The effect of police diversion schemes on offending and health for people suspected of drug-related offences

Analysis protocol

Version 2.0

30 August 2024

Contributors

Name	Role
Ashley Mills: University of Kent; Department of Health & Social Care	Analysis lead
Dan Lewer: University College London; Bradford Institute for Health research; Department of Health & Social Care	Study design / analysis
Jack Cunliffe: University of Kent	Advisor
Brian Eastwood: Department of Health & Social Care	DHSC lead / analysis
Guillermo Reyes Pascual: University of Kent; Department of Health & Social Care	Analysis
Alex Stevens, University of Sheffield	Chief investigator / advisor
Paul Quinton, College of Policing	Chief investigator / advisor

Contact: d.lewer@ucl.ac.uk

Contents

1 Introduction 3
2 Methods 4
3 Statistical power 13
4 Data processing 15
5 Variables required from linked datasets 18
6 Police force-level variables 20
7 Research approvals 21
8 References 22

1 Introduction

Police drug diversion (PDD) schemes may have positive impact on reoffending and other participant outcomes.[1] We define such schemes as alternatives to criminalisation for minor drug-related offences, including—but not limited to—simple possession for personal use. They are alternatives that provide people suspected of such offences with an educative or therapeutic intervention, rather than being processed through prosecution and conviction.

We plan to study the effect of PDD schemes in England and Wales on reoffending and health (with health-related outcomes including entry into drug treatment and hospital episodes related to accidents, drugs, and alcohol).

This study is part of a wider evaluation of PDD schemes that includes qualitative and documentary elements. A protocol for the wider evaluation has been published.[2] This document provides a detailed plan for the quantitative evaluation.

2 Methods

2.1 Research question

This study is part of a wider realist evaluation that aims to understand the "health outcomes, and what works, for whom, in what circumstances and why".[2] This wider evaluation includes both quantitative and qualitative elements. The quantitative component is described in this protocol.

The quantitative component aims to estimate the effect of PDD schemes on suspects. The research question is "what is the effect of police drug diversion schemes on reoffending and health?"

2.2 Study design

We will do a retrospective cohort study. Participants will be classified according to exposure to diversionary policies and followed-up to measure outcomes. Our primary analysis will estimate the effect of diversion policies (a police-force level variable) on these outcomes, with secondary analyses estimating the effect on diverted suspects using an instrumental variable and "per protocol" approach.

2.3 Population

We will study individuals who were suspected of an offence that would mean they are eligible for a diversionary intervention, whether or not such a scheme was available. An example participant would be an individual who was searched on the street and found to be in possession of illegal drugs. We will call this offence the "index offence". Relevant data about these individuals will be collated by police forces.

Participating police forces

Participating police forces and catchment population estimates are shown in Table 1.

Table 1: Police forces in England and Wales, with participation status at 1 July 2024

Police force code	Police force name	Upper tier local authorities (country and unitary authorities)	Population (mid 2022 estimate)[3]	Police-recorded drug offences, year ending Dec 2023[3]	Drug offence rate per 1000	Intervention / control status*
E23000013	Cleveland	Hartlepool, Redcar and Cleveland, Middlesbrough, Stockton-on-Tees	579,300	2,459	4.24	Intervention
E23000008	Durham	Darlington, County Durham	637,600	1,447	2.27	Intervention
E23000007	Northumbria	Gateshead, Newcastle upon Tyne, North Tyneside, South Tyneside, Sunderland, Northumberland	1,466,200	3,350	2.28	Not participating
E23000006	Cheshire	Halton, Warrington, Cheshire East, Cheshire West and Chester	1,108,800	3,246	2.93	Not participating
E23000002	Cumbria	Cumbria	503,000	1,584	3.15	Control
E23000005	Greater Manchester	Bolton, Bury, Manchester, Oldham, Rochdale, Salford, Stockport, Tameside, Trafford, Wigan	2,911,700	14,031	4.82	Control
E23000003	Lancashire	Blackburn with Darwen, Blackpool, Lancashire	1,550,500	2,862	1.85	Not participating
E23000004	Merseyside	Knowsley, Liverpool, Sefton, St. Helens, Wirral	1,442,100	11,364	7.88	Control
E23000012	Humberside	East Riding of Yorkshire, Kingston upon Hull, City of, North East Lincolnshire, North Lincolnshire	943,000	1,596	1.69	Control
E23000009	North Yorkshire	York, North Yorkshire	828,100	1,430	1.73	Not participating
E23000011	South Yorkshire	Barnsley, Doncaster, Rotherham, Sheffield	1,392,100	4,656	3.34	Not participating
E23000010	West Yorkshire	Bradford, Calderdale, Kirklees, Leeds, Wakefield	2,378,100	8,780	3.69	Control**
E23000018	Derbyshire	Derbyshire, Derby	1,067,000	2,311	2.17	Not participating
E23000021	Leicestershire	Leicestershire, Leicester, Rutland	1,136,700	3,381	2.97	Not participating
E23000020	Lincolnshire	Lincolnshire	775,500	1,601	2.06	Not participating
E23000022	Northamptonshire	North Northamptonshire, West Northamptonshire	792,400	2,287	2.89	Control
E23000019	Nottinghamshire	Nottinghamshire, Nottingham	1,163,300	4,061	3.49	Not participating
E23000015	Staffordshire	Staffordshire, Stoke-on-Trent	1,146,200	1,926	1.68	Not participating
E23000017	Warwickshire	Warwickshire	607,600	1,044	1.72	Not participating
E23000016	West Mercia	Herefordshire, County of, Worcestershire, Telford and Wrekin, Shropshire	1,314,000	2,193	1.67	Not participating
E23000014	West Midlands	Birmingham, Coventry, Dudley, Sandwell, Solihull, Walsall, Wolverhampton	2,953,800	7,717	2.61	Intervention
E23000026	Bedfordshire	Bedford, Luton, Central Bedfordshire	715,900	1,821	2.54	Control
E23000023	Cambridgeshire	Cambridgeshire, Peterborough	906,800	1,795	1.98	Intervention
E23000028	Essex	Essex, Southend-on-Sea, Thurrock	1,877,300	5,649	3.01	Not participating
E23000027	Hertfordshire	Hertfordshire	1,204,600	2,135	1.77	Not participating
E23000024	Norfolk	Norfolk	925,300	1,952	2.11	Intervention
E23000025	Suffolk	Suffolk	768,600	1,419	1.85	Intervention
E23000001	Metropolitan	33 London Boroughs	8,855,300	38,918	4.39	Not participating
E23000030	Hampshire	Hampshire, Isle of Wight, Portsmouth, Southampton	2,018,700	5,936	2.94	Control
E23000032	Kent	Kent, Medway	1,875,900	5,271	2.81	Not participating
E23000031	Surrey	Surrey	1,214,500	2,462	2.03	Not participating
E23000033	Sussex	West Sussex, East Sussex	1,721,000	3,928	2.28	Not participating
E23000029	Thames Valley	Buckinghamshire, Bracknell Forest, Oxfordshire, Milton Keynes, Reading, Slough	2,549,700	6,008	2.36	Intervention
E23000036	Avon and Somerset	Bath and North East Somerset, Bristol, City of, North Somerset, South Gloucestershire, Somerset	1,765,400	2,640	1.50	Intervention
E23000035	Devon and Cornwall	Devon, Isles of Scilly, Plymouth, Torbay, Cornwall	1,810,400	3,976	2.20	Not participating
E23000039	Dorset	Bournemouth, Christchurch and Poole, Dorset	785,200	1,016	1.29	Not participating
E23000037	Gloucestershire	Gloucestershire	652,400	1,089	1.67	Not participating
E23000038	Wiltshire	Swindon, Wiltshire	751,500	1,032	1.37	Not participating
W15000004	Dyfed-Powys	Powys, Ceredigion, Pembrokeshire, Carmarthenshire	519,000	1,791	3.45	Not participating
W15000002	Gwent	Caerphilly, Blaenau Gwent, Torfaen, Monmouthshire, Newport	591,400	1,411	2.39	Not participating
W15000001	North Wales	Isle of Anglesey, Gwynedd, Conwy, Denbighshire, Flintshire, Wrexham	688,200	1,384	2.01	Not participating
W15000003	South Wales	Swansea, Neath Port Talbot, Bridgend, Vale of Glamorgan, Cardiff, Rhondda Cynon Taf, Merthyr Tydfil	1,333,100	2,775	2.08	Not participating
	Total		60,227,200	177,734	2.95	
	All research sites		22,791,400	72,836	3.20	
	Intervention		11,086,500	25,437	2.29	
	Control		11,704,900	47,399	4.05	

* Indicative status based on early discussions with police forces. The presence of drug diversion schemes will be established more formally using a structured survey of participating police forces. Participation status is shown at 1 July 2024 and may change.

** West Yorkshire has a relevant diversion scheme in Wakefield.

Inclusion criteria

Participating police forces were asked to select individuals meeting the following criteria:

1. Date: suspected of an offence that occurred between 1 October 2021 and 30 September 2022.
2. Offence: suspected of a qualifying offence, defined as:
    a. GROUP 1: Simple possession of any controlled drug (personal use). This corresponds to Home Office codes:[4] 09261, 09251, 09250, 09260, 09268, 09254, 09259, 09263, 09372, 09253, and 09255
    b. GROUP 2: any of the following offences (Home office codes shown in brackets) in combination with a suspected or proven offence in contravention of the Misuse of Drugs Act 1971 or the Psychoactive Substances Act 2016, in the preceding 3 years, OR a flag in police records for involvement with illicit drugs, OR a positive drugs test on arrest:
        i. shoplifting (04600)
        ii. assault (10501, 10423, 10433, 10508)
        iii. criminal damage (14900, 05800, 05602, 05911)
        iv. drunk and disorderly (14101, 14001, 14112)
        v. any theft (other than burglary) (04910, 04510, 03401, 03900, 05401, 05402, 04400, 03300, 05325, 13001, 03702, 04801, 04300, 04100, 13003, 04700)
3. Age: aged at least 18 years at the date of the offence.
4. Location: the individual lived in the police force area at date of police contact.

Exclusion criteria

We will exclude individuals where:

1. Crucial analytical variables are missing. Crucial variables are: the index offence (eg. possession of cannabis for personal use) and date of index offence.
2. There is identifiable linkage failure. Where linkage algorithms produce a one-to-many or many-to-one match, participants will be excluded from respective analyses. For example, an attempt to link one individual provided in the police data might find two potential matches in the National Drug Treatment Monitoring System (one-to-many); or two different individuals in the police data might match one individual in the National Drug Treatment Monitoring System (many-to-one). Note that this will not eliminate incorrect one-to-zero linkages, meaning that participants who are not matched but do have a record in the target database will be included. This might occur where participants have different name spellings or incorrect dates of birth in one or both databases, for example.
3. Exposure to diversion polices cannot be determined. Police forces may be excluded from the primary analysis if we are unable to determine whether they had a relevant diversion policy during the study.

2.4 Intervention

Diversion schemes vary and often include educational elements, referrals to structured treatment for drug and alcohol use, and case work. We originally planned to evaluate three diversion schemes in Thames Valley, the West Midlands, and Durham. We have developed manuals describing these schemes using the Tidier checklist.[5–7] In this previous plan, we would compare these schemes to three matched ‘control’ forces. Based on geographical and sociodemographic factors, Thames Valley was matched with Hampshire; the West Midlands with Greater Manchester; and Durham with Humberside.

We have since decided to invite other police forces in England and Wales to participate in the study. This is for three reasons: (1) to increase the statistical power; (2) to increase the representativeness of our ‘control’ sample and reduce bias resulting from the 1:1 matching at police force level; (3) to mitigate the risk that some police forces may be unable to provide data for technical reasons or because they do not have capacity. In November 2022 we invited all 43 police forces in England and Wales to participate in the study. All police forces were eligible to participate regardless of whether they had a relevant diversion scheme, since those without such schemes would provide control data. By 1 July 2024, 16 police forces have agreed to participate, of which most have provided data (note the data have not yet been analysed).

We consider a diversion scheme to be a named policy in which individuals meeting certain criteria are offered a diversionary intervention instead of traditional criminal justice pathways. To our knowledge, all police forces offer some kind of diversionary activity, such as referrals to local drug and alcohol services, and out-of-court disposals. However, in most cases these pathways are not offered routinely and can be considered as ‘normal police work’ rather than an explicit diversion scheme.

We have divided diversion schemes into two groups:

1. Group 1 is low-intensity educational programmes designed for people who are found in possession of drugs for personal use. These programmes are typically delivered in a group workshop format (sometimes described as a ‘speed awareness course for drugs’). An example is the ‘DIVERT’ scheme in the West Midlands, in partnership with Cranstoun.[6,8]
2. Group 2 is higher-intensity interventions designed for people with more significant histories of involvement in the criminal justice system. These schemes might include a case worker who assesses the participant’s social needs and creates a package of care together with local drug and alcohol services, housing, and social services. An example is the ‘Checkpoint’ scheme in Durham.[5,9]

2.5 Comparison / control

The control group is individuals from police forces that do not have a relevant scheme.

2.6 Outcomes

Study outcomes are the reoffending rate, entry into drug treatment, and hospital admissions. These are described in Table 2.

Table 2: Study outcomes

Outcome	Description
All-cause reoffending	The count of offences during follow-up. We will include all offences listed in the Police National Computer (not limited to drug-related offences or those in our eligibility criteria).
Entry into drug treatment services	A binary measure of whether the participant started a new episode of structured drug or alcohol treatment. Analysis of this outcome will exclude participants who were enrolled in a structured drug or alcohol treatment programme at the time of the index offence, defined as a live treatment episode/treatment journey (opiate, alcohol, or other) at the time or in the 28 days prior to the index offence. This outcome will be derived from the National Drug Treatment Monitoring System.
Hospital episodes related to drugs, alcohol, and accidents	The count of hospital admissions with a primary diagnosis code of ICD-10 F10-F19 or any code from chapters XIX and XX. This outcome will be derived from the Hospital Episode Statistics database.[10] Inpatient episodes will be processed into 'continuous inpatient spells' (or 'admission').[11]

2.7 Follow-up dates

The entry date will be the date of the index offence (eg. possession of cannabis). The exit date will be the final day on which outcomes could be observed. This is not yet known and will be determined during linkage to national databases. In some analyses, individuals may have earlier censoring dates where they experience a binary outcome before the end of observation.

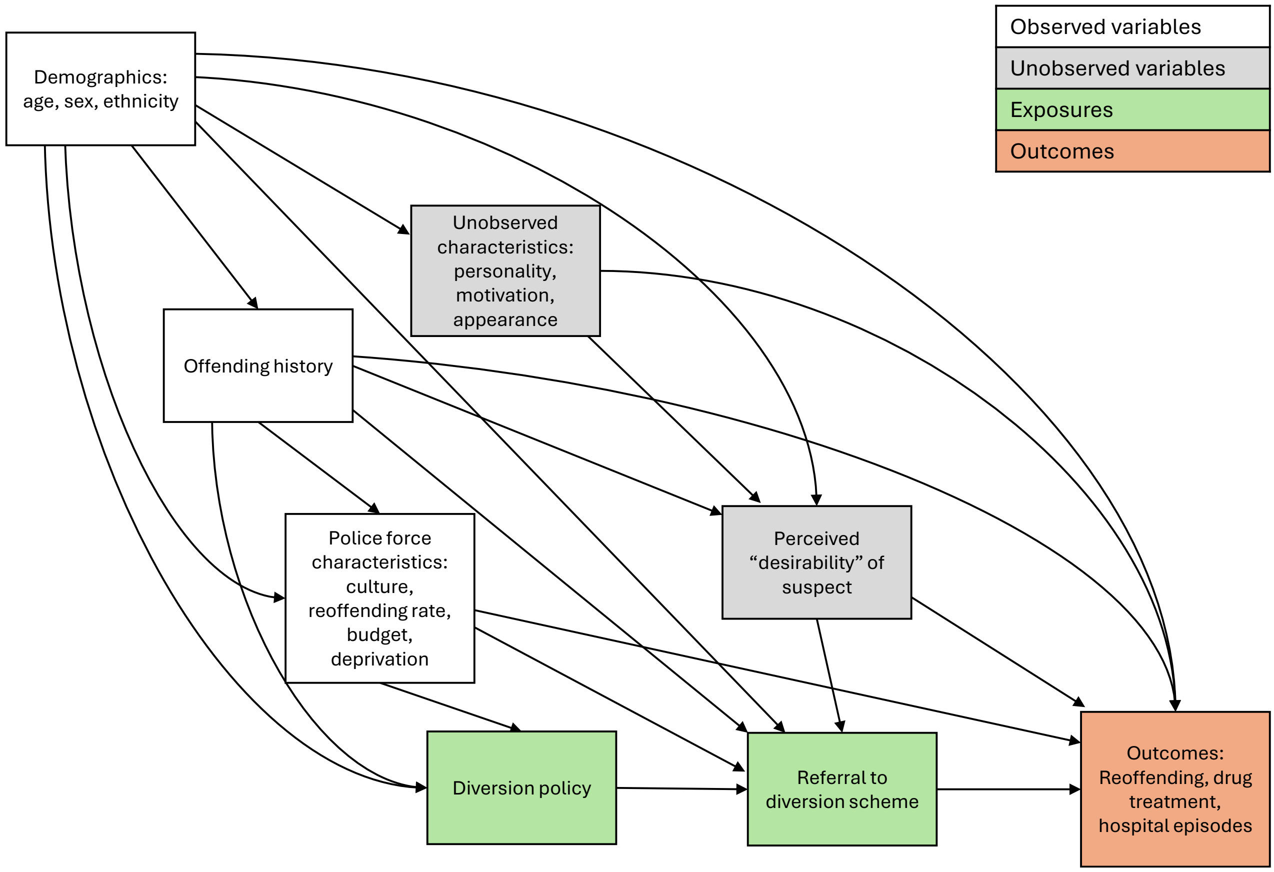
2.8 Causal model

We are trying to estimate the effect of police diversion on outcomes for individuals. A key problem is that police officers select certain types of people for referral diversion schemes. Based on our engagement with police forces, officers are more likely to refer individuals who are perceived to be more likely to attend or complete the intervention, and individuals who are perceived to be more likely to benefit from the intervention. This selection of individuals could be based on unobserved characteristics such as the suspect's personality, motivation, and appearance. We call this process 'officer sorting'. This means that a naïve comparison between suspects who are referred (and/or attend) an intervention with suspects who are not referred is likely to be biased, with the referred group likely to have better outcomes independent of the intervention. It is unlikely that this bias could be controlled by observed variables such as the participant's demographic characteristics and offending history. This type of bias is sometimes called 'confounding-by-indication' and occurs where participants are individually selected for an intervention based on their need for the intervention.

Our assumed causal model is shown as a Directed Acyclic Graph is shown in Figure 1 (see code for daggity.net in appendix).

The effect of a police-force-level diversion policy can be estimated by adjusting for observed individual and force-level characteristics. The unobserved participant characteristics and officer sorting do not confound this analysis because individual officers cannot control the force-level policy at the time of the offence. The exposure is the policy to divert individuals rather the referral or uptake of diversion, meaning that participants in forces with a diversion policy who are not referred are considered exposed.

Figure 1: Directed Acyclic Graph showing our assumed relationships between diversion policies, diversion referrals, and our outcomes



2.9 Analysis methods

We will estimate the effect of police diversion on each of our outcomes in three different ways:

Primary analysis: the effect of diversion policies

In this analysis, the exposure is force level. The analysis will follow two-stages:

1. Force-level analysis: For each intervention force (ie. each force with a PDD policy), we will individually-match participants with control participants from the entire pool of control forces. The control individuals will be exact-matched 1:3 with replacement by age (+/- 3 years), sex, date of index offence (+/- 30 days), and index offence group. The purpose of matching is to create a control group with better face validity, and to improve the independence of estimates across force-level analyses (ie. fewer control participants will be reused across forces). We will then fit a regression model in which the dependent variable is the outcome, the main independent variable is the presence of a diversion scheme, and other independent variables are the potential confounding variables listed in Table 3. Where the dependent variable is binary we will use a binomial model, and where the outcome is a count we will use a negative binomial model. In addition to individual-level confounders, we will adjust for two police-force level variables: the overall reoffending rate from 2018/19 – 2020/21, and number of police officers per capita in September 2021. Given that PDD schemes in our study pre-date the study start date (October 2021), we cannot include the drug-related reoffending rate as the PDD schemes may affect this. We have selected these two variables because we believe they will affect the probability of recorded drug-related reoffending, are relatively stable over time, and are unlikely to be strongly affected by a PDD scheme.

2. Pooling of force-level results: We will plot the effect of diversion from each force on a forest plot and pool the results using random effects meta-analysis. The pooled result will the main effect we are estimating.

We selected this 2-stage approach because (a) it is essential to account for variation across police forces in the frequency of the outcomes. In simulation, we found that analysis methods that do not account for clustering (ie. assume that the intervention is assigned at individual-level) have poor coverage (ie. confidence intervals often exclude the true value) and are too precise; (b) a one-stage approach that has random intercepts for police forces has low power, because it simultaneously tries to estimate the baseline effect of the police force on the outcome and the effect of the intervention, and cannot distinguish between these effects; (c) a two-stage approach without force-level variables (such as the number of police offices and overall reoffending rates) has poor coverage due to differences across police forces in underlying risks of the outcome. If there are strong force-level effects on the outcome, the robustness of our planned analysis depends on the extent to which our force-level variables (number of police officers and overall reoffending rate) represent the underlying force-level risk of drug-related reoffending (note that this cannot be observed empirically because the PDD schemes are already implemented).

This design is supported by simulations of the power and coverage of different methodologies. Code for these simulations is available here:
https://github.com/danlewer/pdd/blob/main/clustered_design_sim.R.

Table 3: Regression model for primary analysis

Formula term	Description
Outcome ~	Dependent variable
PDDPolicyPresence (or WasReferred in secondary analysis B)	Main exposure: PDDPolicyPresence (in primary analysis): True if participant is in an intervention force, false if they are in a control force. WasReferred (in secondary analysis B): True if participant was referred to a diversion scheme, otherwise false.
BaselineReoffendingRate	Baseline reoffending rate for the force within which the offence was committed. See Table 10 for details.
BaselineDrugTreatment	Police-force level penetration of opioid treatment. See Table 10 for details.
ForceFunding	Police force funding per capital in 2021/22
ForceReoffending	Force-level all-cause proven reoffending rate from 2018/19 to 2020/21
offset(LogFollowupDuration)	Log of observation duration (included as an offset term).
IndexOffence	The index offence for which the participant was included in this study, eg. possession of cannabis (categorical variable).
Age	Age in days at the index offence (centred and standardised)
Age-squared	Age in days at the index offence squared (centred and standardised)
Sex	Male / female /other
Ethnicity	Ethnicity either reported by police forces or from linked datasets. This may be participant-defined or officer/professional-defined. Participant-defined ethnicity will be prioritised where both are available. We anticipate differing categorisations and differing values where multiple observations are available. We will aim to clean data so it is consistent with the ONS classification.[12]
HistoricalPosession	Count of simple possession offences in the 5 years prior to the index offence
HistoricalOtherDrugs	Count of other drug-related offences in the 5 years prior to the index offence
HistoricalTheft	Count of theft offences in the 5 years prior to the index offence
HistoricalViolence	Count of violent offences in the 5 years prior to the index offence
HistoticalDrugTreatment	Any structured drug or alcohol treatment episode in the 5 years prior to the index offence (binary)

Secondary analysis A: the effect of diversion referrals (IV approach)

In this analysis, the exposure is individual-level and we will aim to estimate the effect of diversion referrals. We will use an instrumental variable analysis in which the force-level policy is an instrument, assuming that there are limited causal paths other than diversion through which a diversion policy could affect participant outcomes. The purpose is to estimate the effect of diversion on diverted individuals, while controlling for the selection of 'suitable' individuals for diversion ('officer sorting'). Specifically, this analysis estimates the effect on the marginal/additional group that were diverted due to diversion policies. We are not yet sure if a lack of referrals in control forces will affect the feasibility of this analysis, which we will explore when we have collated the data.

Instrumental variables rely on three conditions, with the strength of the instrument depending on how well these conditions are met:[13–16]

1. **Relevance**: the instrument should strongly influence treatment probability. Diversion policies are highly likely to strongly affect the probability of referral. This can be assessed empirically because police forces will report if individuals were offered diversion. Control forces may implement

diversion-type activities even if they have no explicit policy fitting our criteria, but we expect that referrals to diversion programmes will be more common in forces with diversion policies.

2. **Exclusion (or 'exogeneity')**: The instrument should not affect the outcome except through the treatment pathway (ie. diversion programme). One possible pathway in this research is the culture of the police force, which may influence the behaviour of officers.[17] Police officers may observe that they belong to an organisation that aims to limit criminalisation of people who use drugs and support community-based interventions. This may change how police officers engage with people who use drugs, independently of police diversion schemes. We anticipate that such pathways will exist but will be weak in comparison to the actual diversion programmes.

3. **Independence:** the instrument should not be correlated with unmeasured confounders. Possible confounders include (a) the police force's attitude and strategy towards people who use drugs. Police forces that have a more health-oriented view of drugs may be more likely to implement diversion schemes and independently seek non-criminal approaches to drug-related offending; (b) police force per-capita budgets, where those with higher budgets may be more likely to implement discretionary diversion schemes and have better ability to reduce reoffending (which all forces seek to reduce); (c) crime and reoffending rates; for example higher-crime areas may be more likely to implement diversion policies and have higher reoffending rates. We aim to control these force-level confounders (see Table 10).

Secondary analysis B: the effect of diversion referrals (per-protocol approach)

We will compare diverted individuals with non-diverted individuals. This analysis is comparable to a per-protocol approach in an experimental trial, and will provide a 'naive' estimate of the effect of diversion. In terms of causal inference, this approach is flawed because it does not account for the role of police officers selecting 'suitable' individuals for diversion ('officer sorting'). However, it may provide (a) an indication of the degree of residual confounding, for example if we observe large effects in this analysis and small effects in the other analyses; (b) a measure of the difference in outcomes between diverted and non-diverted participants. This analysis will follow the same procedure as the primary analysis, but the referral to diversion will be main exposure (rather than the force-level policy).

Since the referral varies by individual, it will also be possible to do a one-stage analysis in which variables at the police force level are modelling using a hierarchical model.

3 Statistical power

We undertook simulation in March-June 2023 to estimate power of our primary analysis and identify practical issues that may arise. We did the simulation in R[18] and published code in a public repository.[19]

Parameters

Key parameters are shown in Table 4.

Table 4: key parameters for power analysis

Parameter	Value	Source
Number of participating police forces	18 (of which 12 had a relevant diversion policy). Sensitivity analyses with 6 and 12 forces.	Discussions with police forces
Mean per-force participant count (Group 1)	500	Discussions with police forces
Mean per-force participant count (Group 2)	500	Discussions with police forces
Baseline hospital admission rate (Group 1)	0.0248 per person-year	Prior study of hospital admission rates in the general population of young people
Baseline hospital admission rate (Group 2)	0.123 per person-year	Prior study of people who use heroin and crack cocaine[20]
Baseline probability of reoffending within one year	0.25	MOJ Proven Reoffending Statistics (25.8% for the (overall rate of 25.8%, April-June 2022) [21]
Baseline probability of entering drug treatment within one year	0.1	
Standard deviation in police-force level outcome probability	Varied across scenarios	NA

Simulation

We simulated datasets in which outcomes were drawn from a binomial (for binary outcomes) or negative binomial (for count outcome) distributions. Individuals in police forces with a diversion policy had an outcome probability reduced by a specific quantity (the effect size). To estimate power, we simulated the dataset under different effect sizes 1000 times and then fit the regression models shown in Table 5. We then calculated the proportion of simulations in which a significant effect ($p<0.05$) was found. Note that these analyses do not include individual or force-level confounding variables, which will be included in the real analysis.

Table 5: Regression models used for each outcome

Outcome	Model function in R	Formula	Family (assumed distribution)
Reoffending	lme4::glmer	ReoffendedWithinYear ~ PDDPolicy+ OffenceGroup + (1 | PoliceForceID)	binomial (Note in the planned analysis we will use the count of reoffending, but in the power analysis we assumed a binary outcome)
Entering drug treatment	lme4::glmer	EnteredDrugTreatment ~ PDDPolicy + OffenceGroup + (1 | PoliceForceID)	binomial
Hospital admissions	glmmTMB	HospitalEventCount ~ PDDPolicy + OffenceGroup + (1 | PoliceForceID)	nbinom2

Simulation results

Figure 2 (panel A) shows the estimated power of the analysis of reoffending, showing reasonable power with 12-18 police forces and limited between-force variation in the baseline reoffending rate (ie. where the standard deviation in reoffending rate is smaller than 2ppts, compared to an overall reoffending rate of 25%). Power declines rapidly when the intervention effect is obscured by between-force noise. This illustrates the need to include baseline reoffending rates for each force to control in the analysis.

Figure 2: power of analysis, with varying between-force variance in baseline outcome rates and the number of participating forces

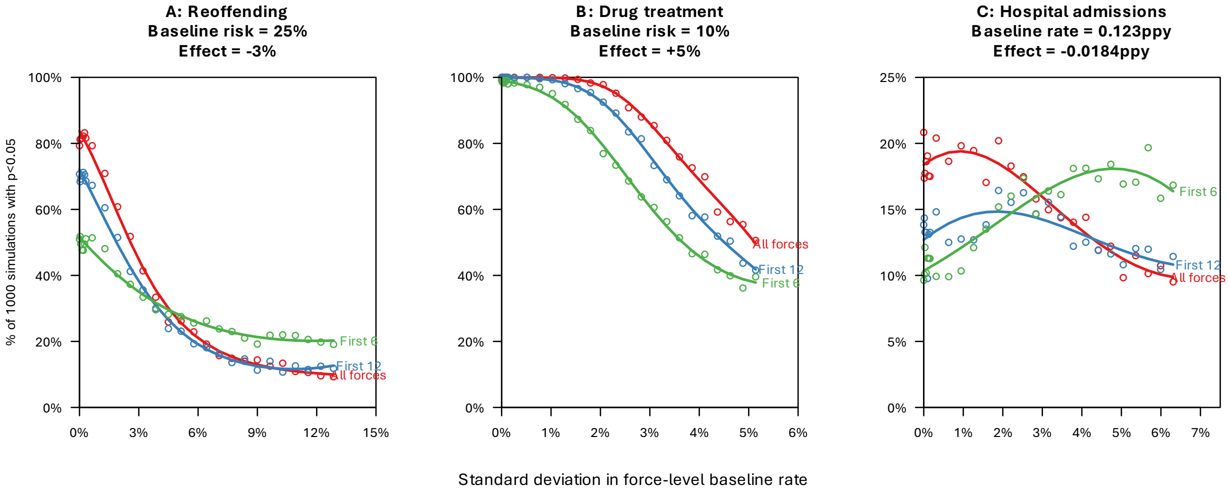


Figure 2 (panel B) shows the estimate power for the analysis of entry into drug treatment. This outcome has higher power, and reasonable power can be achieved with only six forces, assuming limited between-force variation in outcomes. Again, power declines as the between-force variation increases, demonstrating the need to control baseline force rates of drug treatment entry.

Figure 2 (panel C) shows the estimate of power for the analysis of hospital episodes. This shows there is low power under all scenarios. The results for the ‘first 6 forces’ shows an increase in Type 1 errors inflating the apparent power, as random differences in the baseline hospitalisation rates cause a significant difference between exposure groups. This Type 1 (ie. false positive) error could cause a significant result in either direction (ie. some of the ‘significant’ results suggest that diversion schemes are harmful). We also examined doubling the assumed impact of PDD on hospitalisation from a 15% reduction relative to the baseline to a 30% reduction. With zero between-force baseline variation, we estimated 60% power.

4 Data processing

4.1 Identification of participants by police forces

Participating police forces will extract eligible offences together with relevant variables from their record management systems. The eligibility criteria are in section 2.3. For each index offence, the police will include the following variables:

1. Whether the qualifying offence was in Group 1, Group 2, or both
2. First and last name
3. Date of birth
4. Sex/gender
5. Ethnicity, if available (self-identified and/or officer identified)
6. Date of police contact/offence
7. The offence Home Office code
8. The specific drug if the offence was possession (eg. “drugs-possession: cannabis”)
9. Whether the participant was referred to a diversion scheme
10. Whether the participant was arrested
11. Home Office outcome type (eg. “outcome 7: cannabis/khat warning” or “outcome 22: no further action”), which may not be available for more recent participants
12. PNC number, if available

The police forces will encrypt their dataset and send it to DHSC via an access-controlled link to a secure repository where they can upload the encrypted copy. Each participating force will sign a data sharing agreement with DHSC.

4.2 Data cleaning

Cleaning at DHSC will include:

- Checking that data appears valid and all variables have been provided.
- Comparing the rate of offences across participating police forces.
- Comparing the characteristics of offences and participants across police forces.
- Liaising with police forces to understand missing data.
- Removing offences that are not eligible for the study.
- Harmonising variables, particularly ethnicity.

Once the dataset has been cleaned, DHSC will add a unique pseudonymous ID number for all index offences.

4.3 Linkage

The cohort will be linked to four national databases, which are summarised in Table 6.

Table 6: National databases that will be linked to the cohort at individual level

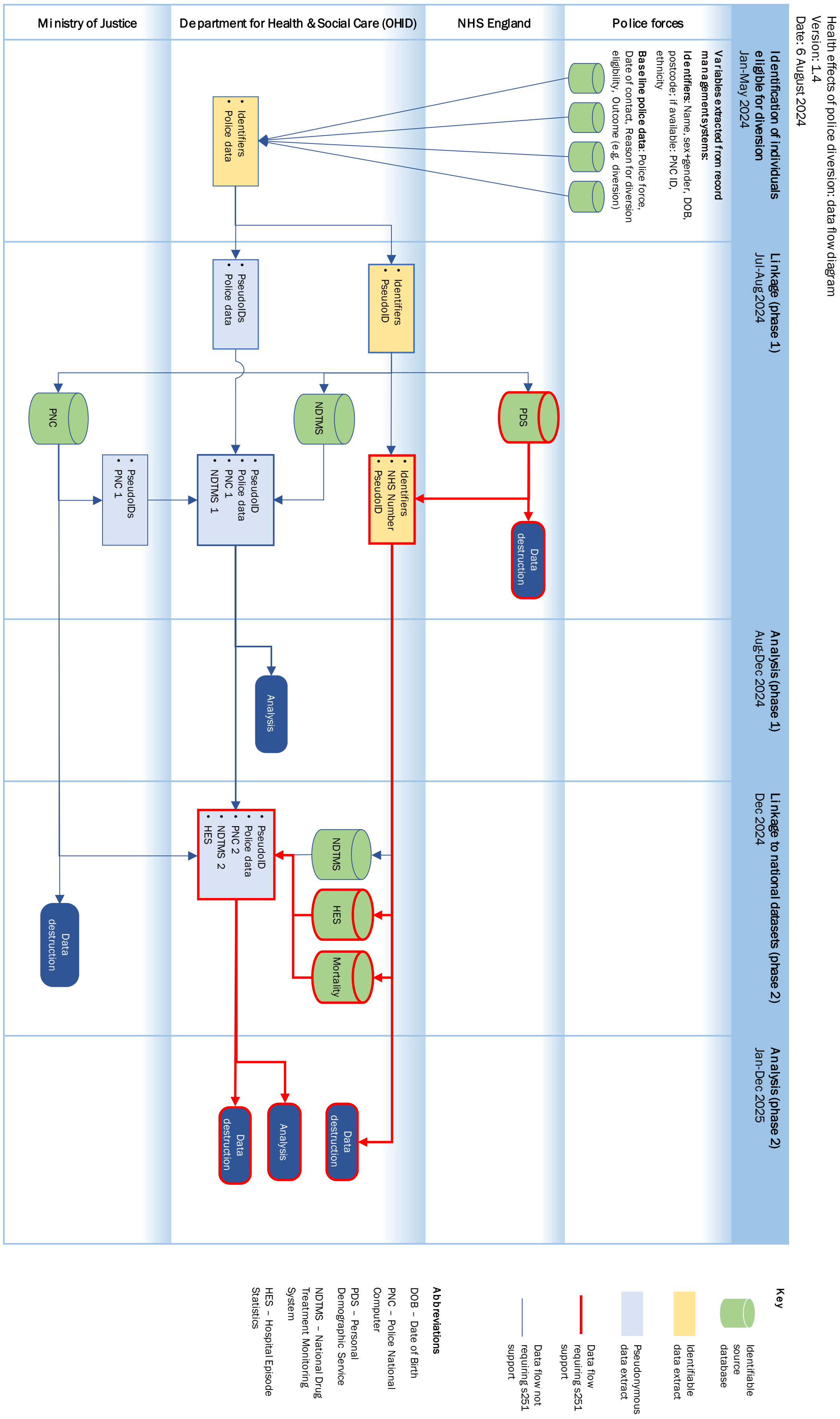
Database	Location	Research variables	Linkage method	Notes
Police National Computer	MOJ	Offending history (in 5 years before index offence) Reoffending	Name, DOB, sex, PNC number, and Pseudo-ID will be sent to MOJ. MOJ will return Pseudo-ID and PNC activity variables	
National Drug Treatment Monitoring System	DHSC	History of drug and alcohol Treatment Entry into drug and alcohol treatment after index offence	Deterministic linkage using initials, DOB, sex, and police force area	
Hospital Episode Statistics	DHSC / NHS England	Comorbidities at study entry Hospital episodes after index offence	Name, DOB, sex, and PseudoID will be sent to NHS England. NHS England will match on Patient Demographic Service to find NHS Number and return PseudoID and NHS Number. DHSC will then link NHS Number to Hospital Episode Statistics	Power may be very low – this will be explored further prior to formal analysis
General Mortality Register	DHSC	Deaths (for censoring follow-up)	Deterministic linkage using name, DOB, sex	Not central to research since the mortality rate will be low in this young population*

* In a cohort of 106,789 people who use illicit opioids (ie. heroin) followed-up from 2000-2018 in England,[22] the all-cause mortality rate for participants aged 18-29 was 61 per 10,000 person years, compared to 6 per 10,000 person-years among people of the same age in the general population. The mortality rate in our cohort is likely to be lower than that of people who use heroin. Therefore, if we have a cohort of 20,000 participants with 1 year follow-up, we expect between 12 and 120 deaths.

4.4 Data deletion

Ministry of Justice and NHS England will delete all personal identifiers once linkage is complete. DHSC will delete personal identifiers when the analytical dataset is complete. DHSC will delete all data when the research is complete.

4.5 Data flow diagram



5 Variables required from linked datasets

Based on our experience of using the linked datasets and existing documentation,[10,23,24] the variables we will request for this research are listed in Table 7, Table 8, and Table 9.

Table 7: Variables required from Hospital Episode Statistics (Admitted Patient Care)

Field	Description	Justification
NHSNOIND	Describes the status of NHS number: verified or being traced	If NHS number is being traced, other identifying principles must be used to link
NEWNHSNO	NHS Number	Primary linkage identifier
HOMEADD	Postcode portion of patient address	Secondary identifiers for linkage when NHS number is unavailable
DOB	DOB of patient	
SEX	Self-reported gender (observed when patient unable to report)	
ETHNOS	Self-reported ethnicity (2001 census codes)	
HESID	Patient identifier - HES generated	Used to find matching records for a given patient
EPIKEY	The record identifier created by the HES system.	
ADMIDATE	The date the patient was admitted to hospital at the start of a hospital spell. ADMIDATE is recorded on all episodes within the spell.	To process continuous inpatient spells and identify admissions
DISDATE	The date on which the patient was discharged from hospital. It is only populated on the final episode of the spell.	
EPISTAT	Episode status (finished or unfinished)	
EPITYPE	Episode type (we are interested in Type 1 - General)	
EPIORDER	Order of the episode in the current spell	
PROCODET	Provider code of treatment	To evaluate the geographical mobility of the study population
ADMIMETH	A code to identify how the patient was admitted to hospital. ADMIMETH is recorded on the first and also all subsequent episodes within the spell (i.e. where the spell is made up of more than one episode).	The outcome is unplanned/emergency admissions
DISMETH	The circumstances under which a patient left hospital. For the majority of patients this is when they are discharged by the consultant. This field is only completed for the last episode in a spell.	To evaluate rates of self-discharge and validate linked ONS mortality data
DIAG_nn	ICD-10 coded diagnosis, nn is the diagnosis order so DIAG_01 is the primary diagnosis, DIAG_02 is the secondary diagnosis and so on.	Identify "drug and alcohol related" episodes and episodes related to injuries and accidents

Table 8: Variables needed from the Police National Computer

Field	Description	Justification
CaseID	Case number	To group offences
OffenceID	Offence number within case	To group offences
OffenceStartDate	Date of offence	Date of event used in research
ForceDescription	Eg. "Metropolitan police"	To evaluate the geographical mobility of the study population
HOOffenceCode	Home Office offence code	To classify historical offences, and for describing reoffending
OffenceDescription	Home Office offence	As above
IsPrimaryOffence	Y/N – used where multiple offence codes are attached to one event	
OffenceStartAge	Age at offence	To validate linkage
Adjudication description	E.g. Guilty / caution	
Ethnicity description	Ethnicity of participant as reported by police force	To help with missing data
Sentence date		
Disposal category	e.g. immediate custody, community sentence, fine, caution	To allow censoring during prison episodes
Disposal rank	Severity of disposal (1 is most severe)	

Table 9: Variables needed from the National Drug Treatment Monitoring System

Variable Name	Description	Justification
FINITIAL	Client's first name initial	Linkage variables
SINITIAL	Client's surname initial	
DOB	Client's date of birth	
SEX	Client sex at birth	
AGNCY	CJIT Agency code[24]	Provides an approximate location, used for linkage
PC	Partial postcode of client, note that no fixed abode is coded as ZZ99 3	
LA	Local authority code[24]	
ETHNIC	Ethnicity (2001 census code)	To help with missing data
REFLD	Date of event which lead to the referral	Linkage - should match the police referral date if present
OFFENCE	Offence that lead to the referral	To describe offences before treatment episodes
CONSENT	Whether the client consented to be included in research	To remove non-consenting participants from analysis
SEX	Client sex at birth	
CPLANDT	Caseload start date	Used to measure duration in treatment
DISD	Closure date	
REFDATE	If they were referred to structured treatment, this is the date at which that occurred	Descriptive analysis

6 Police force-level variables

Table 10: Potentially relevant police-force level variables

Code	Name	Drug offences per 1,000 [i]	Drug-related reoffending rate[ii]	Treatment penetration [iii]	LSOAs in most deprived decile [iv]	Population per km² [v]	Budget per capita, £ [vi]	Police officers per 10,000 population [vii]	Overall reoffending rate[viii]
E23000013	Cleveland	4.24	0.28	0.51	0.31	940	191	24.61	0.33
E23000008	Durham	2.27	0.24	0.43	0.13	262	161	19.26	0.29
E23000007	Northumbria	2.28	0.21	0.45	0.18	260	186	23.78	0.28
E23000006	Cheshire	2.93	0.21	0.44	0.08	466	124	19.76	0.26
E23000002	Cumbria	3.15	0.25	0.52	0.08	70	154	25.32	0.27
E23000005	G Manchester	4.82	0.17	0.40	0.23	2,282	180	24.94	0.22
E23000003	Lancashire	1.85	0.18	0.48	0.20	475	149	20.89	0.24
E23000004	Merseyside	7.88	0.24	0.50	0.34	1,775	210	27.47	0.27
E23000012	Humberside	1.69	0.21	0.43	0.22	258	158	22.29	0.29
E23000009	North Yorkshire	1.73	0.28	0.38	0.02	99	110	18.62	0.25
E23000011	South Yorkshire	3.34	0.20	0.45	0.23	897	164	20.29	0.27
E23000010	West Yorkshire	3.69	0.21	0.49	0.22	1,172	161	23.99	0.27
E23000018	Derbyshire	2.17	0.19	0.57	0.07	406	122	17.93	0.24
E23000021	Leicestershire	2.97	0.19	0.45	0.07	446	120	19.81	0.22
E23000020	Lincolnshire	2.06	0.23	0.50	0.07	127	101	14.61	0.25
E23000022	Northamptonshire	2.89	0.21	0.47	0.06	335	111	17.64	0.23
E23000019	Nottinghamshire	3.49	0.22	0.48	0.13	538	140	19.10	0.26
E23000015	Staffordshire	1.68	0.15	0.50	0.09	422	123	15.52	0.21
E23000017	Warwickshire	1.72	0.17	0.39	0.02	307	106	17.33	0.20
E23000016	West Mercia	1.67	0.17	0.45	0.05	177	111	17.60	0.24
E23000014	West Midlands	2.61	0.17	0.46	0.26	3,276	184	25.21	0.24
E23000026	Bedfordshire	2.54	0.21	0.35	0.02	579	114	19.06	0.23
E23000023	Cambridgeshire	1.98	0.20	0.47	0.04	267	105	18.25	0.24
E23000028	Essex	3.01	0.20	0.38	0.04	475	110	18.51	0.23
E23000027	Hertfordshire	1.77	0.20	0.40	0.00	733	118	18.75	0.24
E23000024	Norfolk	2.11	0.24	0.48	0.07	168	114	19.55	0.27
E23000025	Suffolk	1.85	0.21	0.43	0.05	200	110	17.14	0.23
E23000001	Metropolitan	4.39	0.25	0.26	0.02	5,564	236	37.75	0.25
E23000030	Hampshire	2.94	0.20	0.46	0.04	475	117	15.50	0.22
E23000032	Kent	2.81	0.19	0.40	0.06	480	119	20.68	0.21
E23000031	Surrey	2.03	0.19	0.34	0.00	727	99	16.97	0.20
E23000033	Sussex	2.28	0.19	0.40	0.04	449	115	16.68	0.20
E23000029	Thames Valley	2.36	0.21	0.40	0.01	444	109	17.31	0.23
E23000036	Avon & Somerset	1.50	0.21	0.45	0.06	351	120	17.40	0.23
E23000035	Devon & Cornwall	2.20	0.21	0.49	0.06	173	120	18.33	0.23
E23000039	Dorset	1.29	0.20	0.42	0.03	291	100	16.47	0.23
E23000037	Gloucestershire	1.67	0.17	0.43	0.03	241	109	18.26	0.21
E23000038	Wiltshire	1.37	0.21	0.43	0.03	216	101	14.37	0.23
W15000004	Dyfed-Powys	3.45	0.24	NA	NA	47	66	23.25	0.24
W15000002	Gwent	2.39	0.21	NA	NA	366	87	23.06	0.26
W15000001	North Wales	2.01	0.25	NA	NA	108	72	23.83	0.27
W15000003	South Wales	2.08	0.28	NA	NA	620	90	23.64	0.31

Definitions and sources

i. Drug offences per 1,000 = Police-recorded drug offences in the year ending 2023[3] divided by the mid-year population estimate mid-2022[3] * 1000

ii. Proven reoffending within 12 months for drug-related offences, July 2021-June 2022[21]

iii. Treatment penetration = Number in treatment for opioid use 2019/20[25] / estimate of opiate and crack cocaine users in 2019/20 estimated by capture-recapture methods[26]

iv. Proportion of LSOAs in most deprived decile nationally = Number of LSOAs in most deprived decile of Index of Multiple Deprivation 2019[27] / Number of LSOAs in the police force area

v. Population per km2 = population estimate mid-2022[3] / area of police force area in km2[28]

vi. Budget per capita = Police grant in 2022-23[29] / population estimate mid-2022[3]

vii. Police officers per 10,000 population = police officer FTE in September 2021[30] / population estimate mid-2022[3] * 10,000

viii. Overall reoffending rate = number of proven reoffences / number of offenders (Oct 2018 – Sep 2021)[21]

7 Approvals

This research has the following approvals:

- Ethics. NHS East of England - Cambridge South Research Ethics Committee provided a favourable opinion on 27 June 2023, reference 23/EE/0114. This includes all aspects of the research in this protocol.
- Non-consented processing of data. The Health Research Authority Confidentiality Advisory Group provided ‘section 251 support’ on 29 June 2023, reference 23/CAG/0052. This includes the processing of health and social care data, specifically from the Patient Demographic Service and Hospital Episode Statistics.
- Use of data from the Police National Computer. Our project was approved by the Ministry of Justice Data Access Governance Board on 24 February 2023. This includes linkage of our cohort to the Police National Computer.
- Research sponsorship. The University of Kent is the research sponsor for this project. The University of Kent has produced a Data Protection Impact Assessment (DPIA) that has been approved by its data protection officer. Where necessary, other organisations involved in data processing have also completed DPIAs and sought guidance or local approvals from data protection professionals.
- Local approvals from participating police forces, together with Data Sharing Agreements between police forces and the Department for Health and Social Care.

8 References

1. Stevens A, Hughes CE, Hulme S, Cassidy R. Depenalization, diversion and decriminalization: A realist review and programme theory of alternatives to criminalization for simple drug possession. Eur J Criminol. 2022;19: 29–54. doi:10.1177/1477370819887514
2. Stevens A, Hendrie N, Bacon M, Parrott S, Monaghan M, Williams E, et al. Evaluating police drug diversion in England: protocol for a realist evaluation. Health Justice. 2023;11: 46. doi:10.1186/s40352-023-00249-2
3. Office for National Statistics. Crime in England and Wales: Police Force Area data tables. 2024. Available: https://www.ons.gov.uk/peoplepopulationandcommunity/crimeandjustice/datasets/policeforceareadatatables
4. Home Office UG. Offence group classification - Q2 2023. Available: https://assets.publishing.service.gov.uk/media/65a7e16aed27ca000d27b149/offence_group_classification_june_2023.xlsx
5. Hendrie N, Bacon M, Smith R, Bostock M, Quinton P, Stevens A. The Durham Drug Diversion Scheme. Checkpoint: Descriptive Manual. 2023. Available: https://kar.kent.ac.uk/101848/
6. Hendrie N, Bacon M, Smith R, Bostock M, Quinton P, Stevens A. The West Midlands Police Drug Diversion Scheme DIVERT: Descriptive Manual. 2023. Available: https://kar.kent.ac.uk/101852/
7. Hendrie N, Bacon M, Smith R, Bostock M, Quinton P, Stevens A. The Thames Valley Police Drug Diversion Scheme. TVP: Descriptive Manual. 2023. Available: https://kar.kent.ac.uk/101846/
8. Cranstoun. About Divert. 2023. Available: https://cranstoun.org/help-and-advice/alcohol-other-drugs/divert/
9. Weir K, Routledge G, Kilili S. Checkpoint: An Innovative Programme to Navigate People Away from the Cycle of Reoffending: Implementation Phase Evaluation. Polic J Policy Pract. 2021;15: 508–527. doi:10.1093/police/paz015
10. Herbert A, Wijlaars L, Zylbersztejn A, Cromwell D, Hardelid P. Data Resource Profile: Hospital Episode Statistics Admitted Patient Care (HES APC). Int J Epidemiol. 2017;46: 1093–1093i. doi:10.1093/ije/dyx015
11. Boyd A. Understanding Hospital Episode Statistics (HES). CLOSER; 2018. Available: https://www.closer.ac.uk/wp-content/uploads/CLOSER-resource-understanding-hospital-episode-statistics-2018.pdf
12. Office for National Statistics. Ethnic group, national identity and religion: Measuring equality: A guide for the collection and classification of ethnic group, national identity and religion data in the UK. 2022. Available: https://www.ons.gov.uk/methodology/classificationsandstandards/measuringequality/ethnicgroupnationalidentityandreligion
13. Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference: Instrumental variable methods for causal inference. Stat Med. 2014;33: 2297–2340. doi:10.1002/sim.6128
14. Labrecque J, Swanson SA. Understanding the Assumptions Underlying Instrumental Variable Analyses: a Brief Review of Falsification Strategies and Related Tools. Curr Epidemiol Rep. 2018;5: 214–220. doi:10.1007/s40471-018-0152-1
15. Stel VS, Dekker FW, Zoccali C, Jager KJ. Instrumental variable analysis. Nephrol Dial Transplant. 2013;28: 1694–1699. doi:10.1093/ndt/gfs310
16. Greenland S. An introduction to instrumental variables for epidemiologists. Int J Epidemiol. 2000;29: 722–729. doi:10.1093/ije/29.4.722
17. Wilson J. Varieties of Police Behavior. Boston, MA: Harvard University Press; 1968. Available: https://www.hup.harvard.edu/books/9780674932111
18. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing; 2022. Available: https://www.R-project.org/
19. Mills A, Lewer D, Cunliffe, Jack. PDDModel: A Simulation Study on Police, Drugs & Diversion. 2023. Available: https://github.com/staropram/pddmodel
20. van Hest N, Brothers TD, Williamson A, Lewer D. Health-care resource use among patients who use illicit opioids in England, 2010–20: A descriptive matched cohort study. Addiction. 2023; add.16401. doi:10.1111/add.16401
21. Ministry of Justice. Proven reoffending statistics: April to June 2022. 2024. Available: https://www.gov.uk/government/statistics/proven-reoffending-statistics-april-to-june-2022
22. Lewer D, Brothers TD, Van Hest N, Hickman M, Holland A, Padmanathan P, et al. Causes of death among people who used illicit opioids in England, 2001–18: a matched cohort study. Lancet Public Health. 2022;7: e126–e135. doi:10.1016/S2468-2667(21)00254-1
23. Office for Health Improvement and Disparities. National Drug Treatment Monitoring Service Website. 2024. Available: https://www.ndtms.net/
24. Hughes L. National Drug Treatment Monitoring System (NDTMS) - CJIT technical definition - Core dataset P. Public Health England; 2020. Available: https://assets.publishing.service.gov.uk/media/5e66187e86650c513af60993/NDTMS_CJIT_Technical_Definition_CDS-P_V4.1.pdf
25. Office for Health Improvement and Disparitie, The University of Manchester. National Drug Treatment Monitoring System - Adult Profiles. 2024. Available: https://www.ndtms.net/ViewIt/Adult
26. Office for Health Improvement and Disparitie, UK Health Security Agency. Opiate and crack cocaine use: prevalence estimates. 2023. Available: https://assets.publishing.service.gov.uk/media/6537a3145e47a5000d9898ba/OCU-prevalence-estimates-England_2019-2020.ods
27. Ministry of Housing, Communities & Local Government. English indices of deprivation 2019. 2019. Available: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2019
28. Open Geography Portal. Standard Area Measurements for Administrative Areas (December 2021) in the UK (V2). Office for National Statistics; 2024. Available: https://geoportal.statistics.gov.uk/datasets/ons::standard-area-measurements-for-administrative-areas-december-2021-in-the-uk-v2/about

29. Home Office. Police grants in England and Wales: 2022 to 2023. 2022. Available: https://www.gov.uk/government/publications/police-grants-in-england-and-wales-2022-to-2023
30. Home Office. Police workforce, England and Wales: 30 September 2021. 2022. Available: https://www.gov.uk/government/statistics/police-workforce-england-and-wales-30-september-2021

9 Appendix: daggity.net code

```
dag {
"offending history" [adjusted,pos="-1.203,-0.222"]
"perceived desirability" [latent,pos="-0.476,0.839"]
"police force characteristics" [adjusted,pos="-0.977,0.920"]
"unobserved charactersitics" [latent,pos="-0.797,-0.508"]
demographics [adjusted,pos="-1.930,-0.500"]
diversion [pos="-0.316,1.440"]
pdd_policy [exposure,pos="-1.217,1.452"]
reoffending [outcome,pos="0.564,1.428"]
"offending history" -> "perceived desirability"
"offending history" -> "police force characteristics"
"offending history" -> diversion
"offending history" -> pdd_policy
"offending history" -> reoffending
"perceived desirability" -> diversion
"perceived desirability" -> reoffending
"police force characteristics" -> diversion
"police force characteristics" -> pdd_policy
"police force characteristics" -> reoffending
"unobserved charactersitics" -> "perceived desirability"
"unobserved charactersitics" -> reoffending
demographics -> "offending history"
demographics -> "perceived desirability"
demographics -> "police force characteristics"
demographics -> "unobserved charactersitics"
demographics -> diversion
demographics -> pdd_policy
demographics -> reoffending
diversion -> reoffending
pdd_policy -> diversion
}
```