

# Internal Evaluation of the RFSDi Harm Score and the Integrated Offender Management Model

Ethics Committee (09 November 2022)

## Contents

---

Contents.....	2
1 Executive Summary.....	3
1.1 Recommendations .....	4
2 Introduction .....	5
2.1 Usage of the dashboard .....	7
3 Methodology and Limitations.....	8
4 Participant Responses and Recommendations.....	9
4.1 Broad understanding of the two elements of the model .....	9
4.2 Training .....	9
4.2.1 Training Recommendations: .....	10
4.3 Does the tool assist with decision making? .....	11
4.3.1 Frequency.....	11
4.3.2 Decision making .....	11
4.3.3 Trust .....	12
4.3.4 Use of the predictive IOM element.....	13
4.3.5 User Experience .....	14
4.3.6 Benefits of the tool .....	15
4.3.7 Recommendations for further developments .....	15
5 Wider consultation with LOMUs.....	16
6 Update on model performance .....	16
7 Response to findings of independent evaluation .....	18
8 Appendix: Glossary of Terms .....	24

## 1 Executive Summary

---

The Ethics Committee recommended that an evaluation of the beta testing phase of the RFSDi/IOM model be undertaken.

This paper is the internal evaluation undertaken by the Data Analytics Lab, which complements the external process evaluation undertaken by Alex Babuta. It seeks to provide the Committee with an understanding of how the RFSDi/IOM model has been used by two Local Offender Management Units (LOMUs) during beta testing and whether or not they perceive that it has added value to their processes, along with recommendations for further development.

The RFSDi/IOM model has been tested by two LOMUs since October 2021. Of the 15 people who have access to the dashboard, five inspectors and sergeants were interviewed for this evaluation (Section 4). In addition, a focus group has taken place with LOMU inspectors from across the Force, who do not yet have access, to understand their level of interest and expectations of the tool (Section 5). The report also includes an update on the performance of the model (Section 6) and a response to the findings and recommendations of the external evaluation (Section 7).

This evaluation found that Offender Managers (OMs) are comfortable using the RFSDi element of the model. This is the harm score assigned to each offender. OMs are familiar with the concept of harm scores and the rationale that underpins them. OMs are less confident in their understanding of the predictive IOM element of the model which estimates the probability of a defendant becoming high harm. However, this component is used by Inspectors, to recommend offenders who could be managed in the lower 'flex' cohort by neighbourhood officers. This suggests that although not widely used, the predictive IOM element of the model is being used as intended to direct the appropriate resource towards supporting those who are more likely to start committing more harmful offences.

There was broadly positive feedback about the initial training and ongoing support provided, although there could be more focus on the predictive IOM element. How this would be scaled up and managed if the model is made available to all LOMUs across the Force needs some consideration.

There are some key areas for development which would improve the end-user experience. The ability to apply more filters to the information in the dashboard would facilitate using it as a triage tool when OMs are researching lists of nominals for different cohorts and tiers of management. This could be achieved by adding the ability to filter records by offence type at a more granular level; to see whether a nominal is currently being actively managed and by adding a hyper-link to the source system CONNECT where further research on nominals is carried out prior to any decisions being made

The most significant development requested by end-users is to include data relating to those who are currently under investigation for an offence ('suspects') but for which they have not been charged. This information is routinely considered as part of existing OM decision-making processes, because it helps to identify opportunities for interventions aimed at preventing further criminality. The overwhelming feedback is that the inclusion of 'suspect data' in the RFSDi score would make these decisions more robust. It is not suggested that suspect data would be included in the predictive IOM element.

Overall, this internal evaluation found that OMs see the RFSDi/IOM as a useful addition to their decision-making processes. Depending on the type of cohort managed, use of the tool ranges from it being the initial starting point for determining the level of management for a particular offender; to being a tool which helps to sense check decisions made using other risk assessment tools employed by other agencies or departments and to inform discussions with them.

## 1.1 Recommendations

- For the Offender Management leadership team to plan how training should be delivered in the future, both as it is rolled out to LOMU teams across the force, but also to individual new joiners. Consideration needs to be given to learning styles and the fact that as a live system training examples quickly become outdated. Hands-on learning on the live system in a computer training room (for a group) or in the workplace (for an individual) would be optimal, but this needs to be balanced against capacity.
- Consideration needs to be given as to how on-going support will be provided once the dashboard is made available to all LOMUs across the Force, so as not to overwhelm the Visualisation Developer.
- For the DAL to produce a written guide explaining the methodology in terms which can be understood by end-users, in particular the predictive IOM element. This should be placed on the front of the dashboard as has been done with the Crime Seasonality Planner and could be added to the IOM intranet pages.
- For any future training to take account of the lessons learnt from the LOMUs involved in the beta testing phase.
- For the Ethics Committee to consider whether the RFSDi element of the model could be expanded to include suspects who have not yet been charged with an offence to enable OMs to apply the same methodology across the cohort they manage.
- For the DAL to remove or hide the names of those currently in prison until three months before their release date and to remove names of deceased nominals (where known).
- For the DAL to design in the capability to define cohorts by crime type, such as filtering down to a list of robbery offenders
- For the DAL to include a hyper-link to CONNECT to make it quicker to do additional research in the source system (which can be audited).
- For the DAL to add the status of the Proactive Management Plan (PMP) from CONNECT into the nominal table. This would provide a quick view of which nominals in the list are currently being managed.
- For the DAL to add offending locations to the map view.
- For the DAL to consider adding a network chart to illustrate nominals' associations.

## 2 Introduction

---

The RFSDi/IOM model developed by the West Midlands Police (WMP) Data Analytics Lab (DAL) supports the Force's strategic priority to *Act with Precision* in order to *target the most problematic offenders*. This priority is underpinned by a key performance indicator for Local Policing to achieve a reduction in harm from the Integrated Offender Management Cohort.<sup>1</sup>

The purpose of the model is to assist Offender Managers (OM) to identify the most harmful offenders who should be prioritised for proactive offender management, including those offenders who have not yet become 'high harm' but who should also be considered for offender management as a preventative measure.

Both elements of the model only include data relating to those people who have been charged with an offence<sup>2</sup>, referred to as 'offenders' in police terminology. 'Suspects' are those people who are or have been under investigation for an offence but not charged.<sup>3</sup> Data relating to suspects is not included in the model.

The RFSDi element of the model calculates a harm score for each offender based on their offending history and other information agreed as pertinent with Subject Matter Experts (SMEs)<sup>4</sup>. The score uses **R**ecency, **F**requency, **S**everity (based on the Cambridge Crime Harm Index), information about **D**rugs misuse and their **i**ntelligence footprint.

The IOM element is a statistical model providing a risk assessment which estimates the probability of a defendant becoming high harm, who has not yet reached that threshold. This is the predictive part of the model.

The details of how each element is calculated is available from previous technical papers presented to the Committee.<sup>5</sup>

The output of the model is visualised on a Business Insight (Qlik) dashboard as an interactive tool for Local Offender Management Units (LOMUs) to understand the harm presented by individual offenders they manage and to predict the probability that they will transition from committing low/middle levels of harm to committing high levels of harm.

The RFSDi/IOM model was rolled out for beta-testing in October 2021 to LOMUs in Dudley and Birmingham West Neighbourhood Policing Units (NPU).

*Figure 1* shows part of the overview page of the dashboard. Currently, there are a total of 97,643 offenders in Birmingham West (BW) and Dudley (DY) NPUs.

---

<sup>1</sup> West Midlands Police Strategic Priorities 2022/23

<sup>2</sup> Home Office [Police Recorded Crime and Outcomes: Open data tables user guide 2016](#) Defendants include outcomes 1 - Charge/Summons; 2 - Caution Youths; 3 - Caution adults; 4 - Taken into Consideration; 6 - Penalty Notices for Disorder; 7 - Cannabis/Khat warning and 8 - Community Resolution.

<sup>3</sup> Investigation outcomes such as 9/10/21 - Not in the public interest; 11/12/13 - Prosecution prevented and 14/15/16 - Evidential difficulties may have an identified. See previous link for full details of outcomes.

<sup>4</sup> See previous papers submitted to the [Committee](#), in particular 24/07/2019 (Item 3 IOM Model Answers)

<sup>5</sup> [Ethics Committee Reports and Minutes - West Midlands Police & Crime Commissioner \(westmidlands-pcc.gov.uk\)](#) See papers for Apr 2019; Jul 2019 and Jan 2020.

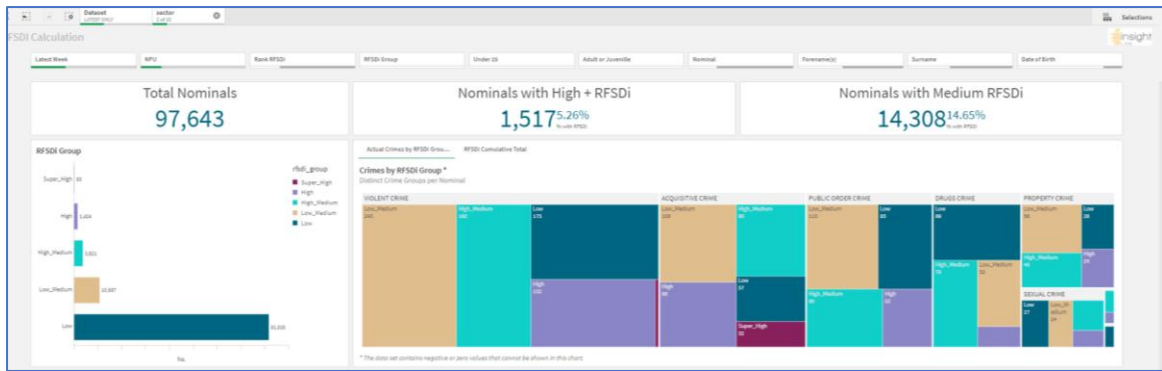


Figure 1: Overview page of the RFSi dashboard for Birmingham West and Dudley LOMUs as at 20/10/2022

Whilst this illustrates that just over 80% of offenders have an RFSi score placing them in the low harm group, in terms of numbers, there are still over 5,000 offenders whose score places them in medium-high, high or super-high groups who need to be considered by OMs. These include 437 offenders who have committed violent crimes, 212 who have committed acquisitive crimes and 31 who have committed sexual offences. Across BW and DY there are 39 OM constables responsible for these offenders (excluding sex offender managers). This illustrates the requirement for a harm score to assist with the prioritisation process.

Additionally, OMs can view a table which displays a harm score for each individual offender and assigns them a rank on the basis of their score. OMs are responsible for particular cohorts of offenders (for example based on a geographical area, a crime category, such as acquisitive crime or an age group such as under 25s). By applying filters to the dashboard, the OMs can create a relevant cohort list and view the most harmful individuals they are responsible for managing.

Figure 2 shows the IOM page of the Business Insight dashboard.

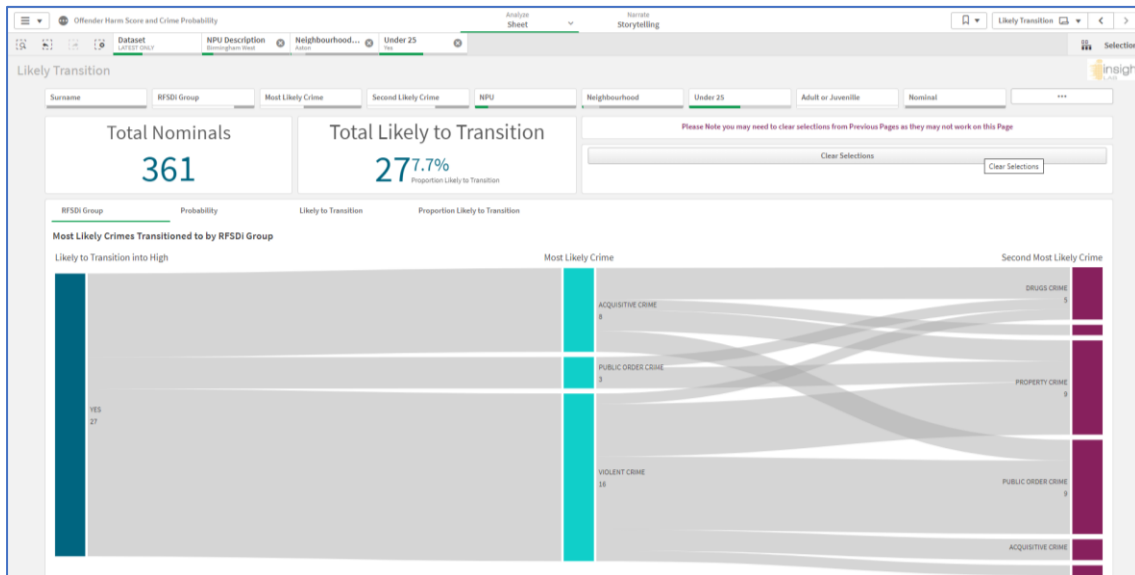


Figure 2: IOM dashboard with selections to display u25 offenders in Aston BW

Figure 2 shows a selection based on nominals aged under 25 in the Aston ward of Birmingham West. This shows that of the 361 nominals in this cohort, 27 are predicted to have a higher probability of transitioning from being a low harm to a high harm offender. It also shows the type of offending that they are most likely to commit. This dashboard can also be filtered down to the level of the individual.

## 2.1 Usage of the dashboard

There are 15 OMs who have been given access to the dashboard during the beta testing phase. This includes inspectors, sergeants and constables in BW and DY LOMUs.

Access Used by Month		
Month	App Stream	App
	Access Type	
	Total	Analyzer
Oct 2022	6	6
Sep 2022	12	12
Aug 2022	7	7
Jul 2022	7	7
Jun 2022	9	9
May 2022	7	7

Figure 3: RFSdi/IOM dashboard usage by LOMUs

Figure 3 shows that around seven OMs have used the dashboard each month between May and August, increasing to 12 users in September (data for October is up to 19<sup>th</sup>). The number of sessions per user is around four to five times a month which suggests it is being used on a weekly basis.

Some users have provided feedback during the beta testing phase and amendments have been made to the visualisation to make it more user-friendly. These include:

- Splitting out names by forename and surname to make searching easier
- A weekly, rather than monthly refresh of the data
- Changing the day of the refresh to fit better with meeting schedules

No amendments have been made to the underlying models, other than the re-coding required due to the move to CONNECT and to a new data platform for the DAL.

### 3 Methodology and Limitations

---

This internal evaluation was conducted by a member of the DAL, although not one involved in the development of the model.

The Offender Management command team were requested to provide names of ten OMs testing the system who could be contacted for the evaluation. Due to operational constraints only eight names were provided and of these five people were available for interview. These took place via skype, between 27 September and 7 October.

<b>NPU</b>	<b>Rank</b>
DY LOMU	Inspector
DY LOMU	Sergeant
DY LOMU	Sergeant
BW LOMU	Inspector
Pan-Birmingham Youth Offender Management	Sergeant

One of the sergeants joined IOM in March 2022 and received an input on the model in April, therefore their answers reflect the perspective of a 'new joiner' to this area of business.

A constable was in the list of names provided for interview, but did not respond to invitations to participate, so only the views of inspectors and sergeants are included in this evaluation.

A semi-structured interview approach was used and was able to take advantage of the fact that a draft of the external evaluation had been received by this stage.

It is acknowledged that the selection of interviewees was not representative and potentially skewed by manager choice and operational constraints. In particular, it has not been possible to reflect the views of constables. Additionally, there may be a different dynamic for end users talking to a member of the development team compared to an independent researcher.

Questions covered four areas:

1. Broad understanding of the two elements of the model
2. Quality of the training received
3. The extent to which the output is used to make decisions about the management of offenders
4. End-user observations for how the model and its output could be improved

The remainder of this evaluation covers these four areas with some recommendations for taking the project forward.



## 4 Participant Responses and Recommendations

---

### 4.1 Broad understanding of the two elements of the model

Participants had a good understanding of what the RFSDi element provides and its uses:

*Highlights nominals for selections who will cause most harm/demand to communities. Helps to map their offending.*

*To inform decisions around our cohorts and see which tier of management they are suitable for – neighbourhood teams or OMs (higher) etc. Includes recency, frequency, severity, drugs and intelligence. The league table of names. There is science behind the selections but it still needs human element and research in CONNECT.*

However, understanding and use of the predictive IOM element of the model was limited. On the whole, OMs had awareness of the predictive element, but lacked confidence in their understanding of it and therefore did not really use it. However, some answers later in the interview demonstrated that inspectors do use of this part of the tool to make recommendations.

*IOM is not exploited as much, I know it exists. For me – not used very much but I know it is there. I use what I know. This is not as well explained – how it is built – not so confident in it.*

It was noted by one participant that the RFSDi element was similar to a previous excel based tool, developed before the DAL existed. This was known as the 'RFG' and assigned domestic abuse offenders and suspects a score based on the **R**ecency, **F**requency and **G**ravity of their offending. Familiarity with this historic tool seems to have given confidence using the RFSDi element and the loss of this tool with the implementation of CONNECT was referred to by many participants who see RFSDi as the potential replacement.

### 4.2 Training

All participants stated that they had received training by the DAL Visualisation Developer who built the dashboard, who has delivered 29 training sessions since February 2019. For some, this was via skype during a Covid lockdown period, for others it was in-person at their work location. One of the on-line training sessions was recorded and some participants referred to this being a useful tool for them to refer back to. One participant who has only recently become an OM stated they had received one-to-one training, but that it was perhaps given too soon after they started alongside a number of other briefings.

Participants stated they had access to the system as soon as they had been trained, so they were able to apply their learning straight away.

Participants stated that whilst there was some explanation of the methodology underpinning the model, most of the training focused on how to navigate the dashboard. Most stated that there was very little explanation of the predictive IOM methodology and that this is one reason why they felt less confident in using it. However, a couple of participants acknowledged that the training sessions were very interactive and to some extent directed by the learners themselves.

*Yes [the training provided an understanding of the methodology] – in basic terms for the rankings, but not so much the predictive part – but we directed the training – we wanted to know about the nuts and bolts.*

The majority of participants felt the training had taught them how to navigate the dashboard and they then continued to self-teach as they used the tool in their daily business.

*Yes, the training gave me the confidence and the basics and then I could use it to find out what it can do for me.*

One participant referenced the fact that a number of new systems have been brought into the Force in a short period (ControlWorks in 2020 and more significantly CONNECT in 2021):

*There is a general training need around Qlik, speaking to other colleagues. As a new supervisor I need to understand all the tools my team are using. A lot of new systems were brought into Force at the same time.*

There is likely to be a diffusion of benefits as more people become familiar with using Qlik since the navigation tools are the same across the platform. Most of the participants use at least one other Qlik dashboard.

Ideas for improving the training offer included some additional reference materials, such as some bite-size thematic videos showing how to do common tasks, although most felt that they learnt best 'by doing'. One participant pointed out that the recording of one training session was quite difficult to learn from some months later because the dashboard is continually refreshed and so the scores and rankings change each week.

*I have looked at the older recorded session, but it was out of date, so I hadn't heard of any of the nominals. So, no.1 then is now in prison and not causing us any harm, so it was confusing.*

Some had contacted the Visualisation Developer at some point if they had an issue with the dashboard or to provide some feedback. All stated they had a positive response and that issues were resolved, additional on-the-spot training provided or feedback acted on where possible. This level of support has been possible given the small number of people accessing the dashboard.

#### 4.2.1 Training Recommendations:

- For the Offender Management leadership team to plan how training should be delivered in the future, both as it is rolled out to LOMU teams across the force, but also to individual new joiners. Consideration needs to be given to learning styles and the fact that as a live system training examples quickly become outdated. Hands-on learning on the live system in a computer training room (for a group) or in the workplace (for an individual) would be optimal, but this needs to be balanced against capacity.
- Consideration needs to be given as to how on-going support will be provided once the dashboard is made available to all LOMUs across the Force, so as not to overwhelm the Visualisation Developer.
- For the DAL to produce a written guide explaining the methodology in terms which can be understood by end-users, in particular the predictive IOM element. This should be placed on the front of the dashboard as has been done with the Crime Seasonality Planner and could be added to the IOM intranet pages.
- For any future training to take account of the lessons learnt from the LOMUs involved in the beta testing phase.

## 4.3 Does the tool assist with decision making?

### 4.3.1 Frequency

Some participants use the dashboard weekly, others monthly. This tends to depend on meeting schedules, whether for internal tasking meetings such as Local Tactical Delivery Boards (LTDB) on NPU; or partnership meetings with Probation or Youth Offending Services (YOS).

*I also use it once a month to extract data from the u25 violence cohort to present at our local tasking.*

*I have a MAPPa meeting second Wed of every month – so I use the system in the week before that.*

Participants felt that the weekly refresh rate was adequate and commented that the refresh day had been changed to fit better with meeting schedules.

### 4.3.2 Decision making

The extent to which the dashboard helps OMs to make decisions about the offenders they manage varies, dependent on the type of cohort they manage.

For example, the risk posed by sexual and violent offenders managed under multi-agency public protection arrangements (MAPPa) has already been identified and the appropriate level of offender management has been determined via these processes.

Similarly, under the Integrated Offender Management (IOM) Strategy<sup>6</sup> offenders leaving prison on licence or serving community orders are assessed using the Probation 'Offender Group Reconviction Score' (OGRS). Those that score over 75 per cent are placed in the 'fixed' cohort and those scoring lower are placed in the 'flex' cohort. Therefore, LOMUs involved in joint management do not use RFSDi to make these decisions, although they do report using it to 'sense check' the OGRS output.<sup>7</sup> The Strategy does allow for individuals to be deselected from the fixed cohort if they are individually assessed as low risk of reoffending. Thus, RFSDi offers an opportunity for OMs to influence decisions made in partnership with Probation with a robust methodology to support their recommendations. Principle 1 of the national strategy states they should have an equal say in such decisions.<sup>8</sup>

*We use the national model for Probation – their OGRS score which provides a % risk harm. This is used by multi-agency team for the management of SAC [serious acquisitive crime] offenders – those with a score of over 76% should be considered for management by IOM. We cross-reference the OGRS score with RFSDi/IOM.*

Some OMs involved in the management of youth offenders also explained that their cohort is to some extent pre-determined because they support children who are under a court order:

*They have either been flagged to us by NHT [neighbourhood team] or they are already charged and therefore open to YOS. So as part of all our checks I would look on RFSDi to see where they sit – it's one component.*

However, for other cohorts, OMs have found RFSDi does help to identify individuals for consideration and helps to determine which tier of support they should receive.

<sup>6</sup> HM Government, [Neighbourhood Crime IOM strategy \(publishing.service.gov.uk\)](https://publishing.service.gov.uk) Dec 2020

<sup>7</sup> Nationally, some forces are using the Office for National Statistics (ONS) Crime Severity Score (CSS) as part of the decision-making process to supplement the Probation OGRS. West Midlands' RFSDi score incorporates the Cambridge Crime Harm score (the 'severity' element of the RFSDi) which performs the same role as the CSS. Discussion of the similarities and differences between these two scores have been presented in previous papers to the Committee.

<sup>8</sup> HM Government, [Neighbourhood Crime IOM strategy \(publishing.service.gov.uk\)](https://publishing.service.gov.uk) Dec 2020. p.13

One example shows how the tool helps to identify those who have been charged but not yet entered the Criminal Justice System, thus providing an opportunity to mitigate the delays in the court system:

*But another nominal was third highest on RFSDi but was not being managed – so the tool highlighted them. They were not being managed because they had not yet been through the court process.*

Without the use of RFSDi, this high-harm offender may not have been managed for a period of time and their offending may have continued.

Both inspectors use the RFSDi to collate a list of offenders who do not meet the criteria to be managed by LOMUs but who would be suitable for management by neighbourhood teams at the next tier down. These lists are fed into LTDB meetings for the consideration of local officers.

*But, for the flex cohort and local non-statutory offenders we would use RFSDi. Not sure if it flags up new names because the list is passed to the NHT – so I don't know if the list we pass them are names that they are already aware of or not.*

This supports the national and Force IOM model which puts a focus on offenders being managed and supported by local officers.

Some OMs were able to give examples where the RFSDi had alerted them to an offender who was not on their radar:

*A recent nominal who was ranked as no.3 – we were not aware of him, but research showed that his offending was not in Dudley that's why we didn't know about him. So, the tool helped us to identify the risk of a person who lived in our area.*

This was a burglary offender who had committed 15 car key burglaries on other NPU's, but whose home address was in Dudley and therefore the responsibility of DY LOMU to manage.

All participants talked about consulting other systems and professionals before making decisions about which tier of management an offender might be assigned to. This includes doing additional research in CONNECT to understand their offending history in greater detail. CONNECT was also used to check any intelligence reports about their associations:

*The gravity of offending really impacts the score – but the actual risk of them re-offending and what they've done in the past is not necessarily the same. We know their risk due to their associations ie. gangs, which is not reflected in the RFSDi data.*

This comment reflects the fact that there could be more training around the predictive IOM element which aims to assist with this question. The predictive model draws from a network of all known offenders to identify associations and calculates their centrality within any networks, however, this is not currently displayed the dashboard.

In addition, OMs check the Probation OGRS and YOS data systems and discuss nominals with their colleagues in these teams before any decisions are made.

#### 4.3.3 Trust

A repeated frustration for OMs is the fact that the model only includes data relating to offenders who have been charged and only the offences for which they have been charged contribute to the RFSDi score. This means the score for prolific offenders does not reflect further offences for which they are currently being investigated. Since charge rates are low, this limits the number of people included in the data set. For example, the current year to date 'positive outcome rate' for residential burglary is 7.49% and for robbery 8.96%. This means that OMs also need to consider other sources of information when considering which tier of management is appropriate. The role of OMs is to prevent further offending, so

managing suspects is as much part of their remit as managing those charged with an offence. This is particularly relevant for OMs managing Domestic Abuse (DA) offenders and young offenders:

*It doesn't work with DA offenders – especially if there are a lot of non-crimes recorded – they wouldn't score highly in RFSDi.*

Currently, OMs have to use a different tool to make decisions about DA offenders. If RFSDi also included suspect data, this would incorporate the requirements of DA OMs.

*In youth arena, for example [one suspect had] 18 outstanding investigations on CONNECT but not showing in RFSDi. So local knowledge tells us there is a very risky young person, but RFSDi doesn't show them.*

*There is a gap around suspects, especially when managing juveniles and not wanting to press charges and criminalise them if possible.*

*No suspects is a concern. As OMs we get asked about crime trends and who should we be looking to get managed. To answer this we have to look at multiple data sources: CONNECT, custody, overnight note, probation – all this could be collapsed into one place.*

*I would cross reference with CONNECT and with Youth Offender Management to see if there is anyone up and coming – because they will not be in the CJ system [ie. RFSDi won't show them because] suspects not charged – they may be suitable for OM.*

*I cross reference with CONNECT investigation notes, what's their offending, what have they been charged with and what have they not been charged with?*

If OMs do not supplement the RFSDi with additional information about offences for which a nominal might be a suspect, they can be professionally embarrassed in meetings with other departments and this risks the level of trust officers have in the tool.

However, the majority of participants did state that they trusted the output of the RFSDi element of the model, with the caveat that it does not reflect the 'whole offending picture'.

*It has got a lot better recently. The first 6/7 months it didn't seem to be up to date. Last week the list produced 70 kids who were high/very high – I would say we do manage half of that list and we know about most of the others.*

*It isn't giving the whole story because we know it doesn't include suspects, but I do trust what it says with that caveat. I find when I go to the FCID meetings they are coming up with the same a list of names [ie. without access to the app].*

These comments reflect the reassurance provided by the RFSDi that names under consideration for management are supported by 'science'.

#### 4.3.4 Use of the predictive IOM element

As highlighted earlier, OMs felt less confident about using the predictive IOM element of the model. For some, it would not be appropriate for their cohort:

*Don't really use the IOM part because the decision to manage an offender is based around their current, substantive behaviour. As OM we manage those who are already CJ statutory offenders in the 'fixed' cohort (since the IOM refresh).*

*I don't use the predictive part because it doesn't fit my cohort – Youth, DA, MAPP. There are different mechanisms for choosing which offenders to manage. Youth via partner meetings; DA was done via RFG (I miss this); MAPP is done by risk management – risk to the community & victim through ViSOR.*

However, it has been used by the Inspectors to provide suggestions of nominals that the neighbourhood teams could consider managing at the next tier down to prevent them moving into higher harm offending:

*But sometimes we use the IOM predictive list to give to the NHT so that they can do preventative work – to add to the ‘flex’ cohort. So NHT will go through the predictive list we give them to see if they match those on their list.*

This use of the IOM element of the model is in line with the stated objectives of the project. The limited number of people using the predictive IOM element is in part reflective of job role and should not be taken as an indicator that it is not useful.

#### 4.3.5 User Experience

Most of the participants are now comfortable navigating the Qlik dashboard. However, all participants stated that using the dashboard can be time consuming. The limited information it provides means that it is not as useful as it could be to conduct an initial triage of offenders.

A repeated frustration is the inability to filter down to a more granular level for offence types. Currently, the dashboard can be filtered by geography down to ward level, RFSDi group (such as low or high harm); aged under 25 or not; adult or juvenile. It can also be filtered by very broad crime categories such as acquisitive crime, drugs crime or sexual crime. However, users familiar with other Qlik dashboards are used to being able to define the crimes they want to view at a lower level:

*Its relatively straightforward but not precise enough – it gives you SAC, but you can’t drill down to burglary – it’s not detailed enough and you have to go into CONNECT*

*Yes – but will always double check, especially to look for offence type, for example to see if they are kids carrying knives.*

For other OMs this does not present as much of an issue:

*I look at high/very high category for BW and filter by age and geography. For me the lack of an offence type is not so much of an issue – so if they are ‘only’ shoplifting this is still important to me – why is a child shoplifting – it still flags as a risk.*

Another frustration is that the dashboard does not take account of and hide those nominals who are currently in prison or who have died.

*Not sure it does what it’s supposed to do – for example I put a list together and got bitten – the feedback was that the nominals were in prison.*

Most OMs referenced this as another caveat they need to be aware of when they are using the tool to make recommendations about which offenders should be managed at which tier.

All participants indicated that the dashboard would be a lot more useful if it gave them more information and enabled them to triage more of the list before going on to do further research in other systems. For example:

*On the table of names, I want to be able to highlight an individual and easily get to their offending history. At the moment when I see new names in my list I have to go into CONNECT to find out about them.*

*Frustrating because you have to come out and go into CONNECT to do your research – it should show you everything.*

*It’s too time consuming – I’m not confident in it so there is a lot of double checking to do.*

*It would help to know if they are already managed – what is their PMP [proactive management plan] status in CONNECT?*

*It would be good if the mapping showed where they offend as well as where they live.*

As with all Qlik dashboards, there are security and audit considerations and some information that the OMs would like to be able to access in a 'one stop shop' would not accord with the general principles applied. However, since the RFSDi/IOM dashboard was originally designed, Qlik software has been upgraded and developers have taken advantage of new features to improve the user experience. This learning could be applied to improve the user experience of the RFSDi/IOM dashboard.

#### 4.3.6 Benefits of the tool

Despite some of the frustrations described above, all participants felt there was benefit in having the tool and were interested in it being further developed:

*It is useful to have some science behind the cohorts and backs up our conversations and agreements with probation.*

*Gave a presentation at an Inspector Away Day – gave reassurance that we are looking at the same/right people – increasing trust in the tool – used regularly at LTDB.*

*Seven years ago, when I last worked in OM I was tasked with identifying our most harmful offenders who cause us the most harm. No-one could say who or why. I did a manual score of 'red, amber, green' using intel and arrests. So the RFSDi is long overdue.*

*Definitely – quick time to work out who is causing the most threat, risk & harm.*

*Overall the idea is good and with a few changes it could be really useful – such as to be able to produce a list of the Top 10 non-managed burglary offenders.*

*It hasn't really speeded up my work because my list of referrals comes to me from YOS or NHT. But it does provide a double check and helps to justify whether we say yes or no to managing the ones they suggest to us.*

#### 4.3.7 Recommendations for further developments

- For the Ethics Committee to consider whether the RFSDi element of the model could be expanded to include suspects who have not yet been charged with an offence to enable OMs to apply the same methodology across the cohort they manage.
- For the DAL to remove or hide the names of those currently in prison until three months before their release date and to remove names of deceased nominals (where known). The exclusion of such information is due to prison data not currently being available to the DAL.
- For the DAL to design in the capability to define cohorts by crime type, such as filtering down to a list of robbery offenders
- For the DAL to include a hyper-link to CONNECT to make it quicker to do additional research in the source system (which can be audited).
- For the DAL to add the status of the Proactive Management Plan (PMP) from CONNECT into the nominal table. This would provide a quick view of which nominals in the list are currently being managed.
- For the DAL to add offending locations to the map view.
- For the DAL to consider adding a network chart to illustrate nominals' associations.

## 5 Wider consultation with LOMUs

A consultation with LOMU inspectors and sergeants<sup>9</sup> took place on 26 October where the model and dashboard were presented to OMs who do not yet have access to the tool, as well as representatives from BW and DY where the beta-testing has taken place.

It was agreed that:

- There should be some level of standardisation about how the tool is applied by the LOMUs across the different NPUs, whilst allowing some flexibility for local context. This would be managed and monitored through meetings such as this one between the LOMU inspectors and the Central IOM inspector.
- If the facility to filter the dashboard by offence type is improved, then the RFSDi can be used to ensure that OMs are making decisions in line with Force priorities about which crime types to focus on. For example, in conjunction with the Crime Seasonality Planner<sup>10</sup>, they can ensure they are focusing on burglary offenders in the Autumn when this criminality tends to increase.
- The LOMUs should use the predictive IOM element to recommend offenders for the 'flex' cohort for management by local neighbourhood officers (probably 2 -3). This would focus resources on those who are identified as having a greater probability of becoming higher harm offenders, and would involve the deployment of normal neighbourhood policing activity.
- That LTDB process should be the mechanism for making and recording recommendations for the flex cohort and whether they are accepted by NHTs. The RFSDi/IOM score should be recorded and monitored via the nominal's PMP as part of the rationale for selection / deselection to specific cohorts. This is where activity to support and divert offenders should be recorded so that an assessment can be made as to the effectiveness of any intervention.

## 6 Update on model performance

The results of the model's estimated probabilities were examined over a period of 8 months from late September 2021 to April 2022.

In terms of comparing against the totality of people who were classified as high harm (essentially the target variable against which the model was originally built), the performance of the model was reasonably maintained with the averages of the main measures being (the cutoff for this model for classifying a predicted class of high harm was 0.6):

Measure	Result
Accuracy	0.92
Sensitivity	0.75
Specificity	0.93
Precision	0.38

As originally built, the results from testing were:

<sup>9</sup> All LOMUs except Solihull were represented at the meeting.

<sup>10</sup> Approved by the Committee in July 2022



Measure	Result
<b>Accuracy</b>	0.97
<b>Sensitivity</b>	0.73
<b>Specificity</b>	0.997
<b>Precision</b>	0.96

Comparing the two tables therefore shows that most of the accuracy measures remain similar to when first built except the precision which has fallen considerably.

Of course, the main interest is how accurate the model is in terms of how well it identifies those who become high harm.

Looking at the predictions from the model run in early October 2021 and comparing to the two following months leads to:

Measure	Result
<b>Accuracy</b>	0.89
<b>Sensitivity</b>	0.56
<b>Specificity</b>	0.89
<b>Precision</b>	0.005

It can be seen that the precision has fallen precipitously. However it should be borne in mind that the model essentially outputs the estimated probability of becoming high harm over a coming two year period and in the period for which the table above was developed, only 345 people became high harm. In short, it is difficult to assess the model's accuracy with so few people having transitioned (and over such a short period).

The above only examines the first three months of the model running as there was a large jump in those classed as high harm in January 2022. The reason for this is not entirely clear, however there have been changes to the system from which the data come (the numbers newly classed as high harm monthly has fallen back to average levels). For this reason we are in the process of re-building the model and early results are closer to the results originally obtained:

Measure	Result
<b>Accuracy</b>	0.95
<b>Sensitivity</b>	0.82
<b>Specificity</b>	0.98
<b>Precision</b>	0.98

## 7 Response to findings of independent evaluation

The table below provides some context for and responses to the findings and recommendations of the independent evaluation<sup>11</sup>:

Findings	
External Evaluation	WMP Comment
The research found that the practical implementation of the application is yet to result in the operational improvements for offender managers that were envisaged at the time it was rolled out for beta-testing. As a result of perceived deficiencies in the application discussed further below, PCs and Sergeants report that the RFSDi score and predictive IOM model are not routinely used as part of offender management processes.	Internal evaluation found that it is used by sergeants and inspectors to differing degrees depending on the type of cohort managed and whether there are other risk assessment tools already mandated (such as Probation OGRS).
The research highlighted a fundamental divergence in views between PCs and Sergeants on the one hand, and Inspectors on the other. Inspectors were significantly more positive and complimentary regarding the new application, reporting that the RFSDi score and IOM model are useful, and represent an improvement over the previous application (Corvus). In stark contrast, no PC or Sergeant reported finding the RFSDi score or IOM model useful, and none agreed that the application represented an improvement over Corvus.	<p>In the internal evaluation, Corvus was only mentioned by one participant. The Force has decommissioned Corvus and its functionality is replaced by CONNECT. However, CONNECT does not provide a harm score for nominals or enable the user to rank them by harm. The old 'IOM Score' provided in Corvus was known to be heavily weighted by intelligence reports.</p> <p>The development of the RFSDi was designed to overcome the known flaws in the old IOM Score and was developed after in-depth discussions with Subject Matter Experts (SMEs). See DAL response to questions raised by the Committee at meeting 03/04/2019 – further paper was provided to the <a href="#">Committee</a> on 24/07/2019 (Item 3 IOM Model Answers).</p> <p>See training recommendations from the internal review.</p>
This divergence in views is partly explained by the fact that Inspectors interviewed for the research had been more closely involved in the development of the application and are therefore likely to have a more detailed understanding of its strengths and limitations. This demonstrates the critical importance of ensuring wide consultation and engagement in the early planning stages of a new data-driven system, to ensure that end-users have had sufficient opportunity to contribute to the early development process.	Noted for further development of this project and for future projects.

<sup>11</sup> Note: this is taken from Draft external report; the final version had not been received at this stage.

<p>Three main reasons were identified for the negative feedback reported by PCs and Sergeants. The first relates to the user experience of the dashboard itself, with officers reporting that the application interface is not user friendly and is difficult to navigate.</p>	<p>This internal evaluation makes specific recommendations for improvements to the dashboard to improve the end-user experience.</p> <p>Familiarity with using Qlik dashboards in general should increase as this tool becomes used more widely across the Force.</p>
<p>The second relates to perceived deficiencies in the underlying statistical application leading to erroneous outputs, damaging officers' confidence in the overall validity of the system. PCs and Sergeants report experiencing a high number of over-classification errors, with a disproportionately large number of nominals being scored as 'high' or 'super high', resulting in an unmanageable list of individuals to review. As well as over-classification errors, officers also report a large number of 'missing nominals'; individuals who should be scored as high-risk but are not being identified by the application (either because they have not yet been charged with an offence, or because the system relies solely on data held on WMP's internal database).</p>	<p>Over-classification errors were not specifically highlighted in the internal review, however, the inability to break the cohort down further by crime type and other filters was a persistent theme and there is a recommendation to address this.</p> <p>The main reasons identified by the internal evaluation for lacking confidence in the output were the lack of suspect data (relating to the 'missing nominals') issue and the inclusion of those in prison affecting the rankings, both of which have recommendations in the report.</p> <p>The RFSDi/IOM model will only ever be based on WMP data. OMs will always need to have professional discussions with colleagues in Probation and YOS to share knowledge about individuals.</p>
<p>Third, officers report receiving insufficient training and written guidance regarding the application before being required to use it operationally.</p>	<p>There is a recommendation in this report for some written guidance and further recommendations for training.</p>
<p>If these issues were to be resolved, Inspectors emphasised three main strengths of the RFSDi harm score which they believe could provide operational benefit. The first is improved precision of targeting, with Inspectors reporting that the tool should enable officers to more precisely monitor changes in an individual's harm score and track an increase or decrease in their level of risk. The second perceived benefit is increased confidence in the risk assessment process, as the factors that contributed to each score can be identified and triangulated across other data sources, providing another source of information to support individual-level risk assessments. The third perceived benefit is the identification of 'hidden risk'; the ability to identify individuals who are not currently subject to offender management orders, but are flagged by the system as requiring further, in-depth risk assessment. It is important to note that these perceived benefits described by Inspectors were not yet reflected at the operational level in the experiences reported by PCs and Sergeants.</p>	<p>No comment required</p>

<p>It is also notable that the majority of interview and survey responses relate specifically to the RFSDi harm score, rather than the predictive IOM model. The distinction between the two components of the application was not clearly recognised by most respondents. When prompted, interviewees had relatively little insight into the predictive modelling element of the application specifically. This is concerning, as the outputs from the predictive model represent a probabilistic forecast associated with an inherent degree of uncertainty, and should therefore be interpreted in a different way to the descriptive RFSDi harm score.</p>	<p>Similar to the findings of the internal evaluation. However, the Inspectors did report using the predictive IOM element in precisely the manner envisaged.</p> <p>There are recommendations in this internal report to improve the training input around the IOM element.</p>
<p>Looking beyond the current performance of the application, the research identified three priority areas for future improvement. The first is training and guidance, with most respondents reporting that they received insufficient training and guidance on the application before being asked to use it.</p>	<p>Addressed in the recommendations in this internal evaluation.</p>
<p>The second is the inclusion of more selection criteria within the dashboard, to allow officers to filter according to particular crime types.</p>	<p>Addressed in the recommendations in this internal evaluation.</p>
<p>The third relates to integration and cross-compatibility with other data systems, most notably the Connect information management system. Several interviewees suggested that RFSDi should be integrated within Connect, to allow officers to view custody images and other intelligence for nominals who are assessed using the RFSDi dashboard.</p>	<p>Qlik and CONNECT are two distinct systems. A recommendation of this report is to include a hyperlink to take users of the dashboard to the relevant record in CONNECT which is the appropriate place for the storage of custody images and intelligence records. Qlik will not be used to store this type of information.</p> <p>These are principles which affect the development of all Qlik dashboards and which are agreed with the Professional Standards Department and Force Executive Team.</p>
<p>In conclusion, the evaluation has not established sufficient positive evidence in favour of deploying the RFSDi dashboard and accompanying IOM model for long-term operational use. The system should be subject to further, detailed evaluation research to conclusively establish its benefits and limitations before it is deployed for wider operational use. This should include the development of a detailed evaluation plan, including measurable criteria against which to assess the ongoing business case for the project and demonstrate that it is delivering its intended outcomes.</p>	<p>There is a demonstrable need for a harm score with which to assess the existing risk posed by nominals. It is considered that the recommendations in this internal evaluation be incorporated and the dashboard be made available to all LOMUs.</p>

Recommendations	
External Evaluation	WMP Comment
The force should establish a clear impact evaluation plan to measure the outcomes of the system on an ongoing basis. This should include developing a basic logic model to describe the intended outputs and outcomes of the project, as outlined in the College of Policing's Evaluation Toolkit. It is important to define measurable evaluation criteria to assess the ongoing business case for the project and demonstrate that it is delivering its intended outcomes.	Initial consultation with IOM inspectors and sergeants on 26/10/2022 started a discussion about how outcomes could be monitored.
Offender Managers who are required to pilot a new data-driven system should be encouraged at an early stage to contribute directly to the application development process. An initial survey of end-users should be distributed, requesting feedback on the limitations of existing processes, and the user interface and design requirements for any new system.	Noted. Consultation meeting with IOM inspectors and sergeants who are not involved in the beta testing to understand their requirements took place on 26/10/2022 should the model should be extended to all LOMUs.
Any future development of the dashboard should focus on surfacing 'hidden risk': the identification of nominals not currently subject to offender management orders who should be subject to more in-depth risk assessment. To avoid the risk of high-priority nominals being erroneously de-selected, individuals already subject to offender management plans should be excluded from the RFSDi harm scoring system.	Addressed in the recommendations in this internal evaluation.  By adding information about the PMP status it will be possible to quickly identify high harm nominals who are not currently being managed.  OMs are likely to still want to be able to view those they are currently managing, so that they can see any changes in their harm score over time and to be able to benchmark 'new names' against those they are familiar with.  Exclusion of known high risk nominals would skew the RFSDi score and would likely reduce trust in the results.
The research was inconclusive regarding the potential benefits offered by the predictive modelling component of the application. The IOM model should be subject to dedicated, controlled evaluation research before it is deployed operationally.	If PMPs are used to record RFSDi/IOM scores, decisions concerning selection/deselection and associated activity this could provide a starting point for such an evaluation.
IOM model predictions should be more clearly distinguished from the descriptive RFSDi harm score. A caveat should be included alongside model outputs, with the following 'health warning': <i>Prediction generated by machine learning model. Accuracy and confidence may vary depending on context. Validate alongside other data sources before taking further action.</i>	Noted – to be included when the dashboard is amended.
Any future development of the dashboard should focus on improving the front-end user experience for officers. The dashboard should incorporate a 'Feedback' section, where users can provide written feedback and submit suggestions for improvement.	Addressed in the recommendations in this internal evaluation.

	<p>Qlik is not interactive in the sense that end-users can add feedback to the system. A feedback process to be agreed with IOM SLT.</p>
<p>The most pertinent data points from other information management systems (most notably custody images) should be included within the RFSDi dashboard. Integrating the dashboard within the existing Connect system is likely to achieve this and should be a priority for any future development of the application.</p>	<p>The inclusion of custody images is not an appropriate use of Qlik software.</p> <p>It should be noted that Connect is a system developed and provided by a third party and as such information developed within a Police force cannot be incorporated into it.</p> <p>It is likely that a hyperlink will take users from the RFSDi dashboard to the nominal record in CONNECT which is the appropriate system for viewing this type of information.</p>
<p>Additional training should be delivered to all officers with access to the RFSDi dashboard and predictive IOM model. This should cover how the dashboard is intended to be used, the input variables used to calculate the RFSDi score and to build the predictive model, and an overview of the inherent limitations of the statistical techniques underpinning the system.</p>	<p>Addressed in the recommendations in this internal evaluation.</p>
<p>Written guidance should be developed for all officers with access to the RFSDi dashboard and IOM model. This guidance should include a summary of how the harm scores and statistical predictions are generated, as well as a workflow diagram of how the data-driven insights should be integrated into existing offender management processes.</p>	<p>Addressed in the recommendations in this internal evaluation.</p>



## 8 Appendix: Glossary of Terms

---

<b>WMP / Law Enforcement Terminology</b>	
BW	Birmingham West NPU
DA	Domestic Abuse
DAL	Data Analytics Lab
DY	Dudley NPU
FET	Force Executive Team
IOM	Integrated Offender Management
LOMU	Local Offender Management Unit
LTDB	Local Tactical Delivery Board
MAPPA	Multi-agency Public Protection Arrangements
NHT	Neighbourhood Team
NPU	Neighbourhood Policing Unit
OGRS	Offender Group Reconviction Score – Probation risk assessment
OM	Offender Manager
OPCC	Office of the Police and Crime Commissioner
PCC	Police and Crime Commissioner
PMP	Proactive Management Plan
RFSDi	Recency, Frequency, Severity, Drugs, intelligence
SAC	Serious Acquisitive Crime (robbery, burglary, vehicle crime)
SME	Subject Matter Expert
WMP	West Midlands Police
YOS	Youth Offending Service