



**GOVERNMENT
OUTCOMES
LAB**

Setting and Measuring Outcomes Academy for Social Justice Commissioning

14th March 2019



@ukgolab
@AcademyJustice

Introducing us



**Andreea Anastasiu, Senior Policy
Engagement Officer, GO Lab**



**Neil Stanworth, GO Lab Fellow of Practice
& Director, ATQ Consultants**

About the GO Lab



Joint partnership
between UK
Government &
Oxford University

Established in 2016

Based at the
Blavatnik School of
Government, in
Oxford

Centre of academic research and practice with a mission to improve the provision of public services to tackle complex social issues, with a focus on outcome based models

Support from the GO Lab



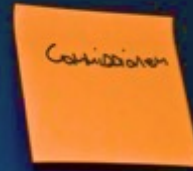
Knowledge Hub for
SIBs & OBC

golab.bsg.ox.ac.uk

Guides &
resources



SIB
Readiness
Framework



SIB
projects
database

HOW TO
GUIDE
CONTRACTING
AND GOVERNANCE

Advice
surgeries



Events &
workshops



Webinars



Regional
SIB
Knowledge
Clubs



Fellows of
Practice

Digital knowledge hub for outcomes commissioning

GOVERNMENT OUTCOMES LAB

About Our Projects Knowledge Guidance News & Events

The GO Lab is a centre of academic research and practice for outcome based contracting & social impact bonds

This is our digital hub for resources and insights from across this emerging field

Event
Save the date: Social Outcomes Conference 2019

Jobs
Career opportunities >

Blog
Has social prescribing come of age? >

- In-depth case studies from UK & beyond
- SIB Projects database - UK
- Policy news, interviews, blogs
- Publications library
- Technical guides
- Webinars
- Advice surgeries...
- ...and much more


Resources for commissioners

Setting and measuring outcomes 2 chapters

Setting and measuring outcomes

A guide to identifying, defining and measuring outcomes for the purposes of outcome based commissioning

Download as PDF



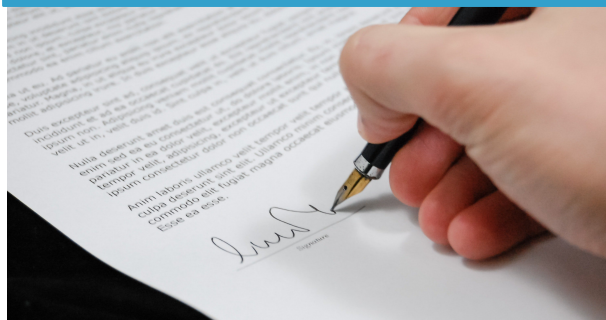
- Chapter 1
- General guidance
- Introduction and overview
- Understanding outcomes, measures and metrics
- Different types of outcome
- Individual versus cohort measurement
- Developing an outcomes framework
- Identifying the right outcomes

About the guide

This guide provides advice on the processes of identifying, defining and measuring outcomes when putting in place an outcomes-based contract or social impact bond. It is aimed mainly at local commissioners, but should also be useful to providers of services and interventions who are involved in

How to guides:
golab.bsg.ox.ac.uk/guidance

WEBINAR: Public contracting for social impact bonds
19TH MARCH 2019
golab.bsg.ox.ac.uk/news-events



BLAVATNIK SCHOOL OF GOVERNMENT UNIVERSITY OF OXFORD GOVERNMENT OUTCOMES LAB



Advice Surgeries

Book online:
golab.bsg.ox.ac.uk/guidance/advice-surgeries

Session overview

- 11.00 *Welcome remarks*
- 11.10 **Session 1 – why measure outcomes**
- 12.00 **Session 2 – identifying outcomes**
- 13.00 *Lunch*
- 13.45 **Session 3 – measuring outcomes**
- 14.45 *Break*
- 15.00 **Session 4 – setting outcomes targets and outcomes-based payment**
- 16.00 *Close*



**GOVERNMENT
OUTCOMES
LAB**

Introductions





**GOVERNMENT
OUTCOMES
LAB**

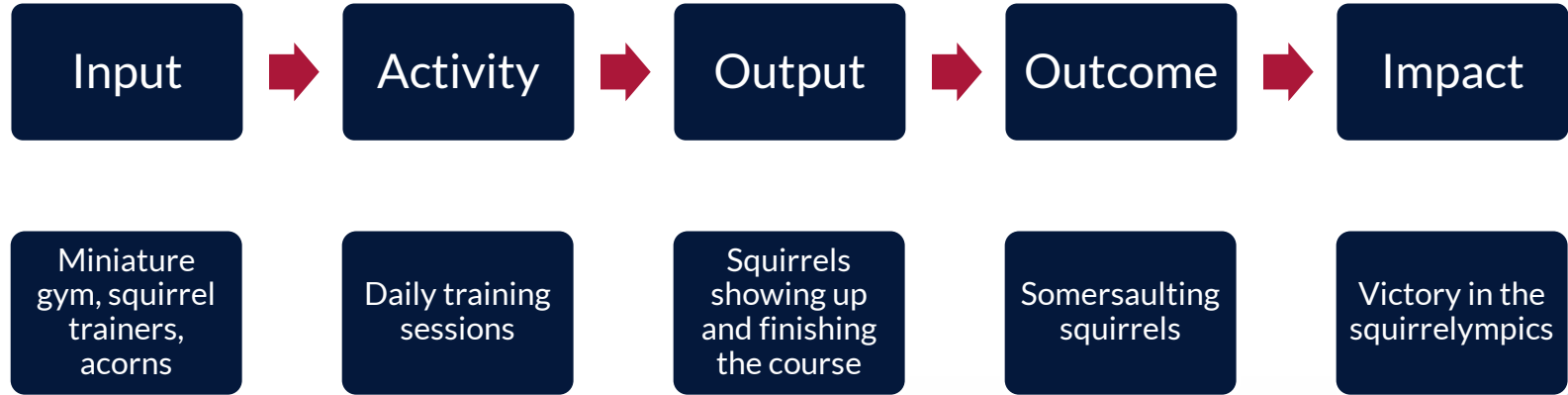
1: Why measure outcomes?



Why set and measure outcomes?

1. To **manage performance** / learn how to get better
2. To **evaluate** whether something works (but not *why* it works – or doesn't)
3. To provide a **target** or means for payment
4. Because I've been told to

Logic model / theory of change



Outcomes vs Impact



Impact measurement answers that “what would have happened anyway?” question

SSIR - Ten Reasons Not to Measure Impact—and What to Do Instead

https://ssir.org/articles/entry/ten_reasons_not_to_measure_impact_and_what_to_do_instead

Key concepts

Outcome	What changes for an individual (or other defined unit, such as a family) as the result of a service or intervention (e.g. improved learning outcomes)
Measure / Indicator	The specific way the commissioner chooses to determine whether that outcome has been achieved (e.g. a test score)
Target / Metric / Trigger	The specific value attached to the measure for the purposes of determining whether satisfactory performance has been achieved (e.g. a test score of 95 out of 100 or improvement of 30 points in a test score over a 5 month period).

Outcomes, measures, targets

Example: Youth unemployment reduction

Outcome	The young person is in employment
Outcome measure <i>(also termed an indicator)</i>	Confirmation from the employer that the person is employed by them
Outcome target <i>(also termed metrics or triggers)</i>	The young person is in continuous employment of a minimum of 16 hours per week for a defined period or That 20% of the total cohort are in continuous employment for a defined period on average

Outcomes, measures, targets

Example: Emotional wellbeing

Outcome	Improved emotional wellbeing of young people
Outcome measure <i>(also termed an indicator)</i>	An identifiable improvement in young people's resilience and ability to deal with challenges using the Strengths and Difficulties Questionnaire (SDQ)
Outcome target <i>(also termed metrics or triggers)</i>	The young person reduces their total SDQ score by a defined number over a specified period or that there is a mean reduction in the average score across the cohort as a whole

Developing an outcomes framework

An outcomes framework should include:

- ✓ the outcomes to be used;
- ✓ the measures to be applied to each outcome;
- ✓ the specific targets to be applied to each measure, that determine the desired level of success; and
- ✓ when measurement takes place.

An outcomes framework involves:

- ✓ An iterative and progressive process
- ✓ Negotiation & consensus - building agreement among the contracting parties
- ✓ Review and revision (if necessary)



GOVERNMENT OUTCOMES LAB

Questions



@ukgolab



**GOVERNMENT
OUTCOMES
LAB**

2: Identifying outcomes



Hard & soft outcomes and measures

Type of outcome	Definition	Advantages	Disadvantages
Hard	Can be measured objectively	<ul style="list-style-type: none">▪ Simpler to measure▪ No risk of disagreement about achievement	<ul style="list-style-type: none">▪ Not always available▪ May not capture sustained impact▪ May not reflect what matters to service users
Soft	Requires subjective assessment	<ul style="list-style-type: none">▪ Useful when no hard outcome is available▪ Can be used to test progress made▪ Measures whether the service meets user expectations	<ul style="list-style-type: none">▪ Consistency of measurement can be difficult▪ Potential for disagreement about achievement

Proxy outcomes & measures

An indirect measure of the desired outcome strongly correlated to that outcome, used when direct measures of the outcome are unavailable or cannot be measured.

Proxy outcomes & measures

Proxy measure

- Reduced hospital admissions
- Off benefit
- Reduced reconvictions

True outcome

- Improved health
- In employment
- Reduced reoffending

Risks

- Reduced attendance might be due to other factors
- Person may cease claiming benefits without finding work
- Many offences go undetected

Lead and final outcomes

Lead/indicative outcome

- Improved school attendance and/or behaviour
- Engagement in part time or voluntary work
- Family attendance at a parenting support programme.

May lead to



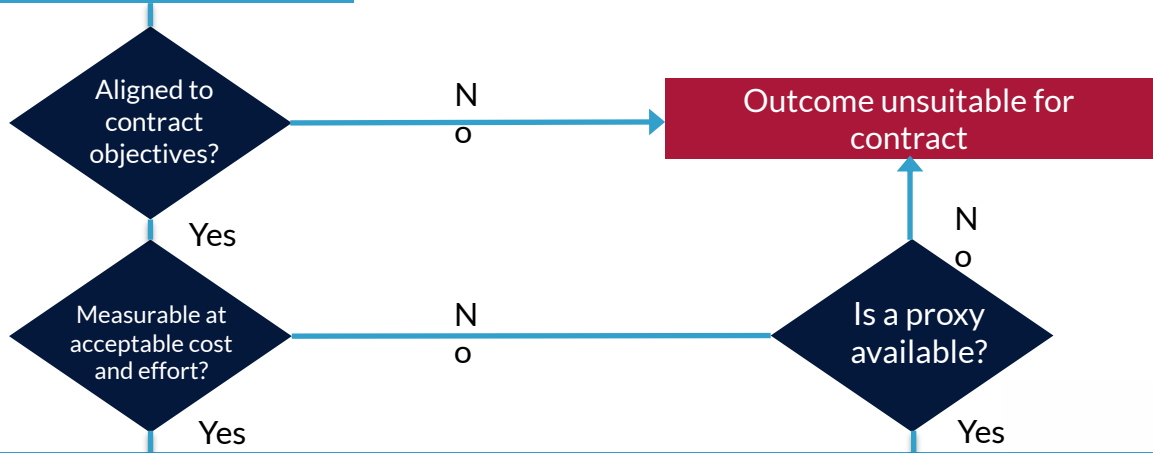
True final outcome

- Improved attainment and reduced risk of exclusion
- Full time employment
- Reduced risk of a child becoming 'in need' or on a Child Protection Plan.

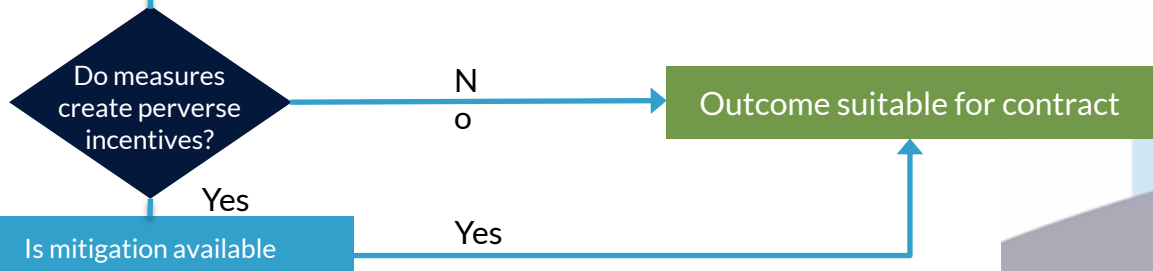
Identifying outcomes

A simplified decision-making tree

Identify possible outcome



Identify requirements for measurement and consider targets that can be used





**GOVERNMENT
OUTCOMES
LAB**

Group exercise





GOVERNMENT
OUTCOMES
LAB

Questions



@ukgolab



**GOVERNMENT
OUTCOMES
LAB**

3: Measuring outcomes



Measuring outcomes

Some general considerations

- ✓ whether the outcome is best measured using an individual or collective (cross-cohort) approach
- ✓ whether the data needed to measure the outcome are already collected for another purpose, for example government statistical returns or internal performance management;
- ✓ if not, whether the data collection requires significant investment in new collection processes and systems;
- ✓ who will be responsible for collecting the data and whether they have the capacity to do so; and
- ✓ whether the data need to be independently checked and validated

Key issues and questions

- Cohort vs individual
- Hard vs. soft
- Binary vs. continuous
- Use of lead / progression measures
- Data availability

Cohort v individual measurement

Cohort outcome measurement	Individual outcome measurement
<ul style="list-style-type: none">▪ Works best when the current adverse outcomes vary across the cohort▪ Usually requires comparison with a group who did not receive the intervention▪ Does not normally require a separate calculation of deadweight	<ul style="list-style-type: none">▪ Works best when the cohort are experience similar adverse outcomes▪ Does not usually involve a comparison group or other baseline▪ Requires good evidence of the likely level of deadweight

Cohort v individual measurement Examples

HMP Peterborough SIB

- 7.5% reduction in reoffending across all SIB cohorts, against a national comparison group compiled using Propensity Score Matching
- 10% reduction in individual annual cohorts



DWP Innovation Fund

NATURE OF OUTCOME	MAXIMUM PRICE OF OUTCOME
Improved attitude towards school	£700
Improved behaviour	£1,300
Improved attendance	£1,400
Entry level qualification	£900
NVQ Level 1 or equivalent	£1,100
NVQ Level 2 or equivalent	£3,300
NVQ Level 3 or equivalent	£5,100
Entry into employment	£3,500
Sustained employment	£2,000

The maximum amount payable per individual is £11,700. The figure is based on 3 years of Annually Managed Expenditure (AME) savings.

Hard & soft outcome measures

Example: reducing entry to local authority care

The Turning the Tide Social Impact Bond (SIB), commissioned by North Somerset Council, measures reductions in the number of older children entering local authority care as a result of family dysfunction and stress.

It uses a combination of:

- **A soft measure** of improved family / parental capacity and resilience, measured through [Triangle Consulting's Family Plus™ Star](#); and
- **A hard outcome measure** relating to the length of time children at risk of entering care are able to remain with their families.

Binary and continuous measures

Type of outcome	Employment	Offending	Child Protection
Binary	In work/not in work	Not convicted/convicted	Looked after/not looked after
Continuous	Length of time in work Quality of job Progression to full time work	Reduced frequency Reduced severity	Length of time not in care Escalation/de-escalation to formal Child Protection

Data collection options

Data type	Pros	Cons
Administrative data	<ul style="list-style-type: none">• Highly accurate• Low cost	<ul style="list-style-type: none">• May not exist• May not cover population of interest• May not directly address question of interest
Primary data	<ul style="list-style-type: none">• Directly addresses question of interest	<ul style="list-style-type: none">• High cost• Possibility of bias

- ✓ Find out whether the required data is already collected for other purposes
- ✓ Do not to make assumptions about the availability of data from other parties or the ability of those parties to collect data on your behalf.



**GOVERNMENT
OUTCOMES
LAB**

Questions





**GOVERNMENT
OUTCOMES
LAB**

Group exercise



@ukgolab



**GOVERNMENT
OUTCOMES
LAB**

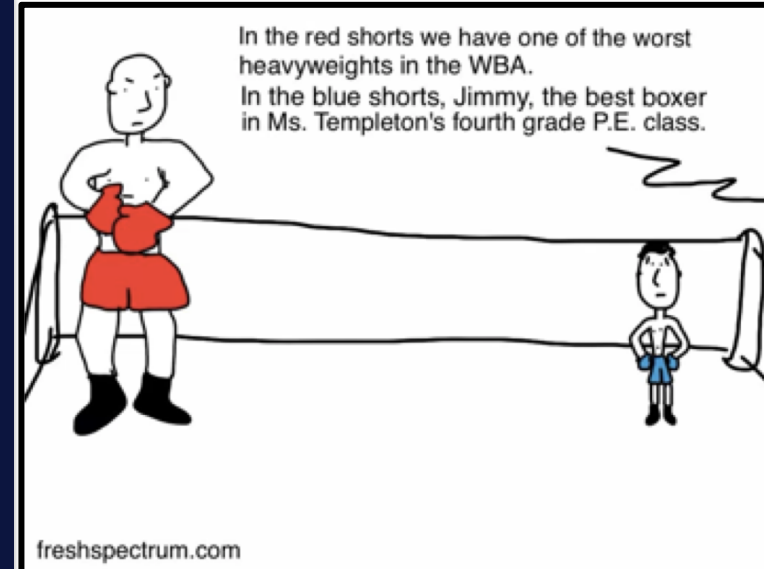
4: Setting outcomes targets and outcomes-based payment



Counterfactual

A measure of what the outcome would have been for programme participants if they had not participated

- What would have happened in the absence of the intervention
- Cannot be observed, so must be estimated using a comparison group
- The challenge is to identify a treatment group and a comparison group that are statistically identical



Source: Chris Lysy, freshspectrum.com/what-is-evaluation-anyway

Deadweight

Deadweight: outcomes which would have occurred without the programme or intervention.

- *Did this programme make a difference or would changes have occurred anyway?*
- *How many /what proportion of outcomes would have been achieved anyway?*



Setting outcome targets

- Specify the level of against which success will be judged / outcome payments will be made
- Must align with the social objectives of the contract (what does good look like?)
- Should take account of existing evidence around the effectiveness of the intervention and the effect of deadweight
- Could be as simple as the achievement of a hard binary outcome
- But should usually also consider:
 - Progression towards the main outcome
 - The sustainment of the outcome and/or further improvement

Achievability

- Find out how successful others have been in the past when attempting to achieve the same (or similar outcome)
- Ask providers what they think is achievable
- Use a combination of lead / progression targets which are easier to achieve, and stretch targets which are harder

Perverse incentives

Perverse incentives are incentives that encourage contract stakeholders to behave in a way that is detrimental to contractual goals even if some outcome metrics improve

Some examples:

- Measures and metrics that link to absence of or reduction in referrals to a statutory agency
- A simplistic binary outcome measure, which can easily be 'failed' by a high proportion of the cohort – this can lead to 'parking' of those who can no longer achieve the outcome
- Setting a metric that is achieved with varying degrees of effort for different members of the cohort – this can lead to creaming (focussing the attention on those easier to help)
- A single time-related milestone at which the person receiving the information is deemed to have achieved a positive outcome – sometimes termed cliff edge

How we have learnt to set better outcomes

Example: children on the edge of care

Metrics relating to the prevention of entry to care have evolved from:

- Cohort-wide measurement of the total number of days in care compared to a baseline or comparison group (complex and time-consuming to measure)
 - *via*
- Bullet payments paid as an individual tariff at intervals (e.g. 6 months) reflecting the length of time out of care (simpler but prone to distortion or perverse incentives) *to*
 - *to*
- A payment per individual for each care day avoided, totted up and paid at intervals (combining the advantages of a tariff with flexibility and avoidance of distortion)
- A soft measure is often used alongside this e.g. the Family Star



GOVERNMENT OUTCOMES LAB

Questions





GOVERNMENT OUTCOMES LAB

Group exercise



@ukgolab

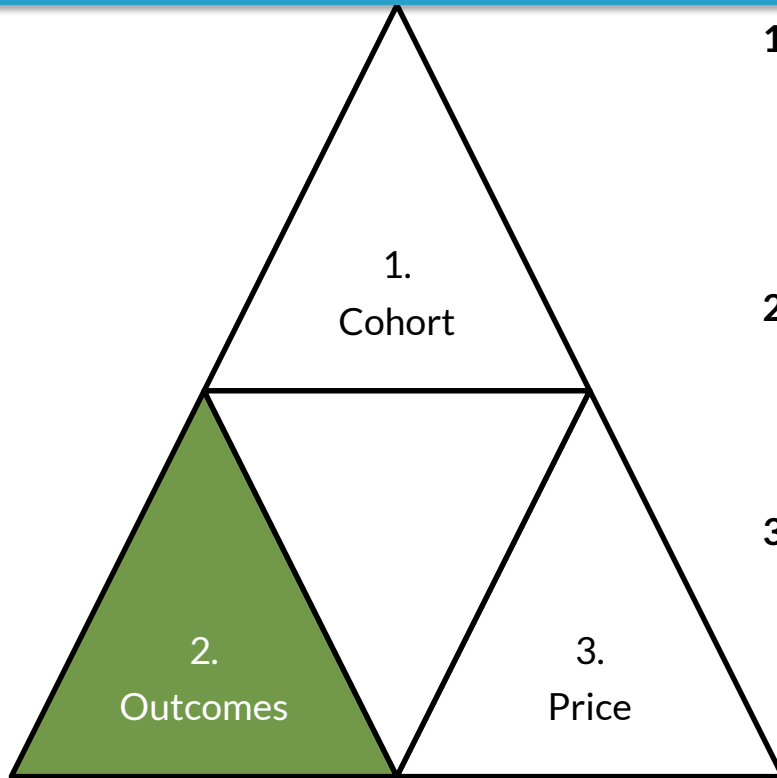
Three top tips for commissioners (and others)

- ✓ Simple is not always best
- ✓ Avoid superficially attractive hard and binary metrics that can cause major problems

- ✓ Be flexible
- ✓ Measures and metrics will change as you develop your framework and talk to stakeholders

- ✓ Learn from others' experience
- ✓ Use existing metrics where they are available – they save time & effort

Designing a robust framework



- 1. Tightly defined eligible cohort**
 - Clear, objective criteria
 - Understanding of how far participants are from the desired outcomes
 - Independent referral / identification mechanism
- 2. Alignment between payable outcomes and policy objectives**
 - Logical link between activity, outputs and outcomes
 - Adequate period of time for tracking
 - A way to tell if the effect has 'stuck'
- 3. Accurate price-setting of outcomes**
 - Robust estimate of likely level of benefit vs what would happen anyway ("deadweight")
 - A way to get confidence that any outcomes are caused by the intervention ("attribution")
 - Suitably long outcome tracking-period

Whilst it is not practical for these aspects to be perfect, commissioners should focus on them to avoid perverse incentives

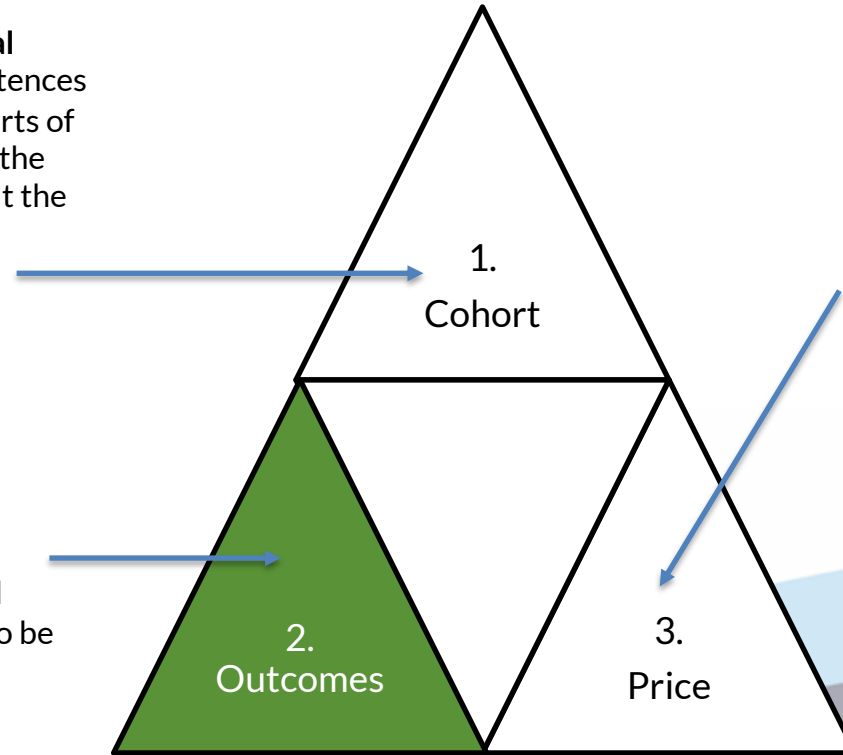
What does 'good' look like?

Example: HMP Peterborough SIB

1. Eligibility **criteria clear and impartial** (offenders in Peterborough prison sentences of less than 12 months over three cohorts of 1,000) and **referral was independent** (the service was offered to everyone who fit the eligibility criteria, participation was voluntary).

2. Policy intent to **reduce reoffending** and the project **paid on a cohort-based reduction in reconviction**, a valid proxy measure tracked over a suitably long period to be durable.

3. Includes **real-time comparator group in the payment mechanism**: commissioners can say with confidence that the **outcomes they paid for are attributable to the ONE service, and additional to business as usual**




Resources for commissioners

Setting and measuring outcomes 2 chapters

Setting and measuring outcomes

A guide to identifying, defining and measuring outcomes for the purposes of outcome based commissioning

Download as PDF



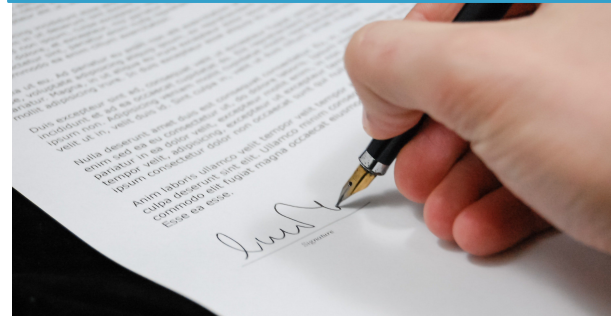
- Chapter 1
- General guidance
- Introduction and overview
- Understanding outcomes, measures and metrics
- Different types of outcome
- Individual versus cohort measurement
- Developing an outcomes framework
- Identifying the right outcomes

About the guide

This guide provides advice on the processes of identifying, defining and measuring outcomes when putting in place an outcomes-based contract or social impact bond. It is aimed mainly at local commissioners, but should also be useful to providers of services and interventions who are involved in

How to guides:
golab.bsg.ox.ac.uk/guidance

WEBINAR: Public contracting for social impact bonds
19TH MARCH 2019
golab.bsg.ox.ac.uk/news-events



BLAVATNIK SCHOOL OF GOVERNMENT UNIVERSITY OF OXFORD GOVERNMENT OUTCOMES LAB



Advice Surgeries

Book online:
golab.bsg.ox.ac.uk/guidance/advice-surgeries

Get in touch



@ukgolab



golab.bsg.ox.ac.uk



golab@bsg.ox.ac.uk