research involves repeated qualitative interviews over a period of time.

#### **5.1.4.3 Case study designs**

Case study designs involve detailed examination of a particular case(s) in context. The case or unit of interest may be an individual, a school, a hospital, a city or even a country. Case study designs may involve a single or multiple measurements from the same case over time, often combining qualitative and quantitative research methods (see ‘Triangulation’) ([Bryman, 2016](#)). Case study designs can be costly if taking multiple measurements from the same case over time and may not generalise to a larger population; however, this largely depends on the selection of the case(s) (see ‘Specifying the sampling method’).

#### **5.1.4.4 Experimental designs**

Experimental designs aim to examine the effect of an independent variable (for example, a novel food ingredient) on a dependent variable (for example, tolerability). Experimental designs are useful when the aim is to understand causal connections between variables. Experimental designs are less susceptible to confounding than observational designs as the independent variable/exposure is randomly allocated. A distinction can be made between laboratory experiments (for example, experiments conducted under highly controlled conditions using a standardised procedure) and field experiments (for example, experiments conducted under real-life conditions). Laboratory experiments allow for close control of the independent variable; field experiments may confer less control of the independent variable. However, field experiments often have higher ecological validity than laboratory experiments (for example, they better reflect real-world conditions). It may not always be ethical or practically feasible to use an experimental design. For example, it is not ethical to directly expose individuals to carcinogenic chemicals if people have not already self-selected to the exposure ([Bryman, 2016](#)).

#### **5.1.4.5 Comparative research**

Comparative research aims to examine if the characteristics of people, places, or events differ by one or more variables. Comparative research is therefore a type

of experimental or cross-sectional design. For example, comparative research can be used to examine if a given treatment leads to improved outcomes compared with a control condition ([Bryman, 2016](#)). While comparative experimental designs are less susceptible to confounding (see 'Experimental designs'), comparative cross-sectional or longitudinal designs (for example, pre- and post-designs) are most useful when time and resources are limited (see 'Cross-sectional designs').

### **5.1.5 Selecting a research method**

A research method is a way to systematise observation and describe the type of tools and techniques to be used during data collection. Research methods also guide the type of analytical techniques to use. Research methods can broadly be grouped into qualitative and quantitative research methods.

When selecting a research method, there needs to be a close match with the research question and research design. For example, if the research aims to gain a better understanding of people's views on a topic and to generalise to a larger group of individuals, it is most appropriate to select a survey method. However, if the research aims to gain an in-depth understanding of people's views on a topic but without aiming to generalise to a larger group of individuals, interviews or focus groups may be more appropriate.

#### **5.1.5.1 Literature reviews**

Different types of literature reviews can be conducted depending on the research question(s) and available resources, including scoping reviews, systematic reviews, and rapid reviews (see 'Scoping the literature' for an overview of scoping reviews and rapid reviews). Systematic reviews involve conducting a literature search through entering queries into bibliographic databases, identifying relevant studies through applying pre-specified eligibility criteria, extracting data from the included studies, appraising the study quality and synthesising and interpreting the evidence. Systematic reviews are useful when the research aims to identify evidence gaps and research needs or to establish whether a given treatment is effective (see 'Meta-analysis') ([Gough et al., 2017](#)).

#### **5.1.5.2 Qualitative research**

##### **a. Interviews**

Interviews involve collecting qualitative data from one participant at a time using a set of structured, semi-structured or unstructured questions relating to the

research question. Selecting interviews over focus groups may be appropriate when the research topic is of a sensitive nature or there are likely to be unequal power dynamics across the sample ([Gill et al., 2008](#); [Moriarty, 2011](#)).

### **b. Focus groups**

Focus groups involve collecting qualitative data from groups of participants at a time using a topic guide relating to the research question. Selecting focus groups over interviews may be appropriate when useful data could be generated through the process of participants hearing and responding to the views of others, as well as to the questions from the researcher. Focus groups may also be selected when there is limited time to devote to qualitative data collection ([Gill et al., 2008](#); [Moriarty, 2011](#)).

### **c. Observation**

Observation involves collecting qualitative data through observation of key events or situations. For example, data could be collected through the observation of meetings within an organisational context or shopper behaviour within a supermarket. Those collecting observational data may be a participant (for example, a member of the meeting) or a non-participant (for example, an external attendee whose only purpose is to collect observational data).

Observational data can be combined with self-report data (for example, that which has been collected through interviews or focus groups) as part of a triangulation process (see 'Triangulation'). Observation may be useful when self-report data is limited by social desirability bias ([Moriarty, 2011](#)).

### **d. Document analysis**

Document analysis is when existing qualitative data is the focus of the study. This might include, for example, policy documents, magazines, newspapers or websites. Data extracted from such documents can be combined with primary data (for example, that which has been collected through interviews or focus groups) as part of a triangulation process (see 'Triangulation'). Document analysis may be useful when the opportunity for primary data collection is limited by resources or expertise or there is a wealth of existing data that could provide helpful insights on the research topic ([Moriarty, 2011](#)).

## **5.1.5.3 Quantitative research**

### **a. Surveys**

A survey is a research method that is used for collecting data from a group of respondents to gain information about phenomena of interest. A survey typically includes a range of questions, which can be open- or closed-ended ([Government Statistical Service, 2020b](#)).

## **b. Tests**

A test is a research method that can be used to assess an individual's attitudes, skills or abilities. Test performance is typically quantified by assigning a test score to each individual. For example, cognitive tests involve different problem-solving tasks. Here, test performance may be a combination of the time taken to complete the task and the response accuracy.

### **5.1.5.4 Mixed methods research**

Mixed methods research involves the study of different phenomena (or the same phenomenon - referred to as 'triangulation') using multiple, complementary research methods. This may, for example, involve the use of a survey in addition to a series of focus groups. Mixed methods research is particularly useful when addressing research questions that neither quantitative nor qualitative methods could answer alone. In addition, as each research method has its own strengths and weaknesses, findings that are consistent across different research methods are less likely to be false positives ([Munafo & Davey Smith, 2018](#)).

### **5.1.6 Selecting a study population and setting**

To ask a testable research question, the researcher needs to specify the target population (for example, the entire group of individuals with the characteristics that are of interest to the researchers) and setting (for example, the physical, social and cultural site in which the target population is located or where the research is being conducted). Research studies typically involve the recruitment and testing of a study sample (for example, a subset of participants from the target population).

In quantitative research, based on the results obtained from the study sample, conclusions can be drawn about the target population with a given level of confidence or precision, following the process of statistical inference (see 'Specifying the sampling method').

#### **5.1.6.1 Eligibility criteria**

It is important to carefully plan and pre-specify eligibility criteria to ensure that only participants from the target population are recruited into the study. Eligibility criteria should be split into inclusion criteria (for example, key features of the target population) and exclusion criteria (for example, features of the study participants who meet the inclusion criteria that may interfere with the study results). Inclusion criteria are typically related to age, sex/gender, and language proficiency. Exclusion criteria may include language proficiency.

### **5.1.7 Specifying the sampling method**

Different methods can be used to obtain the study sample and have different strengths and weaknesses. For example, the ability to generalise to larger groups of individuals depends on the sampling method used but is usually more expensive and may not always be desirable depending on the purpose of the research ([Bryman, 2016](#)).

#### **5.1.7.1 Qualitative research**

##### **a. Purposive sampling**

Purposive sampling involves collecting data from participants who meet the inclusion criteria and have characteristics that are defined for the purpose of the study. These characteristics might include age or gender, for example. This approach can be helpful when researchers have access to a broad range of participants who meet the inclusion criteria ([Moriarty, 2011](#)).

##### **b. Convenience sampling**

Convenience sampling involves collecting data from participants who meet the inclusion criteria and are easiest to access for the researcher (regardless of key characteristics). Part of this process may include asking participants to recommend others who meet the inclusion criteria (referred to as 'snowball sampling'). This approach can be helpful when resources (for example, time) or access to participants are limited ([Moriarty, 2011](#)).

##### **c. Theoretical sampling**

Theoretical sampling involves moving back and forth between data collection and data analysis. The analysis of initial data determines the next phase of data collection in terms of what the researchers might like to collect more information on and who they might like to collect it from. It is often used when the purpose of



the research is to generate new theory about a particular topic ([Mays & Pope, 2000](#); [Moriarty, 2011](#)).

### **5.1.7.2 Quantitative research**

#### **a. Simple random sampling**

Simple random sampling is a type of probability sampling method where there is an equal chance of selecting each unit (for example, individuals, schools, geographic regions) from the target population. The following six steps can be followed to create a simple random sample: i) define the target population; ii) select the sample size; iii) list the population; iv) assign numbers to the units; v) identify random numbers corresponding to the sample size; and vi) select the random sample. Simple random sampling is useful when the research aims to generalise to larger groups of individuals ([Bryman, 2016](#)).

#### **b. Stratified sampling**

Stratified sampling involves dividing units (for example, individuals, schools, geographic regions) from the target population into subgroups (or 'strata') based on shared characteristics (for example, gender, ethnicity, etc). Once divided, each stratum is randomly sampled using, for example, simple random sampling or another probability sampling method. Stratified sampling is useful when the research aims to generalise to larger groups of individuals ([Bryman, 2016](#)).

#### **c. Quota sampling**

Quota sampling involves dividing the target population into subgroups (just as in stratified sampling). Interviewer judgment is then used to select the units from each stratum based on a pre-specified proportion. For example, an interviewer may aim to sample 200 females and 200 males between the age of 45 and 65, with the interviewer selecting who to sample (for example, targeting). This is a non-probability sampling method. Quota sampling is useful when there is limited time to conduct a survey, the research budget is limited, or it is not a priority for the research to generalise to larger groups of individuals ([Bryman, 2016](#)).

#### **d. Convenience sampling**

Convenience sampling is a type of non-probability sampling method that involves the sample being drawn from the part of the target population that is most easy to access for the researchers. Convenience sampling is useful when there is limited time to conduct a survey, the research budget is limited, or it is not a

priority for the research to generalise to larger groups of individuals ([Bryman, 2016](#)).

### **5.1.7.3 Maximising response rates**

The study response rate is the proportion of contacted individuals who take part in the study and respond to study-related measurements. As individuals who take part in studies or respond to measurements can differ from those who decline participation or stop responding to measurements, maximising the response rate is particularly important for research which aims to generalise to a larger group of individuals. Response rates can be improved by including self-addressed stamped envelopes, informing respondents of the importance of the study, ensuring anonymity, and/or providing a financial incentive for participation ([Bland et al., 2012](#)). Incentives should not be the default option and must be adequately justified. If deciding to use incentives, they should be in voucher form or small tokens. Other options may include lottery draws and charitable donations on behalf of the participant (Government Social Research: Guidance on the use of incentives, 2021).

### **5.1.8 Specifying the sample size**

The sample size refers to the number of participants or observations (for example, measurements) in a study.

#### **5.1.8.1 Qualitative research**

In qualitative research, the sample size specification depends on the sampling method and purpose of the research.

##### **a. Data saturation**

Data saturation is when no additional findings or themes are being identified from newly collected data in the process of analysis. This can be interpreted as an indication that no further data collection is required. This approach to estimating the sample size is often pre-specified prior to data collection and analysis and relies on the researcher proceeding to the data analysis stage to make a judgment as to when to stop data collection ([Mays & Pope, 2000](#)). Recent attempts to identify appropriate sample sizes for saturation for commonly used qualitative methods have shown that studies using empirical data reached saturation within a narrow range of interviews (9–17) or focus group discussions (4–8) ([Hennink & Kaiser, 2022](#)). However, data saturation should be considered

alongside the selection of a qualitative analysis approach ([Braun & Clarke, 2021](#)). For example, data saturation is consistent with most types of thematic analysis but not with reflexive thematic analysis (see 'Selecting the analytical approach').

Although data saturation is useful for determining the sample size requirement, the concept of 'information power' was proposed as an alternative method to data saturation for guiding sample size selection for qualitative studies. It is based on the idea that the more information the sample holds that is relevant for the specific study, the lower the number of participants is needed. It has been suggested that the size of a sample with sufficient information power depends on i) how broad or narrow the aim of the study is, ii) the specificity of the study sample (for example, how similar or different participants are), iii) the extent to which established theory has been drawn upon, iv) the strength of the dialogue between researcher(s) and participant(s) during data collection, and v) the analysis approach that is applied to the data ([Malterud et al., 2016](#)).

### **5.1.8.2 Quantitative research**

In quantitative research, the sample size influences two key statistical properties: i) the precision of the model estimates; and ii) the power of the statistical test (see 'Statistical power'). Sample size calculations are required for all quantitative studies but may not be required for preliminary pilot studies.

#### **a. Statistical power**

Statistical power is intrinsically linked to the sample size and is the probability that a statistical test will detect an effect of a certain magnitude if that effect exists. Power is usually set to 80% or greater, and often to 90% for more definitive evaluations. A research question, research design, and research method can be appropriately matched, but without the study being sufficiently powered to detect a given effect, which can seriously limit the quality of the evidence. A study with insufficient power is more likely to miss an effect that actually exists and produce misleading conclusions ([Bland et al., 2012](#)).

#### **b. Information required to calculate the sample size**

The following information must be known or estimated to calculate the sample size: i) the variables of interest in the study (including the type of data; see 'Designing or selecting the research instruments'), ii) the desired power of the statistical test to detect the expected effect (typically set to 80% or higher), iii) the desired significance level (typically set to 5%), iv) the effect size of clinical

importance (which may be expressed as the mean difference between two treatments, the relative risk of a diagnosis if a given exposure is present versus absent), v) the standard deviation of continuous outcome variables, vi) if the analysis will involve one- or two-sided tests, and vii) aspects of the research design (for example, whether the study includes repeated measurements, whether groups are of equal sizes). The sample size calculation should relate to the study's primary outcome variable ([Bland et al., 2012](#)).

### **5.1.9 Designing or selecting the research instruments**

The success of data collection requires careful planning of what instruments to use and how the data will subsequently be analysed. The type of instrument that is most appropriate to use depends largely on the research question, research design and research method

#### **5.1.9.1 Qualitative research**

##### **a. Principles of good topic guide design**

Qualitative research often uses topic guides to provide a consistent structure across participants when collecting data (for example, through interviews or focus groups) and to ensure that the data that is collected aligns with the research question(s). Topic guides may be structured, semi-structured or unstructured. When designing or selecting a topic guide for interviews or focus groups there are several recommended principles that should be adhered to. These include, but are not limited to, use of open and non-leading questions, considered use of prompts to elicit further information where needed and adoption of user-friendly language. Piloting of questions, with those that meet the inclusion criteria for participation, can be a particularly useful approach for identifying areas for improvement ahead of beginning data collection ([Gill et al., 2008](#); [Rosenthal, 2016](#)).

##### **b. Validity**

Validity in qualitative research refers to the extent to which data collection tools and processes as well as the data that they produce are appropriate for the research question. Ensuring validity at the data collection stage should include, but is not limited to, active consideration of the role of the researcher who is collecting the data and the impact that this could have on the findings. For example, this might include the past experiences and beliefs of the researcher as well as their personal characteristics ([Kuper et al., 2008](#); [Malterud, 2001](#); [Mays &](#)

[Pope, 2000](#)).

### **c. Reliability**

Reliability in qualitative research refers to the extent to which data collection and analytic processes are consistent. Ensuring reliability at the data collection stage should include, but is not limited to, use of a topic guide, the generation of detailed field notes to accompany the qualitative data and the use of recording and transcription to formally document the data collected through an interview or focus group ([Gill et al., 2008](#); [Mays & Pope, 2000](#); [Rosenthal, 2016](#)).

## **5.1.9.2 Quantitative research**

### **a. Principles of good survey and test design**

When designing or selecting survey items or tests, it is important to ensure that items and test instructions are clearly worded and concise. Avoid double negatives, leading questions, and ambiguous wording. Piloting of survey items and tests with a small number of the intended sample is important for identifying areas for improvement ahead of the data collection. Where possible, it is useful to make use of existing surveys or tests which have been validated across different populations and settings.

### **b. Validity**

A survey or test is said to be valid if it measures what it intends to measure. In social and psychological research, validity cannot be directly established (as most constructs of interest are not directly measurable but inferred through surveys or tests). There are, however, a few ways in which validity can be indirectly established. A survey or test is said to have good 'face validity' if respondents think that the items measure what they are intended to measure. A survey or test has good 'content validity' if experts (for example, healthcare professionals, policymakers) think that the survey or test contains items or tasks which cover all aspects of the construct being measured. 'Convergent validity' is the extent to which a survey or test is positively correlated with similar constructs. 'Divergent validity' is the extent to which a survey or test is negatively correlated with dissimilar constructs. 'Predictive validity' is the extent to which a survey or test is predictive of an expected outcome.

### **c. Reliability**

A survey or test is said to be reliable if it elicits sufficiently similar responses each time someone takes it. Changes to the question wording or task structure may lead to different responses and can influence the survey's reliability. It is good practice to examine a survey or test's test-retest reliability, which can be obtained through giving a group of people the same survey or test twice over a period of time. The survey or test scores from timepoint 1 and timepoint 2 are then correlated to evaluate the test's reliability, with a strong positive correlation indicative of high reliability.

#### **d. Types of variables**

When designing or selecting research instruments, it is important to specify the type of data gathered from these and their scale of measurement, as this will help determine what analytical approach is appropriate as well as guiding the sample size calculation. Quantitative data can be measured on the interval scale (for example, the data have a natural order and the interval between values has meaning, for example, weight, height, the number of children), ordinal scale (for example, the data have a natural order but the interval between values does not necessarily have meaning, for example, many psychological variables) or nominal scale (for example, categorical data where the categories do not have any natural order, for example, gender). In addition, data measured on an interval scale can be continuous (for example, variables can take all possible values in a given range, for example, weight, height) or discrete (for example, variables can take only a finite number of values in a given range, for example, the number of children) ([Bland et al., 2012](#)). Likert scales are widely used in social science research and commonly have four to seven response options. Likert scales can be treated as interval scales, but strictly speaking they should be treated as ordinal variables with arithmetic operations avoided ([Wu & Leung, 2017](#)).

#### **5.1.10 Selecting the analytical approach**

The type of analytical approach that is most appropriate to use depends on the research question, research design, research method, and research instrument(s). For example, there could be a good match between the research question and method, but a poor match between the method and analytical approach (for example, using a *t*-test when linear regression would be more appropriate). Often, there are identifiable ways of improving the match – for example, by adding a different type of analytical approach or research method to address the research question.

### **5.1.10.1 Qualitative research**

Selecting the appropriate qualitative analysis approach depends on the type of qualitative data collection and the research question(s) at hand.

Irrespective of the type of analytical approach selected, researchers should enhance the trustworthiness and credibility of their qualitative analysis by involving other researchers in the analysis process where appropriate (for example, through double coding or critical review of themes and sub-themes against illustrative data fragments), creating an audit trail of the analysis from start to end, identifying similarities and differences across participants, and not disregarding 'negative cases' (for example, those that represent a contradictory perspective). Once findings have been established, researchers should consider the value of checking their interpretations of the data with the participants ('member checking'), actively seek and consider alternative explanations for the findings and triangulate the findings with those that have come from other data sources. Trustworthiness and credibility should also be considered when reporting the findings (see 'Reporting the results') ([Kuper et al., 2008](#); [Malterud, 2001](#); [Mays & Pope, 2000](#)). The optimal methods for ensuring trustworthiness and credibility should be considered alongside the analysis approach that has been selected. For example, double coding and member checking will be less appropriate when the analysis emphasises an inductive approach in which the interpretation of the individual researcher is prioritised. Member checking will, however, be more appropriate where participants are more heavily involved in the research.

#### **a. Thematic analysis**

Thematic analysis involves grouping together qualitative data fragments to create codes, sub-themes and themes that represent the dataset. It may be useful when the primary aim of the analysis is to describe the qualitative dataset (see 'Selecting a research design') ([Braun & Clarke, 2006, 2019](#)). Different variations of this approach include 'coding reliability', 'codebook' and 'reflexive'. 'Coding reliability' and 'codebook' approaches to thematic analysis are more deductive in nature and may be used when the themes that will be explored have been predetermined to some extent. 'Reflexive' thematic analysis, on the other hand, is more inductive in nature and may be used when the researcher(s) would like to explore patterns of meaning in the dataset ([Braun & Clarke, 2019](#)).

#### **b. Grounded theory**

Grounded theory draws upon theoretical sampling to collect and analyse data in cycles. It involves the inductive coding of data fragments and the grouping of codes into higher level units of meaning. It may be useful when the purpose of the research is to inductively develop a theory about a particular topic of interest ([Moriarty, 2011](#)).

### **c. Framework analysis**

Framework analysis emphasises a deductive approach in which prior theory, models from previous research, detailed breakdown of research aims, or questions are applied to the qualitative dataset. It may be useful when the researcher would like to map the dataset against a pre-existing idea ([Parkinson et al., 2016](#)).

### **5.1.10.2 Quantitative research**

Selecting the appropriate statistical test largely depends on the number and nature of the dependent variables and the nature of the independent variables. Typically, data can be analysed in more than one way to produce a legitimate answer ([Bruin, 2011](#)).

#### **a. Chi-square test, correlation test, t-test and analysis of variance**

The Chi-square test, correlation test, t-test and ANOVA are useful for tests involving one or more independent variable. When the researcher wishes to adjust the analysis for potential confounding variables, however, linear or logistic regression is recommended ([Bruce et al., 2018](#)). See [Bruin \(2011\)](#) for a detailed table to help select an appropriate statistical test.

The Chi-square test can be used to compare two categorical variables from a single population. The Chi-square test tells us whether there is a significant association between the two categorical variables.

The correlation test can be used to test the association between two continuous variables from a single population.

The t-test can be used to compare the means from two groups (paired or unpaired) and if these meaningfully differ from one another. Analysis of variance (ANOVA) can be used to compare the means from multiple (for example, three or more) groups.

#### **b. Linear regression**



Linear regression can be used to describe how one continuous variable (the outcome or dependent variable) depends on another variable (the explanatory or independent variable). The simplest description of a relationship between two variables is a straight line, which is mathematically defined by an intercept (where the line crosses the y-axis) and a slope (the gradient). The intercept is the value of the dependent variable when the independent variable is zero, and the slope is the amount the dependent variable increases for each unit increase in the independent variable. The equation for the linear regression is therefore:  $y$  (dependent variable) =  $a$  (intercept) +  $b$  (slope)\* $x$  (independent variable). Multiple independent variables can be included in a linear regression analysis, which allows researchers to adjust for confounding variables ([Bruce et al., 2018](#)).

### **c. Logistic regression**

Logistic regression is similar to linear regression, but the main difference is that the former can be used to describe how one categorical variable (the dependent variable) depends on another variable (the independent variable). Having a categorical outcome presents a problem for regression analysis, which has one key assumption: a linear association between the dependent and independent variables. To carry out regression analysis with a categorical outcome, a continuous outcome is derived by using the natural logarithm of the odds of, for example, the presence versus absence of a given exposure. The log odds provide an appropriate mathematical transformation that gives a continuous linear function and therefore meets the characteristics required for the regression analysis. Multiple independent variables can be included in a logistic regression analysis, which allows researchers to adjust for confounding variables ([Bruce et al., 2018](#)).

### **d. Survival analysis**

Survival analysis can be used to estimate the time that people in the study survive until dying (mortality), falling ill (morbidity), or experiencing an event of interest (health- or non-health-related). The Kaplan-Meier method or Cox proportionate hazards regression (another form of regression analysis) can be used to estimate survival. If using the Kaplan-Meier method to estimate survival times, the log-rank test can be used to test the hypothesis of no difference in survival between two or more groups. In survival analysis, the probability of an event occurring is called the hazard. This probability can vary over time. Cox regression can be used to estimate the hazard ratio (for example, the estimate of risk at any fixed point in time) ([Bruce et al., 2018](#)).

## **e. Time series analysis**

Time series analysis is an analytical technique which can be used to assess trends in repeated measurements taken at regular intervals and their associations with other trends or events, taking into account the temporal structure of the data. Interrupted time series analysis is a type of time series analysis which can be used to assess if an event or the introduction of a new policy (or treatment) was associated with a change in the trend of an outcome variable. It should be noted that time series analysis can only assess associations at the temporal granularity of the series. For example, if measurements have been taken at weekly intervals, the time series analysis can assess week-by-week changes, and not changes over a shorter timeframe (for example, day-by-day changes). However, data can be aggregated to assess changes over a longer timeframe (for example, month-by-month changes). The data for any time series analysis is typically divided into three main components: i) a trend component, ii) a seasonal component and iii) a random component ([Beard et al., 2019](#)).

## **f. Meta-analysis**

As part of the systematic review process (see 'Selecting a research method'), a meta-analysis can be carried out to obtain a quantitative summary of, for example, a treatment effect or an exposure risk across comparable studies. This is commonly done by combining the results from individual studies or by analysing the raw data from the individual studies if they are available. When combining data from several studies, the sample size is increased and hence also the statistical power to obtain more precise estimates. There are four main steps in carrying out a meta-analysis: i) an assessment of publication bias using a funnel plot (or a statistical analogue) to look for asymmetry, ii) a statistical test for heterogeneity (difference) of the treatment effect between the selected studies, iii) calculating a pooled estimate (for example, risk ratio, odds ratio, standardised mean difference) and 95% confidence interval (CI) for the treatment effect after combining all the trials using a fixed- or random-effects model (depending on whether substantial heterogeneity has been detected), and iv) conducting a hypothesis test to examine if the treatment effect is statistically significant. The results of a meta-analysis (for example, the estimates and CI for each study and the pooled results) should be illustrated in a forest plot ([Bruce et al., 2018](#)).

### **5.1.11 Reporting the results**

The use of standard templates can assist the clear presentation of a study's results and help ensure that all the required information is presented ([FSA Science Council, 2021](#)). The [EQUATOR Network](#) has developed several guidelines and accompanying checklists for the reporting of results from studies using different research designs.

When selecting which EQUATOR Network checklist to use, researchers should ensure that there is a match between the research design used and the research design that the checklist was developed to address. For example, if the research design used is a systematic review, the PRISMA checklist (described in further detail below) should be used. If the research design is a randomised controlled trial, the CONSORT checklist should be used.

### **5.1.11.1 Systematic reviews**

#### **a. PRISMA checklist**

The Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) statement and checklist can help systematic reviewers transparently report why the review was done, what the researchers did, and what they found. For example, the PRISMA checklist asks reviewers to describe the information sources, search strategy, selection process, risk of bias assessment, synthesis method(s), certainty assessment, and the registration information for the review ([Moher et al., 2009](#)).

### **5.1.11.2 Qualitative methods**

#### **a. COREQ checklist**

The COnsolidated criteria for REporting Qualitative research (COREQ) is a checklist of 32 items that should be included in reports of qualitative research to ensure that readers have all necessary details to interpret the results. The items are grouped into three domains: research team and reflexivity, study design, and data analysis and reporting. For example, the COREQ checklist asks researchers to report on key characteristics of those conducting data collection, the setting in which data collection took place, and the number of researchers involved in coding the dataset ([Tong et al., 2007](#)).

#### **b. TIDieR checklist**

The Template for Intervention Description and Replication (TIDieR) statement and checklist aims to improve the completeness of reporting, and ultimately the replicability, of interventions. For example, the TIDieR checklist asks researchers to report the intervention materials, procedures, who provided the intervention, how the intervention was delivered (for example, the mode of delivery), the number of times the intervention was delivered, where the intervention was delivered, and the extent to which the intervention was delivered as planned ( [Hoffmann et al., 2014](#)).

### **5.1.11.3 Quantitative methods**

#### **a. Basic presentation and analysis of results**

Descriptive tables and figures (for example, histograms) are useful for conveying the results. Tables and figures should be concise and easy to read. Avoid shaded backgrounds, unnecessary borders, boxes around legends and other content, patterns, textures and shadows, 3D shapes, unnecessary data markers (relevant mainly for line charts), and thick or dark gridlines. However, statistical commentary is also needed to bring the numbers to life. Good statistical commentary draws attention to important findings and puts them in context ( [Government Statistical Service, 2020a](#)).

#### **b. STROBE checklist**

The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement and checklist provide guidelines for the reporting of observational studies (for example, cross-sectional studies, longitudinal studies). For example, the STROBE checklist asks researchers to describe potential sources of bias, the study size, the statistical method(s), limitations, interpretation of the study results, and generalisability of the study results ( [von Elm et al., 2007](#)).

#### **c. CONSORT checklist**

The Consolidated Standards of Reporting Trials (CONSORT) statement and checklist provide guidelines for the reporting of randomised controlled trials. For example, the CONSORT checklist asks researchers to describe the trial design, the interventions for each group, the sample size, randomisation, blinding, the statistical method(s), the participant flow, the numbers analysed, limitations, interpretation of the study results, and generalisability of the study results ( [Schulz et al., 2010](#)).

## 5.1.12 Interpreting the results

### 5.1.12.1 Statistical significance

Appropriate statistical analysis is essential in quantitative studies. Statistical significance alone is insufficient for data to be important or meaningful. Any data showing a statistically significant effect should be accompanied by an explanation as to why the statistical test used was appropriate as well as the magnitude or relevance of the effect. Statistical point estimates and confidence intervals should be presented alongside statistical significance, as this provides more information to the reader on the magnitude of any effect and its associated uncertainty ([FSA Science Council, 2021](#)).

### 5.1.12.2 External validity/generalisability

Generalisability is the extent to which the results from a quantitative study applies to the target population and not just the sample who took part in the research. Generalisability can also encompass the extent to which the results are applicable to different geographic regions or countries, or different time periods. Typically, researchers aim for population rather than geographic or temporal generalisability. The recruitment of a representative sample is a necessary but insufficient condition for generalisability. For results to be generalisable/externally valid, they also need to be internally valid. Specifically, measurement error (for example, driven by not using valid and reliable research instruments; see 'Designing or selecting the research instruments') and residual confounding are key threats to internal validity.

Traditionally, qualitative studies have not set out to produce generalisable results (at least not in terms of statistical-probabilistic generalisability) and instead have focused on conducting a more in-depth exploration of the study sample. However, more pragmatic approaches may set out to generalise and, in this case, researchers should be mindful of the characteristics of the study sample and the relationship of these to any 'negative cases' identified in the findings. There are four different types of generalisability that qualitative researchers may consider: naturalistic, transferable, theoretical and intersectional generalisations ([Smith, 2017](#)). Naturalistic generalisability refers to when the results resonate with the reader's lived experiences of a given phenomenon. Transferability refers to when the reader can intuitively transfer the research findings to their own actions/behaviours. Theoretical generalisation refers to when the concepts or theories constructed make sense and have significance also in other research,

even if the contexts or populations differ. Intersectional generalisability refers to work that “digs deep and respectfully with a community over time to record the particulars of historically oppressed and/or colonised peoples/communities” ([Smith, 2017](#)).

### **5.1.12.3 Causality**

When interpreting the results from quantitative studies, causal language should be avoided unless the study specifically aimed to examine the causal connection between variables, preferably through randomisation to an intervention and control group as part of an experimental study. Causal inference (for example, uncovering a causal effect of a treatment or exposure on a given outcome) is also possible using cross-sectional or longitudinal research designs, but care must be taken to eliminate confounder bias, measurement error, and differential selection into the study (see ‘Selecting the sampling method’).

### **5.1.12.4 Uncertainty**

Statistical models fitted to quantitative data are simplified idealisations of how the world works. Uncertainty about the event rate, model parameters, and/or alternative model structures is common. The sources of uncertainty include (but are not limited to) essential unpredictability or limitations in information or formalised knowledge. While some of these uncertainties, such as the parameters within models, can be quantified (for example, through presenting the precision of the model estimates), some are more appropriately expressed through a statement of confidence or an outline of the limitations of the model. Evidence derived from modelling should provide a comprehensive statement of uncertainty – including quantification where appropriate, and a statement of its limitations. Sensitivity analyses can provide an indication of the between-model uncertainties arising from both parameters within models and alternative model structures ([Spiegelhalter & Riesch, 2011](#)).

### **5.1.12.5 Strengths and limitations**

Every quantitative and qualitative study – irrespective of the research question, research design, research method, or analytical approach – has specific strengths and weaknesses. It is important to transparently report these to allow for appropriate interpretation of the study results. A high-quality research report should contain a balanced overview of the study’s strengths and limitations. These may include, but are not limited to, the obtained sample size, the obtained

sample characteristics, and the validity and reliability of the research instruments used.

### **5.1.13 Writing the study conclusion**

Study conclusions should consist of one or two sentences that summarise what the study found and should clearly be supported by data ([Government Social Research Profession, 2018](#)). It should not contain any abbreviations or musings on what future studies might try to determine (although statements that explain how the current research will affect future research are fine) ([Addiction Journal, 2022](#)). For FSA/Government reports, longer conclusions with areas for further research may be favoured.

## **5.2 Research management and dissemination**

### **5.2.1 Identifying ethical and data governance issues**

Research should enable participation of the groups it seeks to represent and should be conducted in a manner that minimises personal and social harm ([Government Social Research, 2021](#)).

#### **5.2.1.1 Informed consent**

Anybody participating in a research study should provide their informed consent. This means that they have been given all necessary information (for example, in the form of a participant information sheet) to decide on whether they would like to be part of it. Informed consent can be recorded on paper (for example, through signing a consent form), virtually (for example, through use of an electronic signature on a consent form) or verbally (for example, through stating that consent is given at the start of a recorded interview or focus group) ([Nijhawan et al., 2013](#)).

#### **5.2.1.2 Right to withdraw**

It should be clearly communicated to participants (as part of the informed consent process) up until what point they are able to withdraw their data and discontinue participation. For example, this might be up until the analysis stage of the research study has been initiated ([Nijhawan et al., 2013](#); [Schaefer & Wertheimer, 2010](#)).

### **5.2.1.3 Confidentiality**

It should be clearly communicated to participants (as part of the informed consent process) the extent to which their participation in the study will be confidential. Confidentiality is defined as the extent to which any personal information from participants is protected ([Nijhawan et al., 2013](#)).

### **5.2.1.3 Data protection**

Data should be collected, managed and stored in accordance with the Data Protection Act 2018 and the General Data Protection Regulation (GDPR) ([European Parliament and the Council of the European Union, 2016](#)).

### **5.2.1.4 Equity**

It is best practice for researchers to identify whether there are major, minor or no impacts on equality of opportunity with regards to people of, for example, different religious belief, political opinion or ethnic group. This may refer to equality of opportunity with regards to participating in the research, being involved in the development of research priorities or the impact of a treatment or policy (for example, differential impact of a treatment across people with high and low levels of education, respectively) ([HM Revenue & Customs, 2019](#)). See 'Stakeholder mapping and involvement' for guidance as to when and how in the research process to engage with vulnerable and underserved respondents to increase the likelihood of their research participation.

## **5.2.2 Producing a risk management plan**

Risk assessment involves using a scientific approach to identify and define hazards, and to estimate potential risk to human and/or animal health. This includes evaluating the likely exposure to risks from food and other sources. Risk management is the consideration of potential measures to either prevent or control the risk. It takes into account risk assessment and consumers' wider interests in food to formulate a response ([FSA Science Council, 2021](#)).

Risks to consider include, but are not limited to, data privacy and security, adverse events (including participant distress), health and safety (on the part of the researchers), and financial (for example, who will bear the costs) and reputational risks.



Different types of studies are subject to different FSA sign-off procedures, depending on the level of risk involved (for example, whether the research will have major policy implications, whether there are reputational risks) and the complexity of the research (AU sign-off policy Jan 2022). In addition, an assessment of challenges within the methodological approach (for example, any risks around the data sources, sample size, or implementation) or operational issues should be performed.

### **5.2.3 Openness and transparency**

For evidence to be trustworthy it should be honestly and openly communicated, so that it can be independently interrogated. Any factors that could have influenced the results or the conclusions should be disclosed. Two key criteria in ensuring this are transparency and impartiality ([FSA Science Council, 2021](#)).

The Open Science movement aims to ensure the transparency and reproducibility of scientific research. Open Science has several benefits for research and society at large, including the contribution to high-quality research. When articles, materials, data and analytic code are available and easily accessible, it is easier for other researchers to build on previous research.

#### **5.2.3.1 Pre-registration**

The pre-registration of research involves specifying the research plan in advance of conducting the study and submitting the plan to a registry. Pre-registration separates hypothesis-generating (exploratory) from hypothesis-testing (confirmatory) research. Both are important. But the same data cannot be used to generate and test a hypothesis, which can happen unintentionally and reduces the credibility of the research results. Addressing this problem through planning improves the quality and transparency of the research ([Nosek et al., 2018](#)). Depending on the time and resources available for a project, pre-registration may not always be feasible but should always be used for experimental designs (for example, trials of behavioural interventions).

#### **5.2.3.2 Open materials, data and code**

Openly publishing study and intervention materials, data and code help with error detection and scrutiny, encouraging accurate data collection and labelling practices. In addition, the availability of data allows other researchers to examine the reproducibility of research results. Long-term storage of open materials, data

and code is facilitated through freely available websites and international repositories, such as [www.food.gov](http://www.food.gov), the Open Science Framework (<https://www.osf.io/>), Zenodo (<https://zenodo.org/>), or GitHub (<https://github.com/>).

### **5.2.3.3 Pre-prints**

Benefits of posting pre-prints of scientific articles to relevant platforms (for example, <https://psyarxiv.com/>) enable the rapid dissemination of research results and facilitate open peer review and collaboration. Pre-prints should be subject to the same degree of critical appraisal as evidence that has been formally peer reviewed ([FSA Science Council, 2021](#)). It is the norm for papers to be uploaded to pre-print servers while awaiting publication following formal peer review at academic journals, which can take months or years.

### **5.2.3.4 Open access**

Open access publishing improves the visibility and accessibility of articles. Open access articles are downloaded approximately four times more often than those available behind paywalls. FSA reports should, if possible, be published on [www.food.gov](http://www.food.gov).

## **5.2.4 Identifying pathways to impact**

Establishing the impact of FSA research is key in both ensuring effective strategic prioritisation/resource allocation, and in the facilitation of evidence-led policy development. Research impact has been defined as “the demonstrable contribution that excellent research makes to society and the economy”. Research can also have an academic impact in the contribution that it makes to understanding and advancing scientific, method, theory and application. Research impacts may be achieved at different time points and are rarely linear or immediate. It is therefore key that research commissioners and lead researchers meet at different stages of the research cycle, and after project completion, to discuss intended impact, other potential outcomes, and how these will be achieved. The FSA establishing project impact (EPI) process has been designed to ensure that these conversations are documented and assist in resource prioritisation (Establishing Project Impact: Guidance notes).

### **5.2.4.1 Stakeholder mapping and involvement**

There are many well-documented positive impacts of involving different stakeholders, including patients and the public, in health and social research. Positive impacts include development of user-focused funding priorities and research objectives, development of user-relevant research questions, development of user-friendly information, questionnaires and interview schedules, more appropriate recruitment strategies for studies, consumer-focused interpretation of data and enhanced implementation and dissemination of study results. Involving intended users in the research design can also help maximise response rates (see 'Maximising response rates') ([Brett et al., 2012](#)).

Key stakeholders depend on the research topic and may include ([HM Treasury, 2020](#)):

- those responsible for the intervention under consideration: these are the people who have most to gain from evidence that can reduce risk and uncertainty, and from learning what is working and what is not
- those responsible for future policies: this group will require evidence on what worked (and/or did not), why and how, and on transferable lessons
- those responsible for appraisal analysis: they will have the most insight into what evidence and data were missing from the appraisal of the intervention, and what will be useful for the appraisal of future policies
- those responsible for scrutinising government decisions and spend: those that hold government to account are an eager audience for evidence around the efficacy of the intervention's design and delivery, and its impact and cost
- participants/recipients of the policy/intervention: those affected by the policy/intervention are typically also key participants in the evaluation. Their input is required, but they will also have evidence needs and a perspective on what elements of the policy should be focused on
- those delivering interventions: typically, although policies are often designed in central government, they are delivered by others, in many cases through a long delivery chain. Evaluation should be alive to the needs and issues of all those in the delivery chain
- the public (often via the media): a key line of accountability is to the public who are keen to know that government money is being spent wisely, and that we are learning from past experience
- academics/other researchers: academics and other researchers are often able to spend time scrutinising government data. It is important to work with them to ensure the best use of the research evidence is being made and the maximum learning is being extracted

A good place to start is working out who might benefit from the research through conducting a stakeholder analysis. There is no one right way to prioritise those to engage with. When conducting a stakeholder analysis, the first question to ask yourself is who might be interested in your research. Different individuals/organisations may have different interests. The second question to consider is if there are any groups or organisations who might have the ability to influence your ability to achieve impact indirectly. Influence over impact can work in two ways: those who have the ability to facilitate your impact and those who have the ability to block your impact. Consider how influential each of the interested groups might be, whether they might facilitate or block your impact, rating them high, medium or low. The final question to consider concerns the level and nature of impact for each group who engages with your work. In particular, it is important to consider if there may be a negative impact here, so you can ameliorate this if possible. The final task is to use the information collected by addressing the above three questions to prioritise who to reach out to first ([Reed, 2019](#)).

#### **5.2.4.3 Impact areas**

Research projects can contribute to society and the economy through several impact areas, including (but not limited to) policy development, regulatory change, industry action, change in consumer behaviour, or through broader areas such as international collaboration or improving the evidence base as a foundation for further research (Establishing Project Impact: Guidance notes).

## **5.2 and exploitation of research findings.5 Dissemination**

Dissemination refers to the process of getting research findings to the people who can make use of them, to maximise the benefit of the research without delay ([NIHR, 2019](#)). Before disseminating research, it is important to identify relevant audiences through stakeholder mapping (see 'Stakeholder mapping and involvement'). It is important to produce targeted outputs that are in appropriate format for the target user. For example, different tailored outputs (for example, research reports, policy briefs, newsletters) or activities (for example, webinars, science festivals) will be appropriate for different stakeholders (for example, decision-makers, patients, other researchers, healthcare professionals, the public at large). Internally, results should be disseminated through the appropriate research and evidence programme group and other recommended channels (for example, Yammer, bitesize sessions).

### **5.2.5.1 Scientists and healthcare professionals**

Scientists and healthcare professionals are used to reading scientific articles and reports. Presenting research results at scientific conferences or publicising these on social media (for example, Twitter) can also reach this audience.

### **5.2.5.2 Members of the public**

Members of the public are used to reading or hearing about research findings through social media (for example, Twitter, Facebook, Instagram), charity or community group newsletters or bulletins, and through trusted organisations such as the National Health Service. Public engagement events or science festivals (for example, [SoapboxScience](#)) are useful forums for reaching members of the public.

### **5.2.5.3 Policymakers outside the FSA**

When disseminating research results to decision-makers in local and national government (outside the FSA), the following practical recommendations have been made: (1) make research relevant to policymakers; (2) invest time to develop and maintain relationships with policymakers (for example, through workshops and unstructured discussions); (3) utilise 'windows of opportunity' (for example, find opportunities to regularly present research to policymakers so they become interested and invested in the work); and (4) adapt presentation and communication styles to the audience (for example, communicate via e-mail, slide decks, discussions and presentations in bitesize sessions, posters and strategic discussions and presentations in meetings with senior management).

## **5.3 Research procurement**

### **5.3.1. Promoting tender opportunities**

Reaching a large pool of potential tenderers is important. Potential tenderers can be reached through mailing lists (for example, university mailing lists, small business mailing lists), newsletters and social media (particularly Twitter). It is recommended that as many recruitment channels as possible are used. As timelines for submitting tender applications may be shorter than for academic funding applications, it is important to aim to publicise tender opportunities as quickly as possible after they have been decided upon. Speaking to university department or business Heads/Chairs/CEOs to ensure they are aware of FSA tender opportunities more broadly (for example, not necessarily in relation to a

specific call) is also an important strategy.

### **5.3.2 Engaging with internal and external peer reviewers**

Consulting trusted expert peer reviewers inside and outside the FSA is important for ensuring that quality evidence is produced. The role of the peer reviewer is to provide critical challenge of the study design and data analysis and suggest constructive solutions. The peer reviewer might consider the entire analytical process from the user requirements through to the interpretation of the results or focus on particular aspects of the research project.

When the project timeline is tight, available resource is limited, analytical complexity is low, and risks involved are minimal (see 'Producing a risk management plan'), internal peer review by an FSA colleague (including members of the Advisory Committee for Social Science) is sufficient ([HM Treasury, 2015](#)). In the cases of very complex analysis or analysis that drives a significant business critical decision, however, commissioners of analysis or analytical assurers may wish to request external peer review of a piece of analysis.

#### **5.3.2.1 External academic peer reviewers**

Appropriate external academic peer reviewers are typically experts who have published scientific articles in the area of interest. To give an informed and unbiased opinion on a piece of research, reviewers should have in-depth knowledge of the topic area and methodological expertise. They should not be known to have particularly strong views or opinions on the topic (unless this can be balanced with additional reviews from people with a more neutral stance) or have any conflicts of interest which could bias judgments and lead to lack of objectivity. In addition, they should not currently be working at the same laboratory or institution as the study author ([HM Treasury, 2015](#)).

#### **5.3.2.2 External non-academic peer reviewers**

For many projects, it is important to draw on the expertise of non-academic experts. Appropriate external non-academic peer reviewers are typically professionals who have worked in the area of interest and developed domain-specific expertise (for example, novel foods). To give an informed and unbiased opinion on a piece of research, reviewers should have in-depth knowledge of the topic area and methodological expertise. They should not be known to have particularly strong views or opinions on the topic (unless this can be balanced with one or two additional reviews from reviewers with a more neutral stance)

and should not currently be working at the same business or organisation as the study lead/team ([HM Treasury, 2015](#)).