

# Commissioning for outcomes – a quick guide

Lesley Healey

To successfully commission for outcomes, you need to complete three things that, taken together, form the basis of successful outcome-based commissioning to ensure that you achieve the impact that you want:

- Select the desirable outcomes – decide which are the important outcomes, and how these can be measured (performance monitored)
- Understand how the system works – identify the links amongst and inter-dependencies of inputs, outputs and outcomes.
- Set a baseline against which improvement will be measured – a clear understanding of the starting point and the extent of the risks involved in successful achievement of outcomes

## Selecting outcomes, targets and measures

1. Outcome-based commissioning can be used to:

- Improve the welfare of individuals or
- Improve the operation of systems.

In either case, the performance targets should be expressed in terms of improved outcomes for human beings (individuals or communities).

Targets that are expressed in terms of systems (such as improved value for money) are outputs. However, outputs can make useful proxy measures where outcomes are difficult to ascertain and measure.

### *Personal improvement*

The focus is on the improvement of the individual and the primary objective is to ensure that each service user or each new cohort of service users achieves the desired outcomes.

### *System improvement*

The primary objective is to improve the quality of the system. Performance measures might be used (for example) as a diagnostic to target remedial action, eg, on under-performing staff or services. Because the ultimate goal is to improve outcomes for individuals, performance measures should be expressed in terms of improvements for individuals (or communities)

2. Outcomes can be about the wellbeing of either individuals or communities:

- *Individual outcomes* – the objectives are concerned with the wellbeing of that individual, eg improvements in the length and the quality of life.

- *Collective outcomes* – for services delivered to a community, outcomes need to be specified at a collective level, eg, through sampling public attitudes or feelings of wellbeing, or through some kind of surrogate measure based on outputs.

3. It can be difficult to decide how outcomes should be commissioned and measured:
  - Personal or collective?
  - If individual, should they be commissioned and measured by individuals or by an agency on their behalf?
  - If collective, at what level in society should they be commissioned, and how measured?
  - Should services be commissioned to provide both individual and collective outcomes?

The answers to these will significantly influence the outcomes for which you commission, and need to measure; and will impact on the eventual form of the service.

4. Where outcomes are commissioned for individuals but not by them, the agent needs to introduce co-production with the individual in order to deliver the right outcomes for that person. Outcome measures may be the same, but targets will be different for individuals.
5. Commissioning for outcome avoids traditional (and sometimes unproven) assumptions about the linkages between inputs, outputs and outcomes. However, to be effective as a management tool, outcomes must be:
  - sufficiently broad to allow for innovation in service delivery
  - sufficiently concrete to permit specification and measurement.

Some trade-off will always have to be made:

‘Sufficiently broad’

- Abstract and aspirational outcomes may win broad stakeholder support but they are notoriously difficult to quantify. One famous example aspirational outcome is ‘life, liberty and the pursuit of happiness’.
- Framing outcomes at a high level - so that successful delivery requires a wide range of services and/or central or local government programmes – is too complex in coordination and delivery terms to be achievable.

‘Sufficiently concrete’

- Very narrowly expressed outcomes, eg, that are dependent only on the results of one task, discourage exploration of innovation, different delivery or service integration.

6. In order to be commissioned, measured and rewarded, each outcome must be capable of independent observation and objective measurement (whether quantitatively or

qualitatively). Outcomes that are too abstract or aspirational may not be capable of measurement in any meaningful way.

Where there is a very strong relationship between a set of outputs and the outcome, proxy (or surrogate) measures of outputs can be used to determine a high level outcome. A simple and strong causal relationship may mean just one output measure is sufficient to determine whether an outcome is being achieved. However, it's usually necessary to measure several outputs in order to determine one outcome.

7. There are two approaches to measurement of any target:

- *Absolute* – targets are either achieved, or they are not (eg, someone is or is not re-admitted to hospital within 7 days of discharge, new employees are either retained for at least 13 weeks in their job, or they are not). Rewards are only offered for achievement.
- *Distance travelled* – progress is made towards a target (eg, increased gaps between offences, use of less harmful drugs). Rewards are offered on the basis of partial accomplishment of the eventual target. (See [www.outcomesstar.org.uk/](http://www.outcomesstar.org.uk/) for an example of this approach)

Which is used depends entirely on which works best, and the extent to which measurement is possible and useful in determining whether outcomes are being achieved

Absolute measures can also be used to measure (and reward) partial accomplishment:

- *Statistical measures* – eg, the percentage of people that achieve the target
  - *Comparative measures* – eg, relative performance across providers, or across different cohorts of service users (eg, by area)
8. Outcome measures (or output surrogates) can be expressed in positive (a good thing is being achieved) or negative (a bad thing isn't happening) terms. Negative measures can provide greater clarity, eg, 'not reoffending', but may be de-motivating for providers.
9. It is important to decide how you can motivate achievement (make it worthwhile). The measures themselves must reflect the motivators. Examples of motivators are:
- Financial
  - Reputational
  - Psychological

Motivators and measures must be closely aligned to prevent perverse or 'gaming' behaviours (including collusion across involved parties to achieve rewards) that lead to unintended consequences that derail achievement of outcomes.

Financial incentives are powerful but can be problematic unless:

- Financial rewards are closely aligned with targets
- Output measures (where used) are good surrogates for desired outcomes

- Rewards are proportionately higher for better achievement (ie, significant additional reward for significantly better outcomes)

Competition amongst multiple providers, measured through relative performance, is often effective although again measurement needs to be seen to be a fair representation of true achievement – providers must agree that the measures sufficiently reflect ‘good’.

10. To be a true outcome (good or bad):

a) A change in measures must be:

- Persistent
- Reliable
- Provable

b) You need to understand why that change happened

c) You must be able to replicate the change

*Simple and easy to understand measures are therefore best by far.*

Complex performance frameworks with many measures are counter-productive:

- They won't be used day-to-day by commissioners or providers
- Measurement and reporting are onerous and often disproportionate compared to the contract value
- Measures are less visible, so more open to ‘guesstimates’
- Performance is difficult and time-consuming to audit
- There are greater risks of perverse incentives
- It is far more difficult to compare providers (which of the measures are more important or significant??)

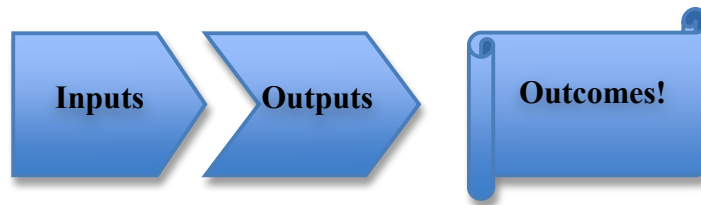
Lies, damned lies and statistics:

- If you use only statistics to judge success, you must always ask ‘why has the change occurred?’
- Apparent improvements understood only by statistics may merely be reflections of natural variations, so rewards may be given for no effective change
- Outcomes may apparently improve as a result of natural ‘wastage’ or change (eg, re-admission to hospital will reduce as people die – a not uncommon outcome for people requiring multiple admissions!). In statistics, this is called regression to the mean and simply indicates that there's no real change.

Short term projects (eg, pilots) may appear to have good results that aren't sustainable over the longer term because:

- Efforts are reduced once the service has bedded down
- Something specific made it work for one cohort that isn't replicable elsewhere (eg, high awareness, or incentives to achieve change, or the placebo effect)
- Change was the result of random or other natural variation, but the timescale was too short to display the true nature of the change

## Understanding the system



11. Outcome-based commissioning invites partnerships and providers to explore different ways of achieving the best outcomes, and establish the most efficient and effective approaches. To enable this, including through wise choices of measures, there must be some understanding of how inputs are connected to outputs, and outputs to outcomes. Efficiency is achieved where there is the ideal ratio between inputs and outcomes.

As commissioners, before you specify a service you must particularly have a fundamental understanding of the relationship between outputs and outcomes. Without this, you will not be able to determine whether and when success is being achieved, and won't select the right outcome measures (or a set of output measures as proxy for outcomes). Providers will need to develop a much more sophisticated understanding of how inputs lead to desired outcomes in order to both deliver targets and live within their budget.

In order to understand the system, commissioners should work with customers, communities, providers, and anybody else who can offer a real insight, or has an interest in the outcome (stakeholders).

12. Some systems are easy to understand, with just one, or even two or three, readily identified causal relationships that if corrected will deliver the desired outcome. Alas this is relatively unusual.

*“Some problems are so complex that you have to be highly intelligent and well informed just to be undecided about them”  
Laurence J Peter*

13. *‘Wicked’ problems*  
There are systems or issues where the relationships amongst inputs, outputs and outcomes are simply not understood sufficiently to enable control of outcomes. They have often been subject to intense research, theorising and speculation. Since we cannot (morally and legally) subject human beings to controlled studies where the outcomes for (or from) the control group are disastrous, effectively the ‘solution’ may never be known.

Problems that are rather better understood but still complex may respond to gradual control through establishment of contributory low-level outcomes or output measures that can be built upon over time.

14. *External influences*

It is important to understand the extent to which the system is under the control of the provider, or partners in general. This is particularly relevant to rewards and downstream actions where outcomes are being under-achieved, or going in the wrong direction. The economy, legislation, actions of government and private sector all impact the context within which a system is delivered. This works both ways: to merit a reward, *outcomes should be better than the average being achieved without additional inputs*. So even in hard times, outcomes may be better than expected, and the provider should then be rewarded. Comparing outcomes from different providers is another way of measuring relative achievements.

15. *Human behaviour*

Few systems are free of the vagaries of human behaviour but it isn't always significant in whether or not outcomes are achieved. In systems that require compliance for success, or where there is a personalised outcome (support to achieve independent living for example) the provider will need to work cooperatively (co-productively) with service users to enable achievement of outcomes and their success will be dependant on their skill in co-production. Where cooperation or compliance is not necessary to achieve an outcome (a clean street, for example) there will be different expectations of the provider.

16. *Environment*

Living environments that impose rules on inhabitants, and penalise breaches of these, provide controlled environments – extreme examples are prison, closed treatment units and locked wards. The application of a system in a controlled environment is less subject to human behaviour (service users have little choice) or external influences (the rules are part of the system). Short-term supported hostels, for example, offer a more controlled environment than does support in people's own homes. Providers working with similar service users but in different environments can expect to have different levels of outcome, and these variations need to be understood by commissioners when setting targets, or comparing providers.

## **Establishing baselines**

17. It is necessary to establish and clearly state the distance between outcomes required and the current position. This enables understanding of whether performance is improving and the extent to which targets are 'stretched', ie, the risk attached to achieving these. Commissioners must set out their expectations in terms of difference made, time periods and the population that will be served. Providers need to make a decision about whether to undertake a commission and the price they consider is reasonable for what is being required.

18. *The starting point*

Research is needed to establish the baseline against which improvement will be measured. The baseline must be capable of being expressed in the same outcome terms as the measures being used. This process also tests out the validity of targets and measures, and enables inclusion in the specification of the performance data that must be collected and reported as part of downstream procurement processes.

It is often also necessary to establish trends – how much change there has been over previous time – to enable both commissioner and provider understand the risk of successful outcomes.

#### 19. *Populations*

Where outcomes are about populations rather than individuals, how measurements of improvement will be made needs to be clear from the start. For example, amongst any population there may be low, medium and high risk individuals, or compliant to chaotic behaviours. Measurements must be consistently based on the same approach to mix of characteristics. Decisions must be taken on:

- The size of sample that is large enough to provide reliable statistical data, but small enough for data collection not to be unreasonably onerous
- Whether the population should be homogenous, or carry a mix of characteristics and therefore risks
- What the profile should be based on, and in what proportions
- Whether the monitored population will be self-selecting (eg, generalised surveys) or targeted

To make this work for populations about whom little is known (eg, where a new problem is identified, or a completely new service required) outcome-based commissioning should only be attempted where there is an appreciation of the need to negotiate targets as part of the downstream process.

#### 20. *Time spans*

The time allowed for outcomes to be achieved (or measurable improvements) requires a careful balance:

- Expectations of quick results may not allow time for the service to take over from a previous provider, settle in, get up to strength and bed in. Unreasonable expectations make the service too risky; providers will struggle and the service may be labelled as having failed when in fact it may have achieved well had there been sufficient time allowed for the service to become functional
- Allowing too much time runs the risk of drift and lack of control. It may also expose the service to extraneous variables without understanding the impact that these are having (including political risk)
- There must be sufficient time for outcomes to be achieved – high level or long-distance outcomes (for example, the impacts on an adult of having received a service as a child) need measures that express progress (distance travelled) or are outputs that indicate the likelihood of eventual outcomes