

Data Explained

Identifying alcohol orders and alcohol-defined crimes in probation data

Enforced alcohol abstinence: does it reduce reoffending?

Author: Dr Carly Lightowlers

Date: April 2024

This Data Explained summarises experiences and learning from working with the Data First probation dataset (Jan 2011 – Dec 2020) while producing research into how alcohol-treatment or monitoring requirements are being used and whether these are reducing re-offending. This publication is intended to help guide future researchers using this data and to provide feedback into future dataset development and documentation.

The administrative data discussed in this Data Explained was made securely available through the Data First programme: a ground-breaking data linkage initiative, led by the Ministry of Justice (MoJ) and funded by ADR UK. The data used in this research project comes from the MoJ and was accessed through the Office for National Statistics (ONS) Secure Research Service. The data was not originally collected for research, and it is expected that there are gaps and inconsistencies in its recording, a number of which are detailed in the following.

Project details

In cases where alcohol consumption has played a role in a defendant's offending, courts in England and Wales may make use of alcohol treatment or abstinence requirements as part of a community sentence. Alcohol treatment requirements are aimed at offenders who are dependent on alcohol, whereas alcohol abstinence monitoring requirements are considered suitable for non-dependent drinkers whose drinking contributed to their offending. Compliance with the latter can be backed up with electronic monitoring devices (and thus the threat of further court hearings, fines or imprisonment). Given the considerable social and economic costs associated with alcohol-related crime, the project's insights will contribute to our understanding of how to effectively respond to this problem.

This Data Explained focuses on lessons learned in making use of the Data First probation dataset to identify the characteristics of those receiving alcohol treatment or abstinence requirements and the potential of this data for measuring and assessing compliance with and completion of such requirements. In particular, it reflects on using this data to discern between different types of offending – in general terms and for those offences involving alcohol.

This is the first of two Data Explained reports associated with this project. The second will expand on this to consider further insights available when linking this data to criminal court data to examine re-offending rates associated with alcohol treatment or abstinence requirements.

Initial research questions

The research is guided by two overarching research questions:

- What is the profile of offenders for whom alcohol related treatment and abstinence requirements are used?
- For whom are alcohol related treatment and abstinence requirements effective?

Research methodology

The study will apply both cross-sectional and longitudinal analysis to:

- assess the profile of offenders issued with alcohol treatment or abstinence requirements and for whom these are effective.
- compare whether those given alcohol treatment or abstinence requirements fare better than those who do not receive them.

It will comprise descriptive interrogation of the data, as well as bivariate tests and regression modelling (to include logistic and negative binomial/Poisson regression and event history analysis). [Further details of the analytical strategy and statistical methods to be deployed can be accessed on the OSF Registries](#) (Lightowlers, 2023).

Data and key variables

Probation data comprises an extract from National Delius (nDelius)¹ for 1 January 2014 to 31 December 2020 (based on the referral date).

The initial flat file contains information about nearly two million (n=1,873,228) events (event_id_hash). This can include offence type, disposal, and key dates. There is one row per offender event. The flat file contains information on offender characteristics, the type of offending and the type of disposals given (disposal_type_description) as well as details of recalls. This detail on the type of disposal does not enable identification of alcohol treatment and/or abstinence requirements to be identified. Rather, this detail is contained in the requirements table.

The offender events in the flat file can be joined (by the event_id_hash) to further information about requirements attached to an offender's sentence and key information, such as start and termination dates, held in the requirements file. This returns a dataset with multiple requirements per offender-event (around one and a half million (n=1,546,184) in the joined dataset), and so allows the specific requirements attached to a given community sentence to be identified.²

Some of the variables drawn upon in this study include those detailed in Table 1 overleaf.

¹ nDelius is used for the management of offenders on probation, or in the community. A service user (offender) is referred to nDelius by a court and an event is created in nDelius.

² Further information on pre-sentence reports and associated recommendations, offender's licence conditions following release from prison, and post-sentence supervision are also available. However they were not used as part of this study.

Table 1: Variables included in this study

Data First probation dataset	Variable description and use
Flat file	
event_id_hash	An identifier given each time a new unique event occurs within a journey through nDelius.
referral_date	A referral date for every unique event. <i>To look at year of disposal with accompanying alcohol treatment or abstinence requirements.</i>
ethnicity_desc	A short description of the ethnic group of the offender's ethnicity. <i>Recoded into those identifying as 'White' to contrast with those identifying as another ethnic group to avoid low cell counts in further analysis.</i>
age_at_offence	The age of the offender when the offence took place. <i>Outliers >118 removed.</i>
gender	Gender of the offender.
probation_area_desc	A description for the probation area - where there are multiple records for each event, only the first probation area in each case is used. <i>To look at where alcohol treatment and abstinence requirements are being used.</i>
lsoa_residence	Local area of residence at lower super output area (LSOA) level based on the offender's postcode. <i>To append information from the Indices of Multiple Deprivation.</i>
mo_code	The main offence code. <i>To discern between types of offending.</i>
disposal_type_description	A short description of the disposal. <i>To look at what type of order - community or suspended sentence - the requirements were attached to.</i>
length_in_days and length	The length in days of the disposal and the length in months of the disposal. <i>Used to look at disposal lengths with accompanying alcohol treatment or abstinence requirements.</i>
disposal_termination_date and disposal_date	The date the disposal was terminated and the date of the disposal.

*Duration (in days) of the disposal could be examined by looking at the difference between the disposal_termination_date and disposal_date.
Used to look at disposal duration by alcohol treatment or abstinence requirements.*

Requirements file

rqmnt_type_main_category_desc	<p>A short description of the main type of requirement (can have multiple requirements per event). <i>Included categories of 'Alcohol Abstinence and Monitoring', 'Alcohol Treatment', amongst others, from which to identify cases where one of these requirements were used.</i> <i>This could also be used to examine which other requirements (such as Unpaid Work, RAR and Curfews) were commonly used alongside such requirements (that is, within the same offender-events).</i></p>
rqmnt_termination_reason_desc	<p>A short description for the reason the requirement was terminated. <i>'Requirement Completed' and 'Expired (Normal)' were used to identify those who had successfully completed their requirements and contrast against those who had not.</i></p>
length	<p>The length of the requirement.³ <i>Used to look at requirement length by alcohol treatment or abstinence requirements.</i></p>
rqmnt_termination_date and rqmnt_commencement_date	<p>The date the requirement was terminated. and the commencement date of the requirement. <i>Duration (in days) of the requirement could be examined by looking at the difference between the rqmnt_termination_date and rqmnt_commencement_date.</i> <i>Used to look at requirement duration by alcohol treatment or abstinence requirements.</i></p>

³ As the length of requirements is measured in different units - e.g. alcohol abstinence requirements in days alcohol treatment requirements in months - these are estimates and may not be accurate.

Key variables: for classifying offence type

To assist in furthering understanding about how and when alcohol treatment and abstinence requirements are used, the researcher looked at this by offence type.

The probation dataset contains detailed offence category codes (and associated descriptions) including: `mo_code`, `mo_category_code`, and `mo_category_description`. However, the `mo_category_description` comprised a possible 200 codes in the joined data file, and there was no variable available in the dataset aggregating these up into broader offence types.

Moreover, these codes do not marry with the offence codes and classifications used in other Data First datasets, such as criminal courts data.⁴ Nor are they consistent over time (e.g. as new codes may be added or taken away). [Further details about understanding criminal justice datasets have been published](#) and discuss the discrepancies in offence codes further.

The offence code lookups are included in the respective Data First data catalogues, as shown in Table 2 below.

Table 2: Offence code lookups in Data First data catalogues

Data First Data Catalogue	Variable name(s)
Probation	"offence_code_lookup" tab with <code>mo_code</code> and <code>mo_category</code> codes.
Magistrates' courts	"cjs_offence_code" tab with CJS offence codes and "ho_offence_codes" tab with Home Office codes.
Crown Court	"offence_code" tab with the CJS offence codes and "offence_ho_code" tab with Home Office codes.

⁴ N.B. the most serious offence in a case as the principal offence at the point of the case being received into the court system.

How data limitations were dealt with

In government publications, such as the Criminal Justice Quarterly statistics publication (informed by the court proceedings database), offences are commonly grouped into 12 offence groups, as well as an additional group where the offence is not known, based on Home Office offence codes. These offence groups are:

- Violence against the person
- Sexual offences
- Robbery
- Theft offences
- Criminal damage and arson
- Drug offences
- Possession of weapons
- Public order offences
- Miscellaneous crimes against society
- Fraud offences
- Summary non-motoring
- Summary motoring
- Not known

Deploying this classification assists in making broad comparisons with findings published elsewhere and the offence types to which they refer. As the criminal court data contain Home Office offence codes, which are commonly used in reporting national crime statistics, it was deemed useful to be able to broadly classify the probation caseload into similar categories.

As such, the 2022 Home Office offence classification file as at November 2022 ([downloaded from GOV.UK](#) (offence-group-classification-june-2022.csv)) was used as a lookup file from which to match mo_codes in the probation data file (see R Code excerpt 1). However, given discrepancies between codes, many cases were subsequently missing a Home Office offence classification.⁵

To avoid small numbers in some offence classifications - for example, in further analyses broken down by whether an alcohol treatment or abstinence requirement was attached to a sentence supervised by probation - offence type categories were aggregated. Fraud offences were added to theft offences. Possession of weapon and robbery offences were added to violence against the person offences.

⁵ The 2022 HO classification lookup table is used as opposed to those issued in 2020 or 2021, as this results in fewer incidents classified as missing.

Further, those crimes for which the offence classification was 'not known' were recoded as missing. This resulted in the following nine offence classifications, for use in further analysis:

- Criminal damage and arson
- Drug offences
- Theft offences
- Miscellaneous crimes against society
- Violence against the person
- Public order offences
- Sexual offences
- Summary motoring
- Summary non-motoring

For the purposes of this study, it was of value to be able to identify *alcohol-related* crimes to assess whether alcohol abstinence requirements or alcohol treatment requirements reduce alcohol-related offending specifically. In the absence of such an indicator, an alternative approach was deployed; namely, separating out *alcohol-defined* offences (namely, those crimes defined by the presence of involvement of alcohol supply and/or consumption) from other offences.

To discern between alcohol-defined and non-alcohol defined offences, the approach taken was to identify a subset of offence codes within the published Home Office offence classification (2022) based on the following stem words: "drink", "drunk", "alcohol", "influence of" and "intox" appearing in the text (Detailed_offence) accompanying the relevant offence codes (Offence_code). This enabled the identification of 123 codes relating to offences, which were then used to derive a new variable giving a binary indication as to whether an offence was alcohol-defined offence or not (see R Code excerpt 2). This information could subsequently be appended to each offender-event in the probation dataset (using steps outlined in R Code excerpt 1).

However, this approach does not distinguish between alcohol supply and consumption (i.e. the role alcohol might play) in the offence. Moreover, some offences refer to being under the influence of either alcohol or drugs or both (e.g. causing death by careless driving while under the influence of drink or drugs) – meaning it is also not possible to identify with certainty whether cases in the probation data pertain to offences perpetrated either wholly or partially attributable to the consumption of alcohol (as opposed to drugs).

Suggested improvements recommended to data owners

Future users of the probation dataset may benefit from guidance on deriving offence categories for the analysis of probation data by crime type. This could simply take the form of offering up the approach(es) (and code) adopted here of appending Home Office Offence Classifications. This may need updating if classification and categorisation codes in the probation database change. However, further details on the offence classification used in the nDelius system, and how these map onto CJS offence codes in the criminal court data, are welcome if this approach is preferred or holds promise for ensuring consistency in classifying crimes and analysing defendant journeys through the criminal justice system.

Future users of criminal justice datasets who are more generally looking to pursue analyses of alcohol-*defined* crimes may benefit from the approach (and code) adopted here. The linguistic approach adopted in this study remains flexible in its application to future iterations of the Home Office Offence Classifications, as it can be re-run on updated offence codes and descriptors as and when they are published.

Additional data which would help to further develop the research

While it has been possible to identify alcohol-defined offences in the probation data using the approach outlined here, it is currently not possible to discern between levels of alcohol consumption and involvement in crime in the criminal justice data provided by Data First.

The available criminal court and probation data currently does not enable a distinction between those offences which are alcohol-*related*. Nor does it offer a means of discerning between the levels of alcohol drinking patterns among defendants and offenders. Future releases of probation data containing Offender Assessment System (OASys) details, which include an assessment of an offender's alcohol needs using the Alcohol Use Disorders Identification Test (AUDIT), would assist in this regard. This could potentially offer opportunities to develop proxy indications as to the volume of crime that is alcohol-related, as well as assess the ways in which alcohol treatment and abstinence requirements are being used according to varying levels of alcohol drinking by defendants. It would thus shed further light on with whom these requirements do (and do not) tend to work effectively.

References

Lightowlers (2023). Enforced alcohol abstinence: does it reduce reoffending? OSF Study Preregistration: <https://doi.org/10.17605/OSF.IO/2BZF3>.

Ministry of Justice, released 09 April 2024, ONS SRS Metadata Catalogue, dataset, Ministry of Justice Data First Probation - England and Wales, DOI: <https://doi.org/10.57906/hvz4-m857>

Disclaimer

This work was produced using administrative data accessed through the ONS Secure Research Service. The use of the data in this work does not imply the endorsement of the ONS Secure Research Service or data owners (e.g. the Ministry of Justice) in relation to the interpretation or analysis. This work uses research datasets which may not exactly reproduce National Statistics aggregates. National Statistics follow consistent statistical conventions over time and cannot be compared to Data First linked datasets.

Acknowledgements

This work is supported by ADR UK (Administrative Data Research UK). ADR UK is a partnership transforming the way researchers access the UK's wealth of public sector data, to enable better informed policy decisions that improve people's lives. ADR UK is an Economic and Social Research Council (ESRC) investment (part of UK Research and Innovation) [Grant number: ES/X003566/1].

Contact

Name: Dr Carly Lightowlers

Email: c.lightowlers@liverpool.ac.uk



Appendix: R Code Excerpts

R Code excerpt 1: Appending Home Office Offence codes

```
# Having read in PROBATION data read in offence_group_classification
HOCODES <- read.csv("offence-group-classification-june-2022.csv")
HOCODES$Offence_code <- as.numeric(HOCODES$Offence_code)
PROBATION$mo_code <- as.numeric(PROBATION$mo_code)
library(dplyr)
PROBATION <- left_join(PROBATION, HOCODES, by=c("mo_code" = "Offence_code"))
PROBATION$Offence.group <- as.factor(PROBATION$Offence.group)
levels(PROBATION$Offence.group)
table(PROBATION$Offence.group, useNA = "ifany") #request NAs
```

R Code excerpt 2: Classifying alcohol-defined offences

```
HOCODES <- read.csv("offence-group-classification-june-2022.csv")
#create vector of stem words to search
stem_words <- c("drink", "drunk", "alcohol", "influence of", "intox")
#find indices of the text variable that contain these stem words
indices <- grep(paste(stem_words, collapse="|"), HOCODES$Detailed_offence,
ignore.case=TRUE)
#use indices to extract associated offence codes
alc_offences <- HOCODES$Offence_code[indices]
#create new variable - binary indicator as to whether an offence is alcohol related or not
HOCODES$alcohol_defined <- ifelse(HOCODES$Offence_code %in% alc_offences, 1, 0)
table(HOCODES$alcohol_defined) #124 codes alcohol defined
#print text for alcohol defined offences
HOCODES$Detailed_offence[HOCODES$alcohol_defined==1]
HOCODES$alcohol_defined <- factor(HOCODES$alcohol_defined, levels=c(0,1), labels=c("Not
alcohol defined", "Alcohol defined")) # or "Absent", "Present"
label(HOCODES$alcohol_defined) <- "Alcohol defined"
table(HOCODES$alcohol_defined)
```