



Ministry  
of Justice

## **Justice Data Lab**

A Peer Review of existing methodology – Response

February 2016

## **Contents**

1. Summary .....	3
2. Introduction .....	4
3. Methodology additions to date .....	4
3.1 New measures.....	4
3.2 Identifying the correct sentence.....	5
3.3 Limitations to the control group.....	5
3.4 Adjustments for interventions for those with probation sentences .....	5
3.5 Sentence matching for mixed cohorts.....	6
4. Peer Review Questions .....	6
5. JDL Conclusion and Forward Look .....	14

## 1. Summary

Whilst there are no inherent problems with the process used by the Justice Data Lab (JDL), it has been several years since the launch of the service and various developments have been implemented during this time. The JDL team always seek to improve the service offered following the initial set up and so a peer review of the existing methodology was commissioned to identify key areas to investigate further to enhance the JDL analyses.

The key messages and conclusions of the review into the current methodology employed by the Justice Data Lab (JDL) are detailed below. The items in bold are the areas identified for investigation, which are discussed in detail in the report and in the 'Conclusion and forward look' (section 5). A programme of work will be implemented to look over these aspects over the next 12 months.

- It was generally deemed that the approach overall was suitable, however several areas of investigation were identified to assess their impact on the JDL process.
- These areas include comparing **alternative statistical techniques**, used elsewhere both in Government and in academia when making comparisons, against propensity score matching with caliper (section 4.1) and assessing the **impact of refining the t-test models** against the clarity of more complex tests that are difficult to convey to a non-technical audience for bespoke comparisons (section 4.10).
- Also, a suggestion to assess how the **JDL model can be refined** (section 4.6) and to **test the overall JDL approach** using dummy data (section 4.11) has also been taken on, with the aim to confirm views from the reviewers that the JDL process is suitable for its purpose.
- Looking at exploring the potential for **testing for unconfoundedness** (section 4.11) as part of extra sensitivity analyses to provide a guide of the impact a potential important unobserved variable may have on the results;
- One key concern was that the variables used in the creation of the JDL control groups would lend itself to having an over-fitted model (see 4.6). The nature of the small treatment groups assessed by the JDL means there is little that can be done about this aspect, with the quality of match being the key outcome to establish reliability of the comparison.
- The current practice of **deleting treatment group data after publication will be reconsidered** to assess whether meta-analysis could be taken forward in the future to provide an over-arching view on what works in reducing reoffending. This would be assessed alongside whether it would affect submissions to the JDL service, with this being the priority (section 4.13).

- Key updates to the process since the original methodology was published following the launch in 2013 have been described so that a full picture of how the JDL process works is fully transparent and publicly available.

## **2. Introduction**

The JDL was launched in 2013<sup>1</sup> and aims to improve the evidence base on successful rehabilitation by enabling organisations to better assess the impact of their work on re-offending, using aggregate re-offending data provided by the JDL service.

The methodology currently being used within the JDL team is based on that published in the original methodology paper<sup>2</sup> and has developed both in terms of the measures produced and the nuances handled within the Data Lab processes.

As such, a peer review of the methodology with focused, specific questions being asked of the reviewers was conducted during autumn 2015. A range of interested parties were identified and approached (from academia, the Justice Sector and cross-Government), with suitable replacements when some of the original review panel were unable to commit to the project. The review aimed to make sure that the methodology in place is fit for purpose, as robust as possible and to identify any areas for improvement.

This paper covers the key developments of the JDL methodology and process since the publication of the original methodology as well as covering the questions asked of the review panel, the responses returned and how the JDL team plan to take forward developments to address any concerns.

## **3. Methodology additions to date**

### **3.1 New measures**

The main focus of the JDL is to assess whether data held by the Ministry of Justice suggests that a change in re-offending behaviour has occurred due to a participating organisation's offender rehabilitation work. As such, the headline measure included from the outset of the JDL work was the one year proven re-offending rate<sup>3</sup> for the cohort of offenders that providers worked with, known as the treatment group, compared against that of a matched control group.

---

<sup>1</sup> With a 2 year pilot phase, becoming a permanent service in April 2015.

<sup>2</sup> [www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/392929/justice-data-lab-methodology.pdf](http://www.gov.uk/government/uploads/system/uploads/attachment_data/file/392929/justice-data-lab-methodology.pdf)

<sup>3</sup> The **one year proven re-offending rate** is defined as the proportion of offenders in a cohort who commit an offence in a one year follow-up period which received a court conviction, caution, reprimand or warning during the one year follow-up or in a further six month waiting period. The one year follow-up period begins when offenders leave custody or start their sentence in the community.

However, to develop the JDL reports to make them as useful as possible for customers, two more measures were developed to improve the information provided to customers and to align with headline re-offending measures published by the Ministry of Justice (MoJ). These are:

- Frequency of one year proven re-offending<sup>4</sup>
- Time to re-offending<sup>5</sup>

Developing from this new baseline set of measures, in October 2015 a further set of measures was published in the October 2015 JDL analysis for Safe Ground<sup>6</sup>, assessing the severity of re-offences and whether re-offences resulted in custodial sentences. These measures are not standard but are provided where applicable.

### **3.2 Identifying the correct sentence**

It is important that JDL requests contain only those who have actually offended (rather than deemed at risk of offending) and those whom the organisation worked with close to their index date (start of probation sentence/release from prison). This is because re-offence information covers only those who have been convicted or cautioned for an offence, and the one year re-offending period starts from the index date (more on this in section 4.13). If an organisation provides details of everyone they have worked with irrelevant of a) whether the people in question have had a proven offence or simply been at risk of offending or b) not taking into account the time-lag between the start of probation sentence/release from prison and the start of the intervention then this can lead to a lower match rate to MoJ's centrally held datasets.

### **3.3 Limitations to the control group**

Whilst this is not an improvement or development of the current processes, this section is to clarify a potential gap in the methodology note where it is not currently specified who is allowed into the control group. The group that goes through the matching stage reflects the observable characteristics of the treatment group – i.e. any limitations to gender or age are taken into account, likewise the severity of the index offence and criminal history is echoed in the control group. This is to make sure that anyone who would be ineligible for the intervention in question is excluded from being part of the control group.

### **3.4 Adjustments for interventions for those with probation sentences**

When the JDL is processing a request for an intervention for those on, for example, community or suspended sentence orders, some adjustments need to be made. As there will be a time delay from the start of the sentence to the start of the intervention, anyone

---

<sup>4</sup> The average number of proven re-offences per individual in the cohort that meet the same definition as the headline measure.

<sup>5</sup> The average number of days between the index date (release date from custody or start of probation date) and the offence date of the first re-offence within the one year follow-up period described in the headline measure. This measure is only calculated for individuals who re-offend

<sup>6</sup> [www.gov.uk/government/statistics/justice-data-lab-statistics-october-2015](http://www.gov.uk/government/statistics/justice-data-lab-statistics-october-2015)

who has re-offended during this period is excluded from the treatment group. This is to make sure that such re-offences are not attributed to the programme in question.

Similarly, additional checks are imposed on the control group so that the matched individuals have similar characteristics. These are to establish that all members of the matched control group could not have committed a proven re-offence during the period that is excluded for their matched treatment group counterparts. Any matches where the control group had committed a proven re-offence during the exclusion period of the counterpart are removed from the analysis.

### **3.5 Sentence matching for mixed cohorts**

Additional matching is also taken into account if a mixed cohort of individuals with both prison and community based sentences are being analysed. This is to make sure that those with custodial sentences are matched to those with custodial sentences, and those with community based sentences are matched to those with community based sentences (note this is not broken down further by type of probation sentences, but the matching quality between the different types are checked when looking at the standardised differences). When numbers are large enough, the JDL team look to provide separate analyses for those with custodial and community based sentences to provide further information about the cohort. These checks make sure that we have greater confidence that the matched control group presents a more accurate counterfactual for comparison.

## **4. Peer Review Questions**

### **4.1 Is one-to-many radius propensity score matching (PSM) (with caliper and with replacement) the most suitable approach for assessing the impact of a treatment/intervention or are there others that we should be considering as well as/instead of?**

The general consensus was that this method was suitable for the task and no radical improvements were suggested. Kernel-based PSM as well as matching without replacement were referenced to as alternatives to the established JDL method, with stronger views being shown about the use of kernel matching than one-to-one matching without replacement.

Separate to the methodology review, coarsened exact matching and regression discontinuity designs have been noted as potential replacement for radius PSM.

#### **The JDL's response:**

Where we have data available (limited due to data deletion after a JDL analysis), the team will seek permission to use the information to test out the methods suggested to produce sensitivity analyses on how results (may) differ under each option. Once such tests have been completed, a recommendation will be made and the JDL matching method will be adjusted if necessary.

**4.2 It is often the case that the JDL control groups are very large and are often being compared to small treatment groups. What is the best way of dealing with such comparisons?**

This did not appear to be a concern of the review panel, as long as the balance between matching members of the treatment group and the quality of the match overall is maintained. Adjusting the one-to-many match rule to one-to-one was suggested as an option.

**The JDL's response:**

As a number of the JDL requests deal with small cohorts (<100), restricting the match to one-to-one may lead to more inconclusive results than already found due to the larger confidence intervals around the control group's outputs. As part of testing the different matching methods previously, how the size of the control group changes will be observed.

**4.3 Currently a caliper of 0.1 is used for matching in the first instance, switching to either 0.01 or 0.2 afterwards if matching needs augmenting. Are these the most appropriate calipers or should a different approach be adopted?**

The reviewers either did not respond to this question or they confirmed this was a logical procedure (albeit with one response stating to reproduce it with a bandwidth for kernel matching rather than the caliper method).

**The JDL's response:**

The current sequential procedure is satisfactory and no changes will be made to the current method (although some changes will have to be made if there is a switch to kernel matching, for example).

**4.4 Within the same intervention evaluation, should the same caliper be used across all analyses for consistency where matching quality allows or would it be fine to have a mix of calipers?**

In some cases, this question was misinterpreted and taken to mean that we would set the same caliper for all interventions. This was argued against, stating that a caliper should not be forced upon a treatment group where the matching quality is not sufficient.

The question was whether the **same** calliper should be used for different analyses within a **single** report (for example, when a report has both national and regional control groups). Where this was understood, the reviewer(s) stated that such analyses should not require the same caliper.

### **The JDL's response:**

Whilst fixing the caliper in advance may aid transparency, the matching quality needs to be met and, as such, the caliper should alter to account for this. As such, the focus will remain on the matching quality of the treatment and control groups rather than aim for seemingly consistent analyses using the same caliper throughout a report.

#### **4.5 We perform 'sentence matching' so that, for an intervention that works with both prisoners and those on community based sentences, these overarching sentence types are matched specifically. Should this approach be adopted more widely for other observed variables or do the propensity scores cover this as reasonably possible?**

The general response was to make sure that the groups are as well as matched possible and that the described 'sentence matching' may offer scope to cover unobserved variables. It is noted that to do such hard matching on too many variables could lead to having few matched offenders per treatment group member and possibly lead to some of the treatment group not being matched at all. For estimating overall impacts, propensity scores should cover this in themselves. One point made in favour of more hard matching was for external approval, as it may improve the credibility of results if people are reassured that participants are not being matched against completely different non-participants.

### **The JDL's response:**

We aim to make our methods as transparent and clear as possible, so that users and other interested parties are assured of our processes. Adding in more hard matching would go some way to reducing the size of the control group (which may be viewed positively by some, when it is compared to the size of the treatment group), it may actually lead to more people being unmatched and dropping out of the treatment group entirely. As we have attrition rates when matching to our administrative datasets already, we aim to keep as many of the treatment group throughout our analyses. Where there are sufficient numbers of those with custodial and community based sentences, we offer a split analysis as well as looking at the group overall to assess whether the impact is greater for one group over the other.

#### **4.6 Currently, the logistic regression models used to establish the propensity scores come from a backward elimination stepwise procedure based on p values (footnotes 4 and 5 on page 16 of the published methodology note). Is this appropriate?**

The panel members found no issue with the backward elimination stepwise procedure based on p values and felt it was appropriate. A key concern from one reviewer was that the model generally would be over-fitted and therefore could be deemed as being unreliable.



### **The JDL's response:**

Over-fitting occurs when a statistical model describes random error or noise instead of the underlying relationship. Over-fitting generally occurs when a model has too many parameters relative to the number of observations. A model that has been over-fit may have poor predictive performance, as it can exaggerate minor fluctuations in the data. This will affect those results based on small treatment groups in particular, and many of those who use the JDL service do provide small groups. It should be noted, however, that whilst some consider model parsimony (which is related to over-fitting) is an important issue, much of the academic community do not and instead focus on matching diagnostics, as does the JDL.

It is fair to point out that some variables within the JDL models may be correlated with each other, although much of the academic community do not consider this an important issue for PSM model. In addition, the programming used would break if any are particularly correlated. As such, the JDL will look into further sensitivity analyses to see how/if the overall model that all JDL analyses begin with can be refined. This overall model is always reduced as variables that are not significant in both the propensity of treatment and of reoffending are removed.

### **4.7 Could the approach taken in estimating the propensity scores using logistic regression be improved? If so, how?**

One comment suggested considering introducing second-order terms of how some of the variables interact with each other. Another put forward a strong case that the JDL must improve this stage, with the single-stage model for the propensity scores looking at the probability of treatment is modelled but is used to control for reoffending in comparing the recidivism rates of the treatment and control groups.

### **The JDL's response:**

Whilst the response to 4.6 states that the JDL focuses on matching diagnostics to assess issues caused by over-fitting the model, introducing interaction terms is likely to aggravate this problem. The single-stage model looks at the probability of treatment and does not include the probability of reoffending as the reoffending rates are the output and should not be included in the modelling of the scores (however they are used in appraising what variables are significant and should be included in the final model).

### **4.8 Should other matching quality diagnostics apart from standardised mean differences be used?**

Performing significance testing on the differences were suggested as, whilst it is generally accepted that the match is adequate if mean differences are sufficiently small, confidence in the results would increase if there was no overall significance in the difference between

treated and comparison groups. It was noted that these standardised mean differences are the most important diagnostics but that it would be good practice to assess the distribution of the propensity scores of both the treatment and control groups to identify outliers.

One point was that such diagnostics focus on calibration and whether the two groups have similar estimated risks but not about discrimination, i.e whether the model is good at identifying those who are treated.

#### **The JDL's response:**

Whilst not explicitly stated in the methodology paper, there are checks on the distributions of the propensity scores to look for any issues, although there is low risk of this when using a tight caliper. The JDL team looks to include all important variables into the JDL models to address both calibration and discrimination aspects, however if you increase the discrimination between the two groups then it would no longer be suitable to compare the treatment and control groups if too dissimilar. For example, the team have taken on a significant piece of development work to incorporate the Offender Assessment System (OASys) data into the JDL analyses<sup>7</sup> to be able to account for specific needs and issues of offenders. The JDL team are committed to expanding the data incorporated into the service to make sure that the information used in analyses is as useful as possible.

#### **4.9 Is the most appropriate significance testing procedure carried out post-matching for the three main measures? For example, when the treatment group is small and the frequency rate is not normally distributed.**

Increasing the minimum treatment group size from 30 to 40 was recommended by one panel member. Another suggested considering permutation testing (a method of re-sampling in which the distribution of the test statistic is obtained by calculating all possible values of the test statistic under rearrangements of the labels on the observed data points). The threshold of 0.05 was questioned but nothing further on this question was suggested by the reviewer. Alternatives of a doubly-robust procedure, where you first implement PSM to get matched treatment and control groups and then use regression analysis to estimate the impact, or implementing bias correction to the differences (i.e. make an adjustment to the mean difference between the two groups that accounts for any remaining differences in covariates) were proposed.

#### **The JDL's response:**

There was no evidence provided to support increasing the treatment group size from 30 to 40 and this would restrict the use of the JDL service to small organisations even more than currently. There are issues with using regression analysis on the matched treatment and control groups in that you won't be able to control for many variables before the model is

---

<sup>7</sup> [www.gov.uk/government/statistics/justice-data-lab-statistics-january-2016](http://www.gov.uk/government/statistics/justice-data-lab-statistics-january-2016)

over-fitted (see previous question), and the impact after controlling for other variables is more difficult to explain and present. This is a key aim for the JDL due to the generally non-technical audience of its' users. However it may be suitable to consider this approach when the JDL cannot get a good match on key important variables.

Permutation testing is a suitable option to look at, however it would offer little benefit to the current processes and would be resource intensive. Bias correction looks to improve the quality of the matching by running regression on both the matched treatment and control groups. With the treatment groups submitted to the JDL generally being small, it is likely that not many covariates could be included in the regression model.

**4.10 When establishing the range of difference between the re-offending rates of the treatment and control groups, we currently compare the extremes of the two 1-sample t-test confidence intervals. Should a 2-sample t-test or other approach be adopted if t-testing were to remain in practice?**

Several panel members agreed that the method had the potential to be incorrect. Cramér's V (a measure of association between two nominal outputs) was proposed by one member and another agreed that a two-sample t-test should replace the current procedure.

**The JDL's response:**

It should be noted that the indication of a result being statistically significant has always been calculated correctly. However, the range of the difference used the two 1-sample t-tests - this was a measured assessment upon the establishment of the JDL procedure, in order to balance both being statistically accurate and also being understandable by a lay-audience (to be able to infer the range of difference from the chart presenting the confidence intervals).

However, since the review was commissioned, this issue was looked into. It was found that it would be highly unlikely that the range would visibly differ when the control group is substantially larger than the treatment group (which is standard in the JDL process) and so the switch was implemented in October 2015. This would not have changed any headline measures published to date. When the two groups are more balanced in size, the range of difference stated between the re-offending rates being compared may differ to the confidence intervals in the relevant report. Thorough commentary would be provided to assist readers' comprehension in such analyses.

**4.11 What extra sensitivity analyses would be appropriate to provide a guide as to the impact a potential important unobserved variable may have on the results?**

There are a number of calculations that one can do to calculate how volatile the estimated outputs are to the presence of a key unobserved variable. These do not provide conclusive evidence but indicate how much weight you may want to place on any estimate of a

particular treatment. It was deemed standard by one to test for unconfoundedness by estimating placebo impacts (comparing the mean difference between the two matched groups in an outcome which cannot plausibly have been affected by the intervention).

One response suggested additional analyses that would test the overall JDL approach by estimating the impact of dummy interventions, i.e. choosing a random group to be considered at the treatment group and complete the JDL process, running this at least 100 times and look for the confidence intervals to span zero in 95% of the cases for confirmation that the procedure is suitable.

#### **The JDL's response:**

Testing for unconfoundedness seems a sensible approach to investigate further, subject to the availability of suitable data. Testing the overall approach as described above also seems suitable to look into.

#### **4.12 Presently the analyses include using both national and regional control groups (where regional control groups are appropriate). Is this the best approach to deal with the issue of area bias and can/should anything more be done?**

The panel appear to be in agreement that producing a separate regional control group should deal with area bias where suitable, with one noting that location should be controlled for as much as possible. It was suggested that looking at LDU (local delivery unit) level information would cover this point.

#### **The JDL's response:**

Regional control groups where possible will continue. Either requesting LDU information from customers or incorporating further datasets to be able to cover this aspect internally would place a burden that does not appear to be warranted. Please note that the JDL also look to perform prison-level or prison category type-level analyses where suitable to help account for any unobserved issues related to the establishment in question.

#### **4.13 Are there any other points you wish to make/suggest for our consideration to enhance the Justice Data Lab methodology?**

Several responses comment that the JDL approach seems highly automated, standardised and non-personal. Whilst the Data Upload Template (used for an organisation to submit their information) is standard, the JDL team work closely with each organisation from the outset of the analysis for their request to make sure that the intervention process is fully understood and to clarify any intricacies that could help inform the best way to take forward the analyses. We make particular aspects of an intervention, such as how offenders are recruited to an intervention, clear in the report and account for anything relevant within the caveats section (such as any possible selection bias when offenders self-refer themselves for a programme).

Relating to the reoffending data, it was noted that the 6 month waiting period (added to the one year follow-up period to allow for offences which are committed towards the end of the follow-up period to be proven by a court, resulting in a conviction, caution, reprimand or final warning) is not necessarily long enough to capture serious offences that could take years to be proven (e.g. serious sex offence, murder). Whilst some cases may not be covered, this is keeping in line with proven re-offence National Statistics publications. Also worth noting is that, whilst it is 6 months for those who commit an offence at the end of the 12 month follow-up period, it is 18 months for those who commit an offence at the beginning. Also, the ideal would be to have reoffending data for each analysis so that the one year reoffending rate is measured from the start of the intervention, rather than the index date of each person in the treatment group. However, this would require significant resource to create such bespoke datasets which is not possible within the JDL service, and it is made clear that the rates are from the index dates. For programmes aimed at those with community based sentences, and therefore likely to see a delay between the index and intervention start dates, a maximum lag of 6 months is allowed in order to retain as large a treatment group as possible without incurring too much bias into the output measures.

It was commented that when a programme demonstrates statistically significant outcomes, it does not necessarily mean that the impact is substantive and meaningful from a behavioural change perspective. The size of the impact is clearly stated throughout the JDL report and the JDL analysis is to act as one piece of evidence, rather than giving the full picture of the impact.

A query was made about whether it was possible to restrict the types of offence considered in the outcome timeframe (for example, for an 'anger management' treatment aimed at domestic abuse offenders you could restrict the relevant outcomes to violent offences). This is worth considering and thoughts on this option will be put forward to the JDL Expert Panel (i.e. JDL steering group) for views. However, this would require bespoke re-offence data that focused on specific re-offences and is subject to the same resource limitations as previously discussed in this section.

One reviewer pointed out that the process assumes we know for certain what the actual propensity scores are and does not account for this estimation, leading to a recommendation that we estimate the variance of the matched outputs. As commented in some of the previous responses, the JDL service aims to strike a balance between technical accuracy and reader comprehension.

This balance is particularly key for JDL reports where the analyses produced are quite technical in their nature. A review of the JDL publication will take such a balance into account and investigations mentioned in previous responses will also focus on this aspect.

Meta-analysis (combining data from multiple studies to obtain a larger sample size than the individuals parts to check that favourable results can be achieved consistently as opposed to

being 'one offs') was proposed. This would be a useful way forward in understanding what really works in reducing reoffending, that would help for the MoJ in its policy making protocols and inform front line organisations to adapt their programmes if/when suitable. However, a key agreement between the JDL and the organisations that use the service is that, once the analysis is published, the data provided is deleted. This was established to address a key concern of the voluntary and community sector. To improve the usefulness of the JDL, the team are looking into this agreement to see what can be done without affecting the trust that users have in the service.

## **5. JDL Conclusion and Forward Look**

The methodology review has proven to be an informative exercise that acknowledges the work put into the establishment of the JDL in 2013 but highlights key areas where improvements can be made. As previously mentioned, the need to keep the reports clear to users is key and any developments will have this aspect in mind.

As with all analysis, the Justice Data Lab seeks to continually improve the service provided and will be looking over the following aspects for investigation over the next 12 months:

- Testing the impact on results using the alternatives discussed in section 4.1 to one-to-many radius matching PSM with caliper;
- Assessing the variables used in the JDL models to see how/if the overall model that all JDL analyses begin with can be refined (see section 4.6);
- Testing the outputs of future bespoke comparisons and assess the impact using the two-sample t-test on the range of difference between the reoffending rates (section 4.10);
- Exploring the potential for testing for unconfoundedness (section 4.11);
- Testing the overall JDL approach by estimating the impact of dummy interventions (section 4.11);
- Considering options for data retention that would allow for meaningful meta-analyses (section 4.13).

Depending on what may be adjusted in the JDL methodology following these investigations, any previous customer to the JDL would be welcome to have their data reanalysed to cover such amendments.

## 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**

Ministry of Justice

Justice Data Lab

Justice Statistical Analytical Services

7<sup>th</sup> Floor

102 Petty France

London

SW1H 9AJ

Tel: 0203 334 4770

E-mail: [justice.datalab@justice.gsi.gov.uk](mailto: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](mailto:statistics.enquiries@justice.gsi.gov.uk)

General information about the official statistics system of the United Kingdom is available from [www.statistics.gov.uk](http://www.statistics.gov.uk)

© Crown copyright 2015

Produced by the Ministry of Justice

You may re-use this information (not including logos) free of charge in any format or medium, under the terms of the Open Government Licence. To view this licence, visit

[www.nationalarchives.gov.uk/doc/open-government-licence/](http://www.nationalarchives.gov.uk/doc/open-government-licence/) or email:

[psi@nationalarchives.gsi.gov.uk](mailto:psi@nationalarchives.gsi.gov.uk)

Where this report identifies any third party copyright material, you will need to obtain permission from the copyright holders concerned.