# Research paper: Data and statistical gaps in criminal justice

## Centre for Public Data, March 2023

# 1. Executive summary

Effective use of data is critical in delivering efficient and dynamic government. This is particularly the case where services operate together as a system, such as the criminal justice system.

Despite this, and despite clear calls from interested parties in Parliament and elsewhere, published data on key areas of the justice system are missing. At the Centre for Public Data, we set out to identify data gaps in criminal justice that were important to both Government and others, identify some of the most pressing to fill, and make recommendations for how to do so.

This report sets out the results of that work. We have found stakeholders, including MPs, repeatedly calling for the same data gaps to be filled, with little being done to address the situation. For example, MPs have asked the Government many times in recent years how many defendants appear in magistrates' courts without legal representation, as have researchers and journalists, but we still do not have any official data or statistics on how many people are affected.

The report includes recommendations (set out in detail in the Appendix) in four areas where we have highlighted significant data gaps:

1. **Remand and bail**, including prisoners on remand and the use of conditional bail.
2. **Sentencing**, including court-level sentencing practice, and sentencing for 'flagged' offences like domestic violence and hate crime.
3. **Court operations**, including legal representation, court backlogs and the effect of interventions.
4. **Low-level offences**, including anti-social behaviour and out-of-court disposals.

The good news is that in most cases, the remedies are fairly straightforward to implement. Most of the data being called for is already being collected at some point in the justice system, often by Common Platform, the courts' new digital case management system. The problem is that this data is not being converted into public statistics that meet users' needs. Generating pre-existing information from administrative systems into reporting statistics is not technically complicated for the MoJ, and crucially, it happens at minimal cost.

We further identified some actions (also noted in Appendix 2) that the MoJ could take to improve its performance on data.

Now is the right time for these issues to be addressed. Justice data is going through a period of intense transformation. HM Courts & Tribunals Service is entering the final phase of its £1.3 billion digital reform programme, and over three-quarters of criminal courts are using Common Platform.[1] Notwithstanding Common Platform's well-documented shortcomings, the system has enabled courts to record far greater amounts of information,

---

[1] National Audit Office, 'Progress on the courts and tribunals reform programme' (2023)

often in useful, structured formats. There is a wealth of administrative data available to statistics producers.

Official producers of statistics have a duty to engage with users, as required by the UK's statistics regulator.[2] In particular, statistics producers must "consider whether to produce new statistics to meet identified information gaps". This report provides an account of where such information gaps exist within the criminal justice system. We argue that the MoJ has a responsibility to recognise and address the gaps identified by stakeholders, as presented in this report.

As well as underpinning efficient operation of the justice system, high-quality criminal justice data sheds light on the experiences of victims and defendants and provides necessary accountability. It also tells stakeholders, civil servants and others how the justice system operates and what is working or not. It is not an optional extra.

Data gaps in criminal justice limit the scope of public inquiry, breed mistrust and can serve to disguise poor treatment of particular groups of people. It is only by listening and responding to user needs that the MoJ can address its key data gaps and facilitate a more meaningful conversation around the state of our criminal justice system. This report explains what it has to do.

# 2. Background

This report is a summary of our research into criminal justice data gaps in England and Wales. The research began in August 2022 with two objectives. First, we aimed to establish a robust, novel and efficient methodology to identify important data gaps that could be reused across multiple policy areas, by both Government and third parties. Second, we aimed to identify some of the most pressing data gaps in criminal justice, and make recommendations for how they could be filled. By exploring the underlying causes of these gaps, we aimed to understand which gaps could potentially be filled, and produce practical, feasible recommendations for relevant statistics and data to support the public good.

## 2.1 What are data gaps?

We define data gaps as "areas where a lack of official data means that questions of significant public interest cannot be answered". These gaps might be caused by the underlying information being difficult to record; the collection or publication of data not being a policy priority; or rarely, policymakers simply choosing to withhold potentially embarrassing information.

We have argued before that the power to not collect data is one of the most important but least-understood sources of power that governments have.[3] By refusing to amass or share knowledge in the first place, decision-makers can exert power. Without quantitative data,

---

[2] Office for Statistics Regulation, 'Code of Practice for Statistics - V1: Relevance to users'
[3] Centre for Public Data, 'Introducing Missing Numbers' (2019)

internal and external stakeholders find it much harder to understand where services are failing or share areas of good practice; journalists find it harder to report stories; and campaigners find it harder to make their case for change.

## 2.2 Needs and challenges in criminal justice data

Justice must not only be done, but must be seen to be done. And with the abundance of data offered by modern digital systems, there are more opportunities than ever to create a transparent, accountable, evidence-based justice system.

Good data matters because practitioners need robust data to understand the impact of their actions. Staff inside magistrates' courts have told us that the lack of data magistrates receive on the outcomes of decisions leaves them unclear whether their decisions will lead to successful outcomes. In 2021, the Sentencing Council, which promotes transparency and consistency in sentencing, warned that its analysis was hampered by a "lack of available data".[4]

Public data is also critical for effective external scrutiny of how the justice system operates, whether certain groups are treated differently and whether crimes are tackled effectively. Criminal justice think tank Transform Justice recently told the House of Commons Justice Committee that "we cannot shine light on the justice system without having access to its actors and data".[5]

Data also helps researchers and policymakers understand the criminal justice system (CJS) as a *system*. Meaningful data gives us an end-to-end view of the impact of changes. For example, the full consequences of the court backlog only become visible if we can also observe trends in the population held on remand, or the numbers of victims withdrawing from hearings. The Law Society has warned that better collection of data is needed throughout the courts system, and that currently courts policy is being made "in the dark, without a clear understanding of its impact".[6]

But stakeholders across the CJS and beyond have long noted challenges with criminal justice data, which is fragmented across multiple agencies. Research by the Legal Education Foundation has highlighted the complexity and challenges involved in current CJS data systems, but also the opportunities to improve data quality and access that are offered by the introduction of Common Platform, the courts' new digital case management system.[7]

---

[4] Sentencing Council, 'A Review of Consistency in Sentencing' (2021)
[5] Transform Justice, 'Written evidence from Transform Justice' (2021)
[6] The Law Society, 'Justice delayed: Five steps to resolve the backlogs in our courts' (2022)
[7] The Legal Education Foundation, 'Digital Justice: HMCTS data strategy and delivering access to justice' (2019)

# 3. Methodology

## 3.1 Our methodology

### 3.1.1 Identifying unanswered questions

We began the project by identifying major stakeholders in the criminal justice system, and examining recent work for requests for better data, or statements that better data was needed to answer questions of interest. We chose to focus on written work, rather than interviewing stakeholders, because we wanted to (i) identify needs identified in the course of broader work, rather than when asked specifically to list missing data; and (ii) produce a novel method that could work across a range of policy areas, was low-cost to reproduce, and could potentially be automated in future.

The sources we included were:

1. **Select Committee reports.** We studied all reports published by the House of Commons Justice Committee,[8] the House of Lords Justice and Home Affairs Committee[9] and the House of Commons Public Accounts Committee[10] (only reports relating to the justice system) in the past five years - producing more than 50 candidate reports in total.
2. **Major independent reports:** We studied major independent reports published in the past five years either commissioned by Government, or published by independent bodies with a statutory role. Reports commissioned by Government were selected by using the Ministry of Justice's publication page on GOV.UK and choosing reports filtered by 'All research and statistics' with documents categorised as 'Independent report'.[11] Reports by independent bodies were limited to include the National Audit Office (NAO) and the Office for Statistics Regulation (OSR) plus the five Criminal Justice Inspectorates.[12] NAO reports were selected by filtering publications by 'Reports' and 'Crime, justice and law',[13] and OSR reports and commentary were selected by filtering publications by 'Systemic/Monitoring Reviews' with the 'Crime and Security' theme.[14] This produced 43 candidate reports in total.

---

[8] Justice Committee reports from the past five years are found at two sources: UK Parliament, 'Reports, special reports and government responses: Justice Committee'; UK Parliament, 'Publications - Justice Committee' [Accessed 1 September 2022]

[9] UK Parliament, 'Reports, special reports and government responses: Justice and Home Affairs Committee' [Accessed 1 September 2022]

[10] UK Parliament, 'Reports, special reports and government responses: Public Accounts Committee' [Accessed 1 September 2022]

[11] GOV.UK, 'Research and statistics' [Accessed 1 September 2022]

[12] GOV.UK, 'Criminal Justice Inspectorates' [Accessed 1 September 2022]

[13] National Audit Office, 'Reports' [Accessed 1 September 2022]

[14] Office for Statistics Regulation, 'Publications: Systemic/Monitoring Reviews' [Accessed 1 September 2022]

3. **Major civil society reports**: To identify relevant stakeholders, we compiled a list of all civil society organisations that had provided oral evidence to the Justice Committee since August 2020.[15] We then compiled relevant research publications from each organisation, producing 160 candidate reports in total.

Once we had identified relevant sources, we read the summary and recommendations in each report from sources 1-3 above, and searched the report for relevant phrases including 'data', 'information', 'gaps', 'lack', 'limited' and others, to find indicators of data gaps. This produced more than 500 potential indicators for data gaps.[16] We published all this research under an open licence, so that others might draw freely on our methodology and findings.

We also studied an important additional data source:

4. **Parliamentary Questions (PQs)**: PQs allow MPs and peers to request information from Government departments, which are then published online. We obtained the text of all written PQs from the House of Commons to the MoJ between December 2019 and July 2022, via web scraping, producing 4,802 PQs. We identified which questions sought quantitative information by searching for the phrases 'how many', 'how much', 'what estimate', 'what proportion' etc, producing 1,663 PQs. We read a random sample of answers to identify common text patterns that indicated the question could not be answered due to a lack of data - this indicated patterns such as 'is not held', 'do not collect', 'unable to answer' etc[17]. We then filtered all answers that contained these and similar phrases, using a regular expression, producing 596 candidate PQs. Finally, we manually reviewed these questions and their answers, giving us a set of 335 PQs where it appeared that data requested by MPs genuinely was unavailable. Again, we published our code and findings as open research.[18]

Finally, we reviewed the full list of identified data gaps and unanswered PQs to identify common themes. We published these findings in an interim report discussing the broad themes, along with our full list of sources and findings.[19] This report highlighted areas including court operations; legal representation in magistrates' courts; sentencing disparities; and a number of other areas. Feedback from stakeholders on this report helped shape the next phase of our work.

## 3.1.2 Identifying priority data gaps

Having established this methodology and proven that it was successful in identifying indicators of potential gaps and highlighting broad themes, the next phase of our research was to speak to stakeholders to understand more about the impact that gaps have on their

---

[15] Centre for Public Data, 'List of civil society organisations for inclusion' (2022)
[16] Centre for Public Data, 'Thematic findings from sources: potential data gaps' (2022)
[17] For all PQ code and methodology, see Centre for Public Data, Jupyter notebook published on github.com, 'Unanswered Parliamentary Questions' (2022)
[18] Centre for Public Data, 'Ministry of Justice written questions' (2022)
[19] Centre for Public Data, 'Data gaps in the justice system' (2022)

work, and to map the underlying data that is recorded, to establish which data gaps it should be possible to fill without excessive amounts of effort.

We used the following factors to prioritise which data gaps to study further:

1. A range of stakeholders expressed a clear desire for better data on the topic.
2. Without the data, stakeholders could not answer a significant question about the operation of the CJS (i.e. an area with significant impact on people or operations), and substitute data could not be obtained elsewhere.
3. Based on our technical understanding of how data is stored and aggregated, it would not be unduly expensive, time-consuming or technically difficult for producers to provide data of adequate quality to fill this data gap.

Working on point 3 involved using our technical knowledge to understand the underlying data held by different CJS bodies. For some areas we were able to piece together information from publicly available material, such as technical guides published alongside official statistics.[20] But for the most part we obtained this understanding through discussions with CJS data teams and statistics producers (where the teams involved would agree to speak to us), Freedom of Information (FOI) requests, talking to system users and CJS staff, and by applying our broader knowledge of how data and reporting systems operate across Government.

Applying what we had learned about the architecture of CJS data, we homed in on a handful of topics to understand why these data gaps existed, and how they could be solved. We did further research about the issue and spoke to expert stakeholders in each area to check our understanding of the need for the data and the systems involved. This research became the focus of our project, and our findings and recommendations are summarised in the rest of this document.

## 3.2 Limitations

Our methodology has the following limitations:

- **Restricted to criminal justice in England & Wales.** Because of time and resource constraints, our research is limited to the criminal justice system in England & Wales.
- **Only covers certain stakeholders.** These sources represent only the concerns of certain types of stakeholders (particularly those who have engaged via Parliament), although we assume they are broadly representative of justice system users more broadly. Time and resource constraints meant that important user groups could not be included in our initial study of needs, including gaps highlighted by journalists and academics, and we were unable to interview defendants, victims and witnesses with first-hand experience of the CJS.

---

[20] For example, Ministry of Justice, 'A Technical Guide to Criminal Justice Statistics Quarterly (CJSQ)' (2023)

- **Gaps may since have been filled.** In our initial list of data gaps identified by stakeholders, we did not research whether the data gaps had since been filled. Our research on topics we explored in greater detail (included in this document) is as up-to-date as possible at the time of writing, but time constraints meant it was not possible to do this for every potential data gap identified in our early research.
- **Limited access to underlying data.** We were somewhat restricted in our understanding of CJS data architecture. The MoJ and HMCTS were unwilling or unable to give us full technical details of the data attributes recorded in administrative court data systems (like Common Platform), and reporting databases (like One Performance Truth). As a result, some of our information on the data recorded by courts is based on the secondary accounts of those who have used such systems. We will be grateful to receive any corrections to our findings.
- **Limited understanding of data quality.** Without limited access to data and staff, even where we knew a particular attribute was recorded in administrative systems, we could not assess the quality of the data, e.g. its completeness, consistency or correctness. However, we could infer from an attribute's presence in MoJ's reporting databases that its quality was considered adequate for internal reporting. And useful reporting can still be generated from limited data, as long as the limitations of the data are clearly understood and explained.

Despite these limitations, we believe our methodology was effective. By analysing a high volume of stakeholder needs, we were able to cover a broad range of issues and it was straightforward to identify recurring themes. When engaging with stakeholders, we received positive feedback on our selection of themes.

# 4. Data gaps in the criminal justice system

In this section, we summarise our findings in four major areas where there is a clear need for better data, as expressed by stakeholders, and where data gaps appear feasible to fill:

1. **Remand and bail**, including prisoners on remand and the use of conditional bail.
2. **Court operations**, including legal representation, backlog data and the effect of interventions.
3. **Sentencing**, including court-level sentencing data, and sentencing for 'flagged' offences like domestic violence and hate crime.
4. **Low-level offences**, including anti-social behaviour and out-of-court disposals.

We discuss why stakeholders have asked for better data, the difference this data would make, and our specific suggestions for change. We include only a summary of each issue here, but reference more detailed briefings we have published on each topic.

This is not and is not intended to be a comprehensive list of data gaps in the CJS. Instead, it is a list of priority areas we identified using the methodology above.

## 4.1 Remand and bail

The first broad area we studied related to the experiences of people who are waiting for a court hearing or trial, and are either remanded in custody or released on bail.

### Custodial remand

The remand population is currently the highest for 50 years, placing additional pressures on overcrowded prisons.[21] But stakeholders are concerned that the lack of data has left them "grappling in the dark to understand what is going on".[22] In particular, basic questions about why defendants are held in custody, or how long they are held for, cannot be answered.

A defendant can be refused bail on the grounds set out by the Bail Act 1976.[23] Courts are required to record the legal grounds on which bail is refused; this is recorded in Common Platform, the administrative data system used by courts. But the MoJ does not publish aggregate data on the use of legal grounds, despite requests from stakeholders who say it would help to understand fluctuations in why remand is used and to understand consistency across the country.[24]

Meanwhile, data on how long people are held in remand has so far only been shared through FOI requests and PQs. In 2022, an FOI from the charity Fair Trials found that 3,879 people (27% of the total remand population) were being held beyond the custodial time limit of 6 months.[25] Multiple stakeholders, including the HM Chief Inspector of Prisons, have called for the MoJ to make this information routinely available.[26]

In January 2023, the Justice Committee urged the Government to improve remand data collection and public data on remand. The Committee recommended that:

> *More data needs to be collected and published on remanded defendants, particularly in relation to the reasons for refusing bail, the length of time people are spending on remand as well as demographic information, including vulnerabilities and protected characteristics amongst the remand population, to increase transparency and improve the information available to decision-makers.*[27]

---

[21] Ministry of Justice, 'Offender management statistics quarterly: July to September 2022' (2023)

[22] House of Commons Justice Committee, 'Oral evidence: The role of adult custodial remand in the criminal justice system, HC 264' (2022) [Q113]

[23] GOV.UK, Bail Act 1976

[24] Tom Smith, Written evidence to Justice Committee (2022)

[25] Fair Trials, 'England and Wales: FOI reveals almost 1,800 people in pre-trial detention for over a year' (2022)

[26] Justice Committee 'Oral evidence: The role of adult custodial remand in the criminal justice system' (2022), Q184; Prison Reform Trust, Written evidence to Justice Committee (2022); Fair Trials, Written evidence to Justice Committee (2022); INQUEST, Written evidence to Justice Committee, (2022); Criminal Justice Alliance, Written evidence to Justice Committee (2022); JUSTICE, Written evidence to Justice Committee (2022)

[27] House of Commons Justice Committee, 'The role of adult custodial remand in the criminal justice system' (2023)

Based on our understanding of the underlying data held, we recommend that:

1. The MoJ should publish statistics on the average length of time defendants are held in custodial remand, plus the number of defendants held beyond 6 months, 1 year, 2 years, and 3+ years. HMPPS's prisoner management system, Prison-NOMIS, records information on remand prisoners, including dates, so it should be straightforward for the MoJ to incorporate this information into their regular extract of the database going forward, broken down by offence type and demographic characteristics (if data quality allows).
2. The MoJ should publish statistics on the reasons why courts refuse bail. Our understanding is that Common Platform stores data on why courts hold defendants in custody, using the grounds set out in the Bail Act 1976. The MoJ could extract this data for publication via the Court Proceedings Database (CPD, the reporting database from which it generates many of its statistical publications). This data should also be broken down by demographic characteristics and offence type.

Our full briefing provides more details.[28]

## Conditional bail

Conditional bail allows courts to impose conditions on defendants who pose a risk of reoffending, harming witnesses or absconding, without holding the defendant in custody. It is an increasingly important aspect of our criminal justice system, with experts suggesting conditions on bail are fast becoming the norm.

However, a lack of data on conditional bail means we know little about how it is used, or whether it is an effective alternative to custodial remand or unconditional bail. There is no official data on how many people receive conditional bail; what types of conditions are imposed and why; the demographic make-up of the conditional bail population; the length of time conditions last; and whether conditions affect rates of offending while on bail.

Numerous researchers have noted the lack of data on conditions, which restrict the liberties of people who have not been convicted of a crime.[29] Professor Anthea Hucklesby warns that the "use of conditions is completely unregulated" and that courts can "attach whatever conditions they want for however long".[30]

Again, our understanding is that some underlying databases do record whether conditions have been attached to bail, and if so what type. We recommend:

---

[28] Centre for Public Data, Custodial remand: data gaps and how to fill them (2022)
[29] Tom Smith, 'Written evidence to the Justice Committee' (2022); Mandeep K. Dhami and Yannick N. van den Brink, 'A Multi-disciplinary and Comparative Approach to Evaluating Pre-trial Detention Decisions: Towards Evidence-Based Reform' (2022); Transform Justice, 'Presumed innocent but behind bars – is remand overused in England and Wales?' (2018)
[30] House of Commons Justice Committee, 'The role of adult custodial remand in the criminal justice system' (2023)

1. The MoJ's bail statistics should distinguish between conditional and unconditional bail. The MoJ could extract this data from the CPD, and combine it with existing demographic data. This would make clear how many people receive conditional bail, as well as a population breakdown.

2. The MoJ should explore generating experimental statistics on offences committed whilst on conditional bail. The MoJ already receives a data extract from the Police National Computer which includes arrests for defendants while on bail. But this data does not include whether the defendant was on conditional bail and, if so, which conditions were attached. We suggest that either the PNC be extended to record information about bail conditions, or the MoJ investigate the possibility of matching PNC and court records to generate this data.

Our full briefing provides more details.[31]

# 4.2 Sentencing

The second area where we found stakeholders consistently highlighting gaps was sentencing - in particular, the lack of sentencing data for individual courts, and sentencing for offences that are not specified in legislation, such as domestic abuse and hate crimes.

## Court-level sentencing data

In 2017, the independent Lammy Review recommended (our emphasis):

> *The Open Justice initiative should be extended and updated so that it is possible to view sentences for individual offences at **individual courts**, broken down by demographic characteristics including gender and ethnicity[32]*

The Government accepted this recommendation[33], and in 2019 announced that all actions relating to the recommendation had been completed.[34]

However, by our assessment, the Government has not completed implementing this recommendation. The Ministry of Justice's sentencing dataset, Criminal Justice Statistics Quarterly (CJSQ), currently allows for sentencing data to be broken down only by Police Force Areas, not by individual courts.[35]

In February 2023, the MoJ were pressed by the Justice Select Committee over when court-level sentencing data would be made available.[36] The MoJ's response clarified there

---

[31] Centre for Public Data, 'Conditional bail in England and Wales' (2023)
[32] David Lammy, 'The Lammy Review' (2017) [Recommendation 12]
[33] Ministry of Justice, 'Government Response to the Lammy Review on the treatment of, and outcomes for, Black, Asian and Minority Ethnic individuals in the Criminal Justice System' (2017)
[34] UK Parliament, 'Treatment of, and Outcomes for, Black, Asian and Minority Ethnic Individuals in the Criminal Justice System Independent Review', (2019)
[35] Ministry of Justice, 'Criminal Justice System statistics quarterly: December 2021' (2022)
[36] Ministry of Justice, 'Evidence to Justice Committee' (2023)

were no technical barriers to providing the data, but made no commitments to publish it routinely.

Stakeholders have suggested this data would help to understand the scale of sentencing disparities across England and Wales, improve institutional accountability, and help the Sentencing Council to meet its statutory obligation to promote awareness of local sentencing practices.[37] We recommend:

1. The MoJ should start publishing sentencing information at the individual court level by using anonymised numerical identifiers for both magistrates' courts and the Crown Courts. This data should be included in existing sentencing publications (the Criminal Justice Statistics Quarterly publication) which already include breakdowns by age, sex and ethnicity (suppressing disclosive results if required). CJSQ currently aggregates sentencing data at Police Force Area level, and we believe it would be relatively straightforward for the MoJ to provide this data at individual court level. The Sentencing Council should use these statistics to analyse, report on and promote awareness of sentencing practices across the country.

Our written evidence provides more details.[38]

## Sentencing data on offences not specified in legislation

Agencies across the criminal justice system use 'crime flags' to improve the recording and data collection of certain crimes that are not specified in legislation, such as hate crimes or domestic abuse. For example, when the police record an assault incident in a domestic setting, the officer will record both the specific offence (e.g assault with injury) as well as flag the incident as related to domestic abuse if appropriate.

The Home Office and the CPS then use their internal flagging systems to publish useful data about these cases, such as the number of Islamophobic incidents reported,[39] or the conviction rates for domestic abuse cases.[40]

The MoJ has its own method of crime flagging, known as 'case markers'.[41] However, it does not use these flags as a reporting tool in its statistics. This means we know very little about sentencing practices for domestic abuse, hate crimes, and other types of offences. Crime flags help the police, Home Office and CPS publish key data but without data from the MoJ on how such cases develop into convictions and sentences, only half the story is being told.

---

[37] Jay Gormley, Tom O'Malley, Julian Roberts, Cassia Spohn and Cyrus Tata, 'Assessing Methodological Approaches to Sentencing Data & Analysis - Report 2: USA, England and Wales, and Scotland' (2022); David Lammy, 'The Lammy Review' (2017)
[38] Centre for Public Data, 'Written evidence to the Justice Committee'' (2022)
[39] Home Office, 'Hate crime, England and Wales, 2020 to 2021' (2021) [Bulletin table 4]
[40] Crown Prosecution Service, 'CPS data summary Quarter 2 2022-2023' (2023)
[41] Ministry of Justice, 'Evidence to Justice Committee' (2023)

There is considerable concern from MPs and other stakeholders about the lack of data on sentencing practices for these offences.[42] In 2015, the CEO of anti-domestic-abuse charity Respect told the APPG on Domestic and Sexual Violence:

> *The system isn't accountable so we don't know if the sentencing guidelines are being properly implemented or not and we can't measure the effectiveness of this sentencing… There is a lot of evidence at police and CPS level but then it just stops.*

The APPG's report commented that "the lack of data on sentencing outcomes was a significant concern to the panel of the APPG Inquiry".[43]

We recommend:

1. The MoJ should use its current system of crime flagging (or 'case markers') to publish experimental sentencing statistics for offences not specified in legislation, such as domestic abuse and hate crimes. We further recommend that Common Platform case markers be extended to include online crimes, as well as sub-flags for religious hate crimes to include the perceived religion of the victim, given the interest from stakeholders in these offences.
2. The case markers in Common Platform should be extended to include online crimes, as well as sub-flags for religious hate crimes to include the perceived religion of the victim, given the interest from stakeholders in these offences.

Our full briefing on this topic is available on request.

# 4.3 Court operations

A third area focused further on data gaps on court activities, examining data gaps around legal representation in magistrates' courts, and the data needed to better understand the court backlog.

## Legal representation in magistrates' courts

There is a clear consensus among stakeholders that unrepresented defendants have a negative impact on court proceedings. Evidence suggests that defendants representing themselves experience harsher justice outcomes, find it harder to participate in proceedings, slow down court operations and have a more negative overall experience in court.[44]

---

[42] TheyWorkForYou, 'Written Questions - Fraud: Telephones' (2022); TheyWorkForYou, 'Written Question - Religious Hatred: Prosecution' (2020); Penelope Gibbs, 'Domestic violence prosecutions clumsy and costly' (2018); House of Commons Justice Committee, 'Draft Sentencing Council guidelines on intimidatory offences and domestic abuse' (2017)
[43] All-Party Parliamentary Group on Domestic and Sexual Violence, 'Women's Access to Justice: from reporting to sentencing' (2015)
[44] Charlotte Rebekah Walker, 'A Study Examining the Experiences of Unrepresented Defendants in the Criminal Courts' (2021);  Sentencing Academy, 'Defendants' Understanding of Sentencing'

But we know very little about the scale and experiences of unrepresented defendants in magistrates' courts. MPs have asked multiple questions on the topic, but been told that data is "not centrally held".[45] It is suggested that there are hundreds of thousands of people representing themselves each year, but estimates range from 13%-30% of the total defendant population.[46] We also do not know which groups of defendants are more likely to represent themselves, by demographic characteristics, location or offence type.

The charity Transform Justice argues that the "lack of data means unrepresented defendants in the magistrates' courts are invisible in policy terms… there is a pressing need to collect and collate this data to ensure that the scale of the situation is better understood".[47]

We recommend that:

1. The MoJ should publish data on the number of unrepresented defendants appearing in magistrates' courts in England and Wales, broken down by demographic characteristics, location and offence type. Currently, the Criminal Court Statistics Quarterly (CCSQ), published by HMCTS, derives operational data from the One Performance Truth (OPT) database, which extracts administrative data from Common Platform. Common Platform records a range of information on defendants, including the legal representation status of defendants. It should therefore be relatively straightforward for HMCTS to extract this information from Common Platform via OPT. This data could then be routinely published as part of CCSQ.

Our full briefing provides more details.[48]

## The courts backlog

Over 400,000 cases are waiting to be heard in criminal courts in England and Wales, with victims waiting longer for justice than any time on record.[49] Poor data undermines efforts to tackle the backlog, making it difficult to understand the composition of the case backlog, and how effective interventions have been in tackling it.

We have identified a range of areas where better data would provide better insight into the backlog. This includes breaking down case timeliness data by individual courts and specific

---

(2021); Transform Justice, 'Justice denied? The experience of unrepresented defendants in the criminal courts' (2016)

[45] UK Parliament, 'Legal Representation' (2021); UK Parliament, 'Legal Representation: Wales' (2019); UK Parliament, 'Legal Representation' (2018); UK Parliament, 'Legal Representation' (2017)

[46] Magistrates Association, 'Written evidence to the Justice Committee' (2020); Transform Justice, 'Justice denied? The experience of unrepresented defendants in the criminal courts' (2016); Lucy Charlotte Welsh, 'Magistrates, Managerialism and Marginalisation: Neoliberalism and Access to Justice in East Kent' (2016); WhatDoTheyKnow, 'Freedom of Information Request'; Dr Vicky Kemp, 'Transforming legal aid: access to criminal defence services' (2010)

[47] Transform Justice, 'Justice denied? The experience of unrepresented defendants in the criminal courts' (2016)

[48] Centre for Public Data, 'Unrepresented defendants in the magistrates' courts' (2023)

[49] Institute for Government, 'Performance Tracker 2022' (2022)

offences, so it is possible to see where cases are taking longer to progress and for which type of cases. Other areas include being able to break down the magistrates' court backlog by offence type, knowing the reasons why trials are vacated, and publishing the pleas of defendants for cases in the Crown Court backlog to better understand the complexity of cases waiting to be heard.

Stakeholders are also concerned the MoJ is not doing enough to set out the evidence for how court reforms are impacting the backlog. The Legal Education Foundation argues that if better data is not collected, "the Ministry will remain unable to accurately assess which policies are working to reduce the backlog. Lack of data will continue to undermine the ability of the department to understand what works, design effective policies and deliver value for money for the taxpayer".[50]

We recommend:

1. The MoJ should publish the following data to provide a more comprehensive account of the nature of the court backlog:
    a. Timeliness data should be broken down by individual courts and specific offence, to help stakeholders identify which courts are taking the longest to progress which types of case, and direct resources and support accordingly.
    b. Data for receipts, disposals and outstanding cases in the magistrates' courts should be broken down by offence type, to understand the type of cases waiting to be heard.
    c. Data on the reasons why trials are vacated so that stakeholders can understand the typical reasons why trials might not go ahead as planned, and mitigate against such reasons going forward.
    d. Data on the pleas of defendants currently in the Crown Court backlog, to understand the proportion of cases that will result in more complex, time-consuming jury trials.
2. The MoJ should set out for review how it intends to measure and report the outcomes for its upcoming evaluations of the Nightingale courts programme (extra courts that opened during the Covid pandemic), and publish robust data on the use of remote hearings. We also recommend that plans for reporting are built into the design of future measures to tackle the backlog. This means specifying details of the metrics to be reported in the initial business case or equivalent.

Our full briefing provides more details.[51]

## 4.4 Low-level offences

The final area we highlight relates to low-level offences that are largely resolved outside the formal courts system, typically dealt with by the police, local authorities or other local bodies.

---

[50] Legal Education Foundation, 'Written evidence to the Public Accounts Committee' (2021)
[51] Centre for Public Data, 'Data gaps on the court backlog in England and Wales' (2023)

## Anti-social behaviour

Anti-social behaviour (ASB) is a matter of intense political and public concern. Two-fifths of the public report experiencing or witnessing ASB on an annual basis, and political parties have endorsed strengthening powers to sanction offenders.[52]

However, there is no centrally published data on how each of the six available powers to tackle ASB are used. Researchers have warned that we therefore know little about how often reported incidents of ASB result in sanctions, and if so, which sanctions are used.[53] We also don't know how frequently some types of sanctions are breached, or whether ASB powers are used more for some demographic groups or areas than others.

The paucity of data on ASB powers has been criticised by justice stakeholders.[54] MPs have repeatedly asked for data but have received little information.[55] Resolve, an organisation advocating for effective ASB remedies, told us that "[w]e are calling for all authorities who use the powers to report – what that will do will highlight their use, effectiveness and also areas where they are not used well".

There is also no national data, and patchy local data, on the use of the Community Trigger, a mechanism where ASB victims can request that local service providers review the handling of their cases. Local bodies are legally required to publish data on the use of the Community Trigger, but stakeholders have found that many bodies do not meet this obligation.[56]

We recommend:

1. The MoJ and Home Office should work together to explore collecting and publishing better data on the response to ASB. This should include:
   a. How frequently each of the six ASB powers are being issued. To do this, the Home Office should require police forces to routinely report data on the use of ASB powers to the Home Office Data Hub, supported by recording guidance. This data should then be disaggregated by demographic characteristics such as age, sex, ethnicity and location. To get a complete picture of how ASB powers are issued, data would also need to be collected from Councils and Housing Associations. In the short-term, however,

---

[52] Office for National Statistics, 'Crime in England and Wales: year ending March 2020' (2020)

[53] See e.g.

[54] For example: Transform Justice, 'Should we imprison people for being a pain in the neck?' (2023); The Bureau of Investigative Journalism, 'Sent to jail for feeding the pigeons: the broken system of antisocial behaviour laws' (2022); Civil Justice Council, 'Anti-social behaviour and the Civil Courts' (2020); Centre for Crime and Justice Studies, 'Anti-social behaviour powers and young adults' (2018); HMICFRS, 'PEEL: Police effectiveness 2017' (2018)

[55] The only power where a PQ has successfully been answered is for criminal behaviour orders - UK Parliament, 'Criminal Behaviour Orders' (2022). Examples of unsuccessful questions include: UK Parliament, 'Anti-social Behaviour' (2023); UK Parliament, 'Anti-social Behaviour' (2022); UK Parliament, 'Anti-social Behaviour: Social Rented Housing' (2021)

[56] ASB Help, 'The Community Trigger: Where We Are Today' (2019)

        collecting police force data would be a helpful way to identify trends in usage.

    b. How frequently civil injunctions are breached, making it easier to assess how the power is being implemented and enforced

    c. How often ASB victims are asking for case reviews ('Community Trigger'). As recommended by ASB Help, the Home Office should consider appointing an officer to take national responsibility for the Community Trigger, ensuring that all relevant local bodies are fulfilling their statutory duty of publishing annual statistics on usage.

Our full briefing provides more details.[57]

## Out-of-court disposals

Over 200,000 out-of-court disposals (OOCDs)[58] are handed out annually by the police, to deal with low-level offenders.[59] Advocates suggest that OOCDs improve efficiency and effectiveness in the CJS, giving police forces a simple and proportionate way to deal with minor crimes, without recourse to courts or prisons.

However, stakeholders want greater clarity on how OOCDs are being implemented throughout the country.[60] In 2022, the Magistrates' Association concluded that OOCDs' "popularity has led to a rapid and largely uncontrolled expansion of their use that has not been accompanied by sufficient checks and balances".[61]

In particular, there is little national data on how diversionary schemes are used, and who receives them. Stakeholders also want to see a more meaningful evaluation of recent reforms to OOCDs, which the Government estimates will cost the criminal justice system an extra £15.6 million a year to administer, despite the available evidence suggesting the system is unlikely to have a major impact on reoffending.[62]

We recommend that:

1. The MoJ and Home Office should publish more comprehensive national data on the use of diversionary schemes. In particular:
    a. From 2023, police forces are mandated to submit data to the Home Office on Outcome 22, an outcome on the PNC that records information on the use of diversionary schemes. The Home Office should commit to using the new data as the basis for future routine and detailed publications on diversionary schemes nationwide. This could be combined with data collected on

---

[57] Centre for Public Data, 'Data gaps on anti-social behaviour in England & Wales' (2023)

[58] We use 'out-of-court disposals' as an umbrella term to include the two formal disposals (diversionary cautions and community cautions) as well as informal disposals (a range of diversionary schemes which are not specified in legislation, including community resolutions).

[59] Ministry of Justice, 'Criminal Justice Statistics Quarterly: June 2022' (2022)

[60] Co-op, 'Breaking the Cycle' (2021); Crest Advisory, 'Making the criminal justice system work better: how to improve out-of-court disposals and diversion schemes' (2022)

[61] Magistrates Association, 'Out of court disposals: Fit for purpose or in need of reform' (2022)

[62] House of Commons Library, 'Police, Crime, Sentencing and Courts Bill: Part 6, Cautions' (2021)

Outcome 8 (community resolutions) to create a complete and well-rounded picture of which informal OOCDs police forces are using. This data should include the number of people who have received informal OOCDs, the type of scheme offered and demographic information.

b. All current and future data published on informal OOCDs should include essential demographic information on the offender, including their age and sex, and that the Home Office explores how to improve the quality of data held on ethnicity.

2. The MoJ and Home Office should publish better data on and evaluations of the effectiveness of OOCDs. This includes incorporating the new OOCD into current MoJ reoffending statistics, as well as setting out a more detailed monitoring and evaluation plan, establishing criteria against which it plans to evaluate the new regime, and with details of the data it will collect to perform this evaluation.

Our full briefing provides more details.[63]

# 5. Conclusions and wider recommendations

The focus of this project was to research specific data gaps in our criminal justice system and provide technical recommendations for how they could be closed. Section 4 above lists these gaps.

During the course of our work, we also reached broader conclusions. We will not repeat the findings of others about the improvements needed to CJS data systems as a whole, although we agree with them.[64] However, we can add three recommendations specifically relating to our focus on data gaps.

## 5.1 Data gaps in the CJS really matter

Throughout this project, we have seen how important clear and credible public data is to the work of CJS stakeholders, both inside and outside the official system. We found many examples of experts and practitioners whose work was restricted by data gaps, highlighted in our research outputs and briefings, with clear examples of the harm this caused. Take conditional bail - without knowing what impact conditions have on outcomes, magistrates are largely operating in the dark. Our research has aggregated evidence on many data gaps.

**We recommend that the next iteration of HMCTS's data strategy needs to prioritise the problem of data gaps, using a similar approach to that described in this report to identify where they exist and where they can be fixed.**

---

[63] Centre for Public Data, 'Out of court disposals' (2023)
[64] The Legal Education Foundation, 'Digital Justice: HMCTS data strategy and delivering access to justice' (2019); CREST Advisory, 'Joining up Justice with Real World Solutions' (2022)

## 5.2 These gaps are fixable

The gaps we identify in this report are largely gaps in statistical reporting, not data collection, and thus should not be technically difficult to fix. We have not focussed on gaps where the solution would be to collect new data; instead, many of the gaps we have identified could be filled by aggregating information already in administrative systems.

The MoJ's Data First and Justice Data Lab programmes show that the MoJ has the technical capacity to aggregate and share high-quality data, though for highly restricted, self-selected use cases and audiences[65]; the question now is whether the MoJ is willing to use its existing data to meet the needs expressed by stakeholders more broadly.

The UK statistical system allows for the development of 'experimental statistics' - statistics that are being tested and developed, or that offer partial coverage.[66] We suggest the publication of experimental statistics would be a suitable way to tackle gaps for which there is a clear ongoing demand; in other cases, one-off releases of management information might be more suitable.

**We recommend MoJ investigate which of these data gaps can be addressed via the publication of experimental statistics based on existing administrative data.**

## 5.3 MoJ has a formal duty to address these gaps

The MoJ has a formal duty to address these data gaps. Our research often found calls from multiple stakeholders for the same information, but in many cases limited response from the MoJ (and in some cases, like legal representation, clear reluctance to engage).

As producers of official statistics, the MoJ has formal duties under the UK's Code of Practice for Statistics to ensure its statistics are "relevant" to users, and to consider whether to produce new statistics to fill "identified information gaps". Under the Code, the MoJ is also required to give users feedback about whether their needs can and cannot be met, being transparent about the reasons for its decisions.[67]

**We recommend the MoJ consider the data gaps identified in this report and associated briefings in the light of its duties under the Code of Practice for Statistics, and provide transparent information on whether stakeholders' needs for better data can be met.**

---

[65] GOV.UK, 'Ministry of Justice: Data First' (2023); 'Accessing the Justice Data Lab service' (2023)
[66] Office for National Statistics, 'Guide to experimental statistics' (2023)
[67] Office for Statistics Regulation, 'Code of Practice for Statistics - V1: Relevance to users', sections V1.3-V1.5

# 6. About us and acknowledgements

The Centre for Public Data is a non-partisan, non-profit research and advocacy organisation that works to improve the quality of UK public data. We have a particular interest in data gaps - areas where a lack of publicly available data or statistics means that questions of significant public interest cannot be answered.

This report was written by Gideon Leibowitz and Anna Powell-Smith. We will be grateful to receive comments and corrections: please contact us at contact@centreforpublicdata.org.

Our work on criminal justice data is funded by the Justice Lab, an initiative of the Legal Education Foundation, as part of their ongoing programme of research and advocacy to improve the quality and availability of justice system data. We are very grateful for their support.

We are also very grateful to the many people and institutions who shared their expertise with us during this project - their knowledge and guidance was invaluable. Any errors are our own. In particular, we would like to thank the following people, listed alphabetically:

- Dr Jonathan Bild (Sentencing Academy)
- Phil Bowen (Centre for Justice Innovation)
- Rebecca Bryant (Resolve)
- Martin Brand (HMCTS)
- Nick Davies (Institute for Government)
- Griff Ferris (Fair Trials)
- Penelope Gibbs (Transform Justice)
- Anne Gronan (Resolve)
- Professor Anthea Hucklesby (University of Birmingham)
- Paul Jackson (HMCTS)
- Stephanie Needleman (JUSTICE)
- Dr Jose Pina-Sánchez (Sentencing Academy)
- Professor Julian Roberts (Sentencing Academy)
- Job de Roij (Office for Statistics Regulation)
- Lucy Slade (Centre for Justice Innovation)
- Dr Thomas Smith (University of the West of England)
- Emma Snell (JUSTICE)
- Nicola Webb (HMCTS)
- Court staff who preferred to remain anonymous

# Appendix 1: Full list of recommendations

1. The MoJ should publish statistics on the average length of time defendants are held in custodial remand, plus the number of defendants held beyond 6 months, 1 year, 2 years, and 3+ years. HMPSS' prisoner management system, Prison-NOMIS, records information on remand prisoners, including their dates, so it should be straightforward for the MoJ to incorporate this information into their regular extract of the database going forward, broken down by demographic characteristics and offence type.

2. The MoJ should publish statistics on the reasons why courts refuse bail. Our understanding is that Common Platform already stores data on why courts hold defendants in custody, using the grounds set out in the Bail Act 1976. The MoJ could extract this data for publication via the Court Proceedings Database. This data should also be broken down by demographic characteristics and offence type.

3. The MoJ's bail statistics should distinguish between conditional and unconditional bail. The MoJ could extract this data from CPD, and combine it with existing demographic data. This would make clear how many people receive conditional bail, as well as a population breakdown.

4. The MoJ should explore generating experimental statistics on offences committed whilst on conditional bail. The MoJ already receives a data extract from the Police National Computer on the number of arrests for defendants whilst on bail. But this data does not include whether the defendant was on conditional bail and, if so, which conditions were attached. We suggest that either the PNC be extended to record information about bail conditions and include this as part of the MoJ data extract, or the MoJ investigate the possibility of matching PNC and court records to understand the effectiveness of conditions in mitigating offending.

5. The MoJ should start publishing sentencing information at the individual court level by using anonymised numerical identifiers for both magistrates' courts and the Crown Courts. This data should be included in existing sentencing publications (the Criminal Justice Statistics Quarterly publication) which already include breakdowns by age, sex and ethnicity (suppressing disclosive results if required). CJSQ currently aggregates sentencing data up to Police Force Area level, and we believe it would be relatively straightforward for the MoJ to provide this data at individual court level. The Sentencing Council should use these statistics to analyse, report on and promote awareness of sentencing practices across the country.

6. The MoJ should use its current system of crime flagging (or 'case markers') to publish experimental sentencing statistics for offences not specified in legislation, such as domestic abuse and hate crimes. We further recommend that Common Platform case markers be extended to include online crimes, as well as sub-flags for

religious hate crimes to include the perceived religion of the victim, given the interest from stakeholders in these offences.

7. The case markers in Common Platform should be extended to include online crimes, as well as sub-flags for religious hate crimes to include the perceived religion of the victim, given the interest from stakeholders in these offences.

8. The MoJ should publish data on the number of unrepresented defendants appearing in magistrates' courts in England and Wales, broken down by demographic characteristics, location and offence type. Currently, the Criminal Court Statistics Quarterly (CCSQ), published by HMCTS, derives operational data from the One Performance Truth (OPT) database, which extracts administrative data from Common Platform. Common Platform records a range of information on defendants, including the legal representation status of defendants. It should therefore be relatively straightforward for HMCTS to extract this information from Common Platform via OPT. This data could then be routinely published as part of CCSQ.

9. The MoJ should publish the following data to provide a more comprehensive account of the nature of the court backlog:
   a. Timeliness data should be broken down by individual courts to help stakeholders identify which courts are taking the longest to progress cases, and direct resources and support accordingly.
   b. Timeliness data should be separately broken down by specific offences to help stakeholders identify delays affecting particularly serious or complex offences, or case types of particular interest.
   c. Data for receipts, disposals and outstanding cases in the magistrates' courts should be broken down by offence type, to understand the type of cases waiting to be heard.
   d. Data on the reasons why trials are vacated so that stakeholders can understand the typical reasons why trials might not go ahead as planned, and mitigate against such reasons going forward.
   e. Data on the pleas of defendants currently in the Crown Court backlog, to understand the proportion of cases that will result in more complex, time-consuming jury trials.

10. The MoJ should set out for review how it intends to measure and report the outcomes for its upcoming evaluations of the Nightingale courts programme (extra courts that opened during the Covid pandemic), and publish robust data on the use of remote hearings. We also recommend that plans for reporting are built into the design of future measures to tackle the backlog. This means specifying details of the metrics to be reported in the initial business case or equivalent.

11. The MoJ and Home Office should work together to explore collecting and publishing better data on the response to ASB. This should include:

a. How frequently each of the six ASB powers are being issued. To do this, the Home Office should require police forces to routinely report data on the use of ASB powers to the Home Office Data Hub, supported by recording guidance. This data should then be disaggregated by demographic characteristics such as age, sex, ethnicity and location. To get a complete picture of how ASB powers are issued, data would also need to be collected from Councils and Housing Associations. In the short-term, however, collecting police force data would be a helpful way to identify trends in usage.

b. How frequently civil injunctions are breached, making it easier to assess how the power is being implemented and enforced

c. How often ASB victims are asking for case reviews ('Community Trigger'). As recommended by ASB Help, the Home Office should consider appointing an officer to take national responsibility for the Community Trigger, ensuring that all relevant local bodies are fulfilling their statutory duty of publishing annual statistics on usage.

12. We recommend that the next iteration of HMCTS's data strategy needs to prioritise the problem of data gaps, using a similar approach to that described in this report to identify where they exist and where they can be fixed.

13. We recommend MoJ investigate which of these data gaps can be addressed via the publication of experimental statistics based on existing administrative data.

14. We recommend the MoJ consider the data gaps identified in this report and associated briefings in the light of its duties under the Code of Practice for Statistics, and provide transparent information on whether stakeholders' needs for better data can be met.

# Appendix 2: Full list of research outputs

**Initial report outputs**

- 'Thematic findings from CJS sources: potential data gaps' (2022)
- 'List of civil society organisations for inclusion' (2022)
- 'Interim report: data gaps in the justice system' (2022)

**Full topic-specific briefings**

- 'Custodial remand: data gaps and how to fill them' (2023)
- 'Conditional bail in England and Wales' (2023)
- 'Unrepresented defendants in the magistrates' courts' (2023)
- 'Data gaps on the court backlog in England and Wales' (2023)
- 'Data gaps on anti-social behaviour in England & Wales' (2023)
- 'Out of court disposals' (2023)