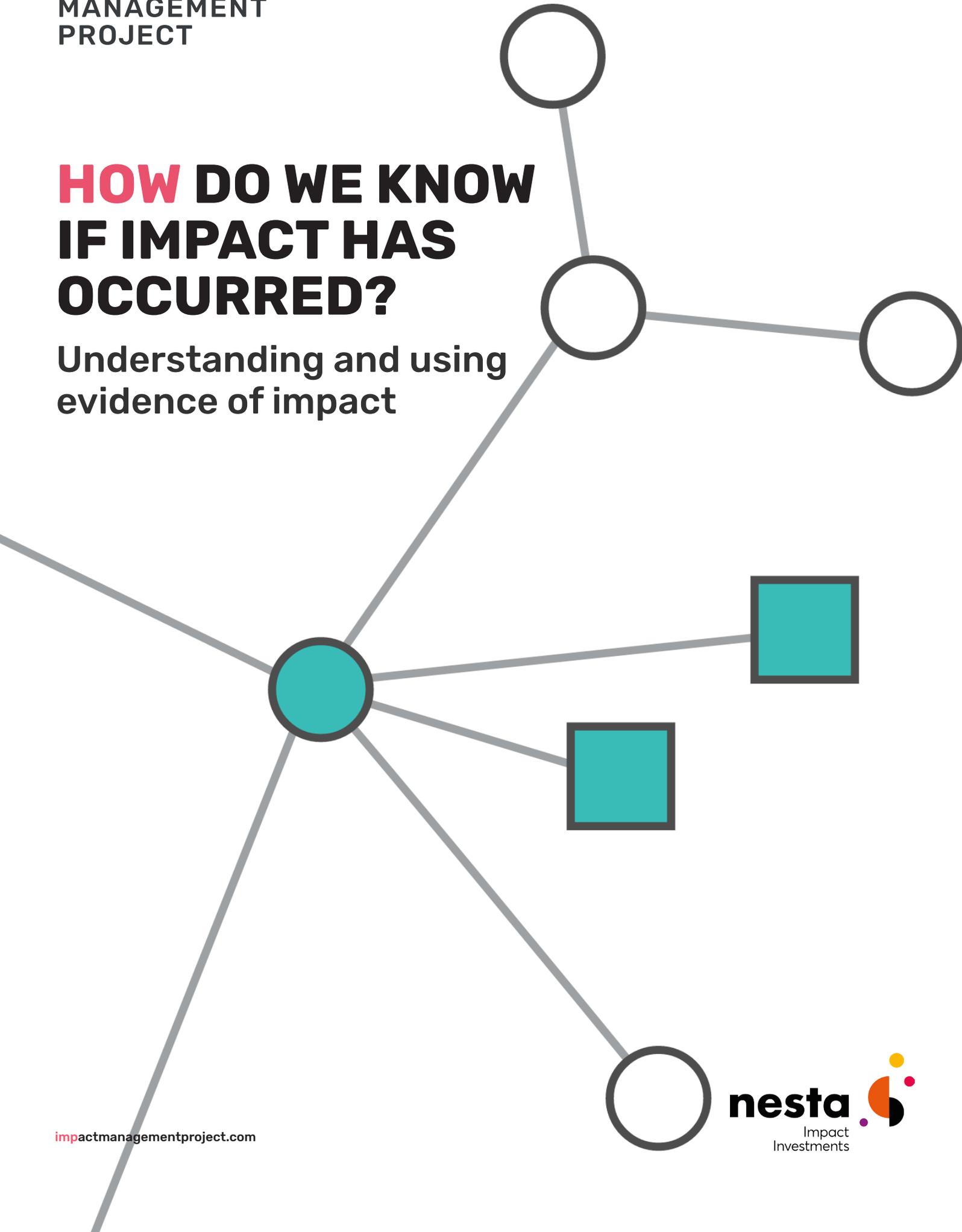


HOW DO WE KNOW IF IMPACT HAS OCCURRED?

Understanding and using
evidence of impact



INTRODUCTION

Contents

INTRODUCTION.....	2
CONTEXT SETTING.....	3
Why does evidence matter?.....	3
What is evidence of impact?.....	3
EXISTING EVIDENCE.....	4
Collecting existing evidence.....	4
Using existing evidence.....	4
NEW EVIDENCE.....	5
Collecting our own evidence.....	5
MAXIMISING THE VALUE OF EVIDENCE.....	7
Choosing the best methodology.....	7
Sharing evidence.....	7
MORE INFORMATION.....	8

The Impact Management Project is a collaborative effort by over 700 organisations, from different context and countries, to agree on shared fundamentals for how we talk about, measure and manage impact - and therefore our goals and performance.

Shared fundamentals for describing the effects that different businesses - or portfolios of businesses - have on people and planet help different actors in a value chain figure out what information they need to manage those effects. Collecting evidence of impact is an important step towards figuring out which material positive and negative effects are occurring, before making decisions to improve the experience of people and planet.

Over the last six months, Nesta has been working with the Impact Management Project to explore how we think about evidence of impact. This document provides guidance on both using existing evidence of impact, and building your own evidence base.

This report is the product of a series of discussions among a wide range of practitioners as part of the Impact Management Project. The report has been co-authored by Nesta and the Impact Management Project Team. Please direct any feedback or further enquiries about this report to:

team@impactmanagementproject.com

Acknowledgements

The Impact Management Project is deeply grateful to all project participants, who provided excellent insight into the topic of evidence and to Nesta co-author Lucy Heady.

CONTEXT SETTING

Why does evidence matter?

The assertion that we should use evidence to guide our decisions seems so obvious as to be hardly worth saying. Why wouldn't we use what we already know to guide our decision-making? Scratch the surface, however, and things immediately seem less clear. Different disciplines prioritise different types of evidence. Techniques from one field might be imported to a new context without appreciating important differences. And when different types of evidence disagree, or there is very little evidence that anything works, it can seem easier and cheaper to abandon evidence altogether.

But without evidence we are doomed to repeat our mistakes. Ideas with strong intuitive appeal, such as scaring children off crime by having them spend time in prisons or putting teenagers off motherhood by giving them artificial crying babies, turn out to be actively harmful¹. Achieving social impact is hard; we must never assume that something works because the logic is appealing. By building the evidence base, across contexts and with different approaches, and learning and adapting based on what we find, we can achieve positive impact, and are more likely to attract investment.

Using and collecting evidence of impact need not require a PhD in social science or a million-dollar budget. The norms below set out principles for interpreting existing evidence and collecting new evidence in a cost-effective way. Wherever possible we have signposted to more detailed resources that demonstrate the practical application of these principles.

What is evidence of impact?

We define 'evidence of impact' as:

"The available body of facts or information that can be used to judge to what extent (or not) impact has occurred."²

We collect lots of different types of information through the impact management process, only some of which is used to understand the experience of people and planet, and is therefore 'impact evidence'. Impact evidence helps us understand what impact has occurred, and enables us to reduce our level of evidence risk.³

When using evidence to make judgements we take into account how good the evidence is likely to be. There are many attributes which help us determine whether we can have strong grounds for believing information about impact constitutes 'good' evidence. Frameworks like [Standards of Evidence](#) help us assess what level of confidence we have, based on the type of evidence we collect.

¹For example: see the [evaluation](#) of the juvenile delinquency prevention programme 'Scared straight'

²Based on Oxford English Dictionary definition, adapted to be specific to whether impact (positive or negative) has occurred.

³There are many other types of evidence used to make decisions within the impact management process - such as expert opinion - but these do not constitute evidence of impact

EXISTING EVIDENCE

Collecting existing evidence

When setting our impact goals and selecting or designing delivery models (e.g. a business or programme), we often draw on existing evidence. To do this we consider what available evidence exists across all of the [dimensions of impact](#).

For example, when evaluating an existing enterprise, perhaps as a prospective investor, we might consider how far data exists across all the dimensions (Figure 1, below) and what this assessment tells us about an enterprise's ability to meet a [set of goals](#). This helps us work out where gaps, or weak evidence, exist in the existing evidence base for one or more group of people (or the planet) affected, and where new evidence may be needed.

Using existing evidence

There are a number of considerations which help assess whether existing evidence of impact may be 'good' enough:

COMPLETENESS

- Is the data set complete and comprehensive?

- Has data been collected from a range of relevant stakeholders?

ACCURACY

- Does the evidence conform to established criteria for quality and rigour?⁴
- Is the evidence the product of external review or an independent assessment?⁵

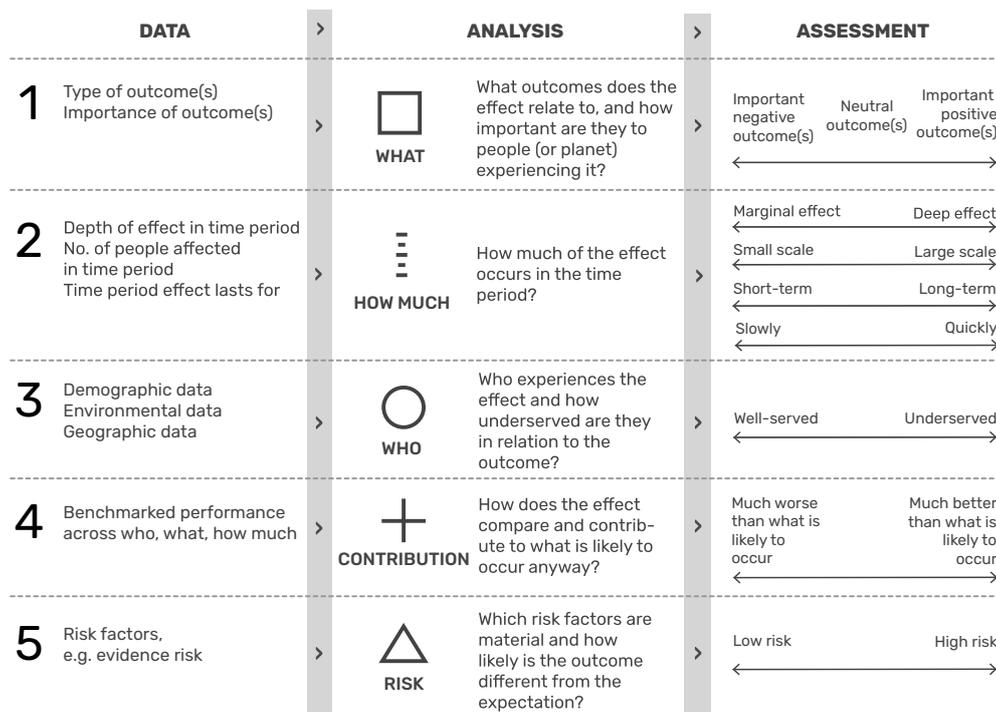
RELEVANCE

- Does the evidence directly correspond to the same impact goals across all the dimensions of impact? Are there any gaps?

When examining an existing evidence base, the more individual studies (or sets of information) that demonstrate these features and share the same findings, the stronger the evidence base.⁶

However, existing evidence will rarely tell us everything we need to know about the likely impact of a delivery model in a given context. We therefore often need to collect our own evidence to re-affirm what we know, or fill in evidence gaps. Resources such as [Zie's Evidence Gap Maps](#) help us identify where these evidence gaps are likely to be.

Figure 1: The dimensions of impact



⁴ See: [USAID data quality checklist](#)

⁵ See: [Nesta standards of evidence](#); [Levels of evidence in Medicine](#); [What Works?](#)

⁶ See: [DFID's 'Assessing Strength of Evidence' guide](#)

NEW EVIDENCE

Collecting our own evidence

Collecting our own evidence better enables us to substantiate or refute existing evidence, and further understand what impact is being delivered, in order to improve the experience of people and planet. We may also need to understand whether this experience changes after engagement with an enterprise has ended. All of the evidence we gather informs the decisions we make about our strategy and the goals we set and re-set.

The type of evidence collected will vary according to what level of evidence is necessary and proportional for an enterprise to understand their effects on people and planet. There are two broad dimensions along which we can think about proportionality:

1. What type of data we collect: the amount of confidence we have in our data is often related to the type of data we use.

- Direct measurement of change in effects enables a greater level of certainty.

For example, change in nutrition status or reduction in CO2 emissions.

However, sometimes these measures are not possible, very difficult or very expensive to collect for every demographic and outcome type. This could be due to the timescale within which effects are realised, customer confidentiality, a lack of appropriate measurement techniques or simply cost. In these cases, it is common to collect data on proxies for the final effect, for example home improvements might be a proxy for increased incomes or hospital waiting lists might be a proxy for a well-functioning health service.

- Activity or output data is often used as a proxy as it is typically much easier to collect this type of data.

For example, counting a child that has received a vaccination against measles as a proxy for that child not developing measles.

In some cases, the existing evidence of a link between the activity and the outcome is so strong that activity or output data can serve as a proxy for outcomes, for example the number of children receiving vaccinations against measles. When selecting our own indicators or using proxies, we are at risk of enforcing causality assumptions. Where an enterprise is taking a new approach, or the evidence base linking outputs to outcomes is inconclusive, non-substantive or incomparable, these proxy measures will not constitute good evidence of impact, even though they may provide useful evidence that some parts of a [theory of change](#) are valid.

- We also recognise that people and planet will likely experience other impact too, both positive and negative and therefore draw on existing information about the material effects that comparable approaches delivered in comparable contexts to help us judge other effects we might want either to mitigate (if negative) or increase deliberately (if positive).

For example, by consulting existing research on the effects of offshore wind farms, we might learn that some marine animals experience negative impact due to the noise and regular disturbance. This information helps us to dedicate resource to trying to collect new information to uncover and mitigate these negative effects.

- Feedback from people experiencing the effects through a questionnaire or surveying methodology is the most direct and sometimes most reliable indicator that any change in effect is occurring. This method is most reliable where there is existing evidence that people's feedback is strongly correlated with the desired effect. This data may be a proxy (e.g. 'did you take your medicine on time?'), or may be the best metric to indicate what effect is occurring across any of the dimensions (e.g. 'how far has your quality of life improved?' is a good indicator of depth for a wellbeing outcome).

NEW EVIDENCE cont'd

2. How we collect it: the amount of confidence we have in our data is also informed by the methodology we use.

- At a minimum, we collect data (direct measures or proxies) before engagement with an enterprise and after that engagement. However, this approach still leaves potential bias as any change detected is unlikely to be caused only by the engagement, for example if some people would have improved their situation anyway or if there is some other engagement causing change.

For example, maternal health clinics have tracked that 470 mothers received treatment in 2008, 89% of whom safely delivered a healthy child and 97% of whom reported satisfaction with the service.

- The greatest confidence is often obtained through an external evaluation that removes selection bias, considers attribution and/or uncovers other effects.

For example, an evaluation shows that, of all of the treatments available, mothers receiving supplemental nutrition were most likely to give birth to a healthy-weight child.

- Evaluations that use a control group, such as Randomised Control Trials or Propensity Score Matching, help us establish causality. While innovative approaches are bringing the costs of such techniques down, they do often still require a higher financial investment and more technical expertise.

For example, a survey of a similar group of mothers in a neighbouring region reported a 65% safe, healthy delivery rate in the same period, while the maternal mortality rate is 43% below the regional average.

- Being clear about how context affects the delivery of these effects enables consistent replication.

For example, there are multiple drivers of positive maternal health outcomes – including access to good nutrition and a sanitary delivery environment – and people in different contexts will be underserved in different ways in relation to this outcome. This will influence the effectiveness and/or applicability of pre- and post-natal care.

Standards of Evidence are helpful frameworks for assessing the confidence we have in our evidence based on the methodology used.

Figure 2: An example Standards of Evidence⁸



⁸ Puttick & Ludlow (2013) '[Nesta: Standards of Evidence](#)'. These Standards were based on the Maryland Scientific Methods Scale, and the Project Oracle Standards of Evidence.

MAXIMISING THE VALUE OF EVIDENCE

Choosing the best methodology

It is not always appropriate to use the most resource-intensive and complex methodology for collecting evidence. Considering our confidence in existing evidence, and the level of impact risk parties are willing to take, is a helpful starting point.

It is often sensible to use less expensive approaches to test whether it is plausible that impact has occurred while a strategy is still being developed, and then invest in something more rigorous once there is more confidence that the strategy is of sufficient quality.

Collecting, analysing and communicating evidence all come at a cost and so it is important that the evidence we collect is both useful and collected in a cost-effective way.

The following considerations can help us to collect useful evidence:

1. Is there a clear purpose for the evidence?

This might be:

- improving the experience of people or planet
- increasing publicly available evidence about which strategies work (or don't) in relation to a specific population and/or outcome
- reducing risk and uncertainty to increase the likelihood that impact is being delivered as expected

2. Is the value of the evidence likely to be proportionate to its cost? The most valuable evidence is likely to be:

- evidence that helps us manage our [most material effects](#) on people and planet
- focused on building the evidence base where there is greatest uncertainty

3. Is the methodology appropriate for the type of evidence sought?⁹ Draw on the existing evidence base to learn about:

- gaps in existing evidence, and why they exist
- likely risks involved (e.g. ethical, cultural or practical)
- what data-collection methods are often used in this context

4. Who will benefit most from the evidence?

This should inform:

- the standards of evidence employed
- who carries the cost¹⁰

Sharing evidence

Sharing evidence maximises its value by enabling more enterprises to learn and improve their delivery models. How well evidence is used is directly related to how it is shared.

To maximise the value of evidence, we should adhere to the following principles when reporting evidence:

- **TRANSPARENT:** Reports should include what was tested, time-stamped results, methodology, data quality, context and stakeholders.
- **TIMELY:** Evidence should be shared as soon as possible to make it useful to others, where possible considering the timeline of decisions the data may influence.
- **COMPREHENSIVE:** Reports should make clear how well the evidence can say what would have happened in the absence of the enterprise, report any negative findings and any assumptions involved in arriving at estimates of impact.
- **CONTEXTUALISED:** New evidence should always be presented in the context of other existing evidence and our original goals. This allows us to see where results differ and what new lessons have been learned.

⁹ See: [IDInsight decision-focused evaluation](#)

¹⁰ Where a resource-intensive methodology is being employed to answer questions that will help build the field, it is less likely that it is appropriate for the enterprise to carry the cost alone.

MORE INFORMATION

Suggested resources

Identifying good evidence

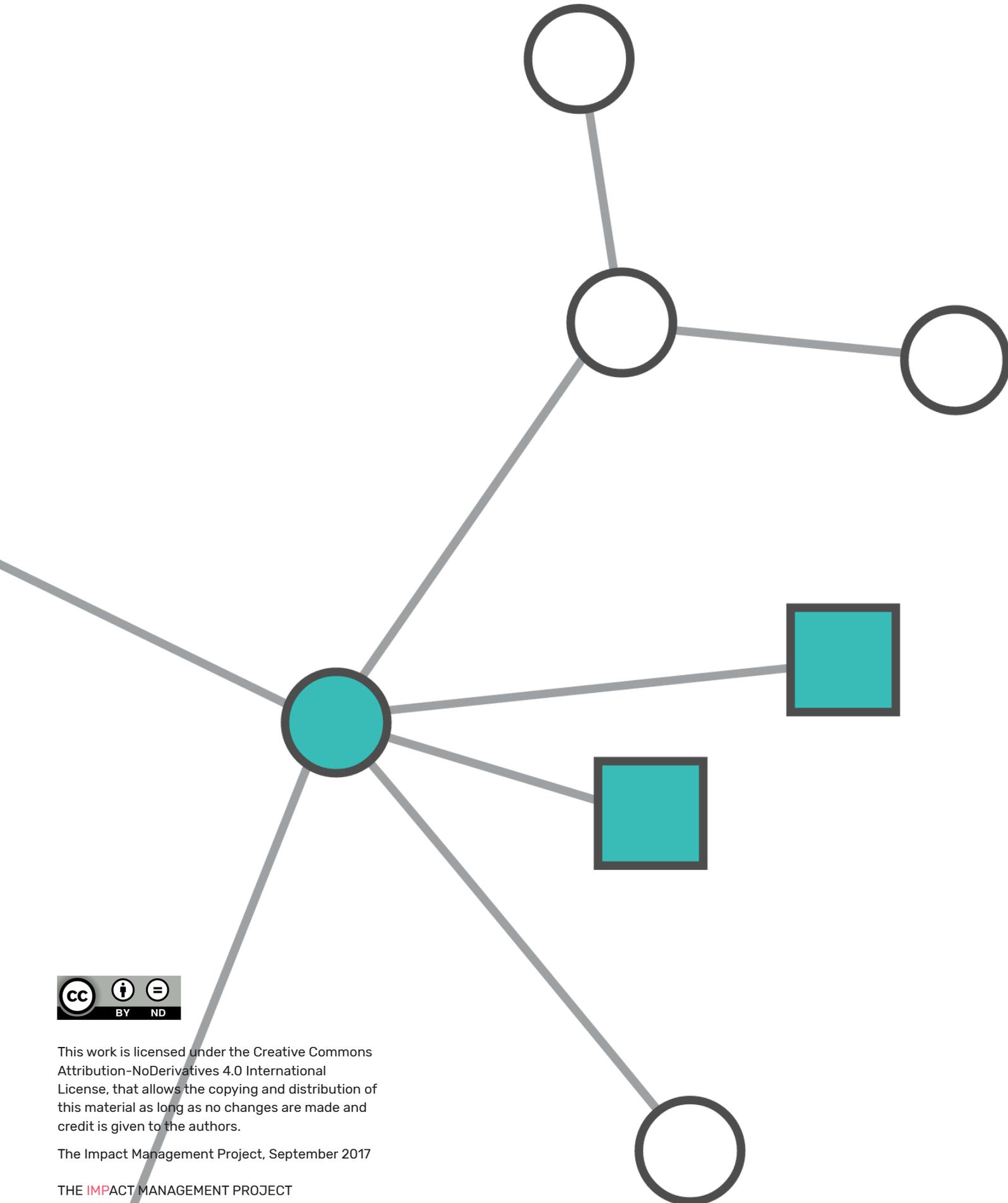
- Bond [principles for assessing the quality of evidence](#)
- DFID [Assessing the Strength of Evidence guide](#)
- Nesta [Standards of Evidence](#)
- USAID [Data Quality Checklist](#)
- Evidence-Based Medicine [resources](#)
- Alliance for Useful Evidence [Using Research Evidence](#)
- ImpactMatters [Impact Audit Standard](#)

Finding existing evidence

- Dartington [Social Research Unit](#)
- Beam Exchange [Evidence Map](#)
- 3ie [Evidence Gap Maps](#)
- What Works Centre [papers](#)
- Campbell Collaboration [Library](#)

Choosing a methodology

- IDInsight [Decision-Focused Evaluation](#)
- CDI [Review of evaluation methodologies](#)
- IPA [Research resources](#)
- SSIR [How Impact Investors Measure Impact](#)
- OHRC [Collecting Data Ethically](#)



This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International License, that allows the copying and distribution of this material as long as no changes are made and credit is given to the authors.

The Impact Management Project, September 2017

THE **IMPACT** MANAGEMENT PROJECT