Ministry of Justice

Justice Data Lab: General annex to accompany re-offending reports

This document provides answers to frequently asked questions about Justice Data Lab (JDL) analyses and reports, to help customers and other interested parties to understand the background to the analyses and how to interpret the results.

Contents

What is the purpose of a JDL analysis? 2

How is re-offending measured? 2

How is the impact of an intervention analysed? 2

What caveats should be considered when reading a JDL report? 3

What measures of re-offending are used in the analysis? 4

How should the numbers in the report be interpreted? 5

What do the confidence intervals on the graphs mean? 5

What is the relationship between the measurements and the estimates? 6

Can the true impact of the intervention ever be outside the estimated range? 6

What are percentage points? 7

What is the meaning of statistical significance? 7

If the results of an analysis are non-significant, does this mean that the intervention has no impact? 7

How can a measure show a significant result when another measure shows a non-significant result for the same intervention? 8

Why does a JDL report contain more than one analysis? 8

Contact points 9

What is the purpose of a JDL analysis?

A JDL analysis adds to the evidence about the way in which an intervention with offenders affects their re-offending behaviour. This helps to determine whether the intervention reduces re-offending among its participants, and by how much. The analysis also examines the impact of the intervention on re-offending at different levels of severity.

The results of a JDL analysis can be used to estimate the impact that an intervention would have on people who are similar to those who received it. This does not mean that the intervention would have the same impact on all offenders.

How is re-offending measured?

In order to be included in a JDL analysis, a person must previously have committed a criminal offence that resulted in a court conviction, caution, reprimand or warning in England or Wales. This is called their 'index offence'. Re-offending is recorded from the 'index date', which is the date on which the person left custody or received a non-custodial sentence, caution, reprimand or warning as a result of the index offence. If there is more than one offence relating to the same index date, the most severe one is used as the index offence.

JDL analyses look at 'one-year proven re-offending', which records offences that the person committed during a one-year period starting on the index date and that resulted in a court conviction, caution, reprimand or warning in England or Wales during the same period or a further six-month waiting period.

If a person commits an offence more than one year after their index date, or if they are convicted of an offence more than eighteen months after their index date (or never convicted), this will not be recorded as a re-offence in a JDL analysis.

See pages 6-8 of www.gov.uk/government/uploads/system/uploads/attachment_data/file/472535/proven-reoffending-definitions-measurement-Oct15.pdf for a full description of the proven re-offending statistics used by the JDL.

How is the impact of an intervention analysed?

A JDL analysis provides *measurements* of the re-offending behaviour of a 'treatment group', which is a group of people who received an intervention around the time of their index date or during their index prison sentence, and whose identifying details are provided by the organisation running the intervention. It also provides measurements of the re-offending behaviour of a larger 'matched control group', which contains people who are similar to those in the treatment group but who have not received the intervention. The measured differences between the groups are used to make *estimates* of the impact of the intervention on the re-offending behaviour of the

treatment group. By extension, the estimates also provide evidence of the impact that the intervention would have on the re-offending behaviour of any people who are similar to those in the treatment and control groups.

The control group is expected to exhibit the re-offending behaviour that the treatment group would have exhibited if they had not received the intervention. The members of the control group are therefore chosen for their similarity to the members of the treatment group on a wide range of characteristics that are known to be generally related to offending behaviour, including demographics, employment history, criminal history and individual risks and needs. They are selected from a pool of almost all the offender records in England and Wales that have index dates in the same years as those of the treatment group.

What caveats should be considered when reading a JDL report?

When interpreting the evidence presented in a JDL report, the following caveats should be considered:

- The differences between the re-offending behaviour of the treatment and control groups could be due to unobserved factors as well as to the impact of the intervention. This is because the control group must be selected to match the people in the treatment group, and the matching can only take account of characteristics that have been observed in the data sets used. The matching process uses individual information about offenders in England and Wales, which is drawn from reliable administrative data sets. This covers a wide range of characteristics that are related to offending behaviour, including demographics, employment history, criminal history and individual risks and needs. However, it is still possible that the re-offending behaviour of the treatment and control groups will differ due to unobserved characteristics such as the impact of other interventions, motivation to change offending behaviour or the complexity of personal problems.
- Reliable results depend on good matching between the treatment and control groups. 'Standardised differences' are provided to indicate the quality of matching in an analysis, and are individually rated as 'good' (-5% to 5%), 'reasonable' (-10% to -5% or 5% to 10%) or 'poor' (below -10% or above 10%). Each standardised difference is a measure of the difference between the treatment and control group averages for one characteristic. The more standardised differences are rated as good, the more reliable the results of the analysis are. JDL analyses aim to achieve the best possible matching with the data available.
- It may not be possible for the treatment group to include everyone who has received the intervention, and the impact of the intervention on those who are included may be different to the impact on those who are not. The impact of the intervention could also be quite different on an entirely different group of people. For this reason, the re-offending behaviour measured

in an analysis should not be directly compared either to the re-offending behaviour measured in any other analysis or to figures such as national averages.

- The impact of the intervention may differ for each person who receives it. A JDL analysis can only estimate the impact of the intervention on the re-offending behaviour of the treatment group as a whole.
- Some re-offences committed during a one-year period are not recorded in JDL analyses. Only proven re-offending is recorded, which is generally an underestimate of actual re-offending. In addition, a re-offence is not included if it is proven more than six months after the end of a person's one-year re-offending period. This means that a re-offence is more likely to be included if it is less severe, because the time between offence and conviction is generally shorter for less severe offences. Unrecorded offences affect the measured re-offending rates of both the treatment and control groups.

What measures of re-offending are used in the analysis?

In a JDL report, measurements and estimates are presented for up to fourteen measures of one-year proven re-offending. Three are headline measures of overall re-offending, nine are measures of re-offending severity and two are measures of custodial sentencing. They are:

- *Re-offending rate* – the number of people who commit a proven re-offence, expressed as a percentage of the group
- *Re-offending frequency* – the number of proven re-offences committed, expressed per person
- *Average time to first re-offence* – the average number of days between a person's index date and the date on which they commit their first proven re-offence, including only those who re-offend
- *Tier 1 re-offending rate* – the number of people who commit their first proven re-offence in severity tier 1, expressed as a percentage of the re-offenders
- *Tier 2 re-offending rate* – the number of people who commit their first proven re-offence in severity tier 2, expressed as a percentage of the re-offenders
- *Tier 3 re-offending rate* – the number of people who commit their first proven re-offence in severity tier 3, expressed as a percentage of the re-offenders
- *Tier 1 re-offending frequency* – the number of proven tier 1 re-offences committed, expressed per re-offender
- *Tier 2 re-offending frequency* – the number of proven tier 2 re-offences committed, expressed per re-offender

- *Tier 3 re-offending frequency* – the number of proven tier 3 re-offences committed, expressed per re-offender
- *More severe re-offending rate* – the number of people who commit their first proven re-offence in a more severe tier than their index offence, expressed as a percentage of the re-offenders
- *Same-severity re-offending rate* – the number of people who commit their first proven re-offence in the same severity tier as their index offence, expressed as a percentage of the re-offenders
- *Less severe re-offending rate* – the number of people who commit their first proven re-offence in a less severe tier than their index offence, expressed as a percentage of the re-offenders
- *Custody rate* – the number of people who receive a custodial sentence for their first proven re-offence, expressed as a percentage of the re-offenders
- *Custody frequency* – the number of custodial sentences received as a result of re-offending, expressed per re-offender

The three overall measures are included in every report, and the others are included if there are enough people in each category to allow reliable estimates to be made. Offences are classified into three tiers of severity, with tier 1 being the most severe. See pages 22-26 of www.gov.uk/government/uploads/system/uploads/attachment_data/file/472535/proven-reoffending-definitions-measurement-Oct15.pdf for a list of offence types in tiers 1 and 2 (serious acquisitive offences are tier 2).

How should the numbers in the report be interpreted?

Summary boxes and graphs in JDL reports express some measures per 100 people instead of per person or as a percentage, in order to make the numbers more meaningful to the reader. For example, a re-offending rate of 40% and a re-offending frequency of 1.2 offences per person in the treatment group may be written as: "For 100 typical people in the treatment group, 40 people committed a proven re-offence within a one-year period. They committed 120 proven re-offences during the year."

See www.justice.gov.uk/downloads/justice-data-lab/justice-data-lab-methodology.pdf and https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/506327/methodology-review-response.pdf for a full description of JDL methodology.

What do the confidence intervals on the graphs mean?

A confidence interval shows the estimated range of a re-offending measure for one group. For example, the re-offending rate of a treatment group could be 40% and the confidence interval could cover the range from 35% to 45%. This would mean that the treatment group had a re-

offending rate of 40%, but that similar people who received the intervention would be expected to have a re-offending rate that is between 35% and 45%.

What i**s the relationship between the measurements and the estimates?**

The measurements give re-offending information for the treatment and control groups used in the analysis. The re-offending behaviour that is measured depends on the impact of the intervention, but it also depends on random factors affecting the particular people in those groups.

The estimates quantify the impact of the intervention, without the random factors. But, because they are based on the measurements, the estimates cannot be exact. For this reason, each estimate is given as a range of numbers, and the size of the impact is expected to be somewhere within this range.

For example: the measurement of the difference in the re-offending frequency could be -1.5 offences per person, with the impact estimated to be between between -2.0 and -1.0 offences per person. This would mean that the treatment group committed an average of 1.5 fewer re-offences per person than the control group, and that the intervention would be expected to prevent between 1.0 and 2.0 re-offences per person if it were provided to similar people in future.

In a JDL analysis, an estimated range is always centred on the measurement that it is based upon. In general, the range becomes narrower as the treatment group becomes larger. This means that the impact of an intervention can be estimated more precisely when the treatment group is large.

Can the true impact of the intervention ever be outside the estimated range?

Yes. The estimates are calculated using a standard method called '95% confidence'. This means that the underlying difference between the re-offending behaviour of the treatment and control groups is expected to be outside the estimated range in 5% of cases: in 2.5% of cases the underlying difference will be lower than estimated, and in 2.5% of cases it will be higher. It does not mean that there is a 95% probability of the underlying difference being within the estimated range in every case, because each case should be judged individually using all the available evidence.

It is also possible that the underlying difference between the re-offending behaviour of the treatment and control groups may be caused by unobserved differences between the groups as well as by the impact of the intervention.

What are percentage points?

Percentage points (pp) are used to express the difference between two percentages. For example, the difference between 5% and 10% is 5 percentage points, and the difference between 85% and 90% is also 5 percentage points.

What is the meaning of statistical significance?

A statistically significant result means that random factors appear to be an unlikely explanation for the measured difference between the treatment and control groups. It is robust evidence that there is an underlying difference between the re-offending behaviour of the groups. For example, the estimated range for the difference in the re-offending rate could be -5 percentage points to +3 percentage points. This is not statistically significant because it appears plausible that the difference is zero. But if the estimated range is -10 percentage points to -2 percentage points, this is a statistically significant result because it appears unlikely that the difference is zero.

Statistical significance is a guideline. It acts as a flag to highlight the most convincing pieces of evidence. See http://www.thinknpc.org/publications/understanding-statistical-significance/npc_understanding-statistical-significance_final/ for more guidance on understanding statistical significance.

If the results of an analysis are non-significant, does this mean that the intervention has no impact?

No. A non-significant result means that it is plausible that the intervention has a positive impact, a negative impact or no impact on re-offending, based on the JDL analysis alone. For example, the estimated range for the difference in the re-offending rate could be -5 percentage points to +3 percentage points. It would then be plausible that the impact was -5 percentage points, or +3 percentage points, or anything in between. The only way to get robust evidence of the direction of the impact would be to include more people in the analysis.

A non-significant result provides useful information. In the example above, it appears unlikely that the impact is greater than -5 percentage points in the negative direction or greater than +3 percentage points in the positive direction.

Every measurement provides evidence about an intervention, even if the result is non-significant. The measurement is at the centre of the estimated range. In the example above, the measured difference in the re-offending rate would be -1 percentage point. This means that the most plausible conclusion would be that the intervention reduces the re-offending rate by 1 percentage point. It is a promising result, but the lack of statistical significance means that the conclusion would not be very robust.

Results from a JDL analysis should always be placed into the wider context of offender rehabilitation. Each result provides a piece of evidence about the effectiveness of an intervention. Other important pieces of evidence include the methodology of the intervention and accounts of its impact on specific individuals. Effectiveness can also be demonstrated through outcomes other than re-offending behaviour, such as health or relationships, which may lead to a reduction in re-offending after a period of more than one year. The overall picture of an intervention is made up of all these pieces. A JDL report is designed to provide reliable, quantitative evidence with which to judge the impact of the intervention in the specific area of re-offending behaviour over a one-year period.

How can a measure show a significant result when another measure shows a non-significant result for the same intervention?

A statistically significant result occurs when the direction of an impact becomes apparent from its estimated range. Each measure records a different aspect of re-offending behaviour, and the impact of the intervention on each aspect can be different. This means that the impact on one measure may be large enough for its direction to become apparent, while the impact on another measure in the same analysis may not be large enough. It does not necessarily mean that the impacts act in opposite directions.

For example, the estimated impact of an intervention on the re-offending rate could be between -10 percentage points and +1 percentage point (non-significant), while the estimated impact on the re-offending frequency could be between -1.5 and -0.5 offences per person (significant). It appears likely that the intervention reduces both the rate and frequency of re-offending, but there is a small chance that it reduces the frequency without reducing the rate.

Why does a JDL report contain more than one analysis?

Each JDL report contains a 'complex' and a 'standard' analysis. It also contains 'national' and 'regional' analyses if the intervention takes place in a specific geographical area. All of these look at almost the same treatment group, but each has a different matched control group. Multiple analyses broaden the comparison between people who have received an intervention and people who have not, providing more detailed and robust evidence of the impact of the intervention.

Standard analyses use characteristics such as demographics, employment history and criminal history when matching the control group to the treatment group. Complex analyses use risks and needs as well as the same characteristics as the standard analyses. The additional information comes from the Offender Assessment System (OASys), which records data from individual offenders on factors such as accommodation, education and work skills, substance misuse, mental health, relationships and attitudes towards offending. These factors can be important in matching a control group that is similar to the treatment group in terms of offending behaviour,

but they are not usually available for everyone in the treatment and control groups, unlike the standard characteristics. For this reason, both standard and complex analyses are presented. See https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/491688/oasys-methodology.pdf for a full description of the integration of OASys data into JDL analyses.

National analyses use a control group of people who have been convicted of offences anywhere in England or Wales. If an intervention takes place in a specific geographical area, a regional analysis is also conducted using a control group of people who have been convicted in the same area. The national analysis selects the control group from the largest possible pool of people, while the regional analysis avoids any unobserved differences in offending behaviour between people in different areas.

Contact points

Press enquiries should be directed to the Ministry of Justice press office:

Tel: 020 3334 3555

Other enquiries about the analysis should be directed to:

Sarah French
Justice Data Lab Team
Justice Statistical Analytical Services
Ministry of Justice
7th Floor
102 Petty France
London
SW1H 9AJ

Tel: 07967 592428

E-mail: justice.datalab@justice.gsi.gov.uk

General enquiries about the statistical work of the Ministry of Justice can be e-mailed to: statistics.enquiries@justice.gsi.gov.uk
General information about the official statistics system of the United Kingdom is available from www.statisticsauthority.gov.uk/about-the-authority/uk-statistical-system