



PTRS 13/14 Specification

Will Gibberd
NOMS PAG

V1-00
March 2013

| Version Control | | | | |
|-----------------|----------------------------|-----------|------------|----------------|
| Version | Status | Author | Date | Comments |
| V1-00 | Subject to Approval at NAB | I Hansson | March 2013 | 2013/14 update |
| | | | | |
| | | | | |

1 Introduction

This document sets out the specification for the Probation Trust Rating System (PTRS) for 2013/14. It is intended to provide a detailed statement of how PTRS will operate, allowing all stakeholders to understand how results are generated from the performance data supplied.

2 Summary of Changes

2.1 Aims of PTRS Development for 2013/14

PTRS is intended to provide Probation Trusts, Contract Managers and the NOMS Agency with a picture of performance built on the general principles of:

- Rounded picture of performance
- Fair and robust Assessment
- Encourage on-going improvement
- Timely & accurate
- Multiple customers
- Interactive & user friendly product
- Uses trusted data
- Streamlined production

PTRS has been developed from the 2010/11 PTRS, which was based on the 2009/10 Integrated Probation Performance Framework (IPPF). These frameworks have been updated over a number of annual cycles, with stakeholders reviewing the framework each year to provide advice and requirements for further development for the succeeding year.

PTRS 2013/14

- PTRS will now be produced only twice during the year: an indicative view at 6 months (Q2) and the usual full-year version at 12 months (Q4). Data for Q1 and Q3 will still be available on the PAG Performance Hub.
- The five metrics corresponding to OMI2 scores will be removed from PTRS for 2013/14:
 - OMI2 Assessment and Sentence Planning
 - OMI2 Risk of Harm
 - OMI2 Likelihood of Reoffending
 - OMI2 Interventions
 - OMI2 Enforcement and Compliance
- The weight corresponding to each of the OMI2 metrics will be redistributed among the remaining metrics in the domain in which they are present for PTRS 2012/13.
- The overall structure of the model will remain the same for 2013/14.
- The Probation Contract Measures & Management Information (PCMI) report will no longer be included in PTRS. The information contained in this report will continue to be available in the Performance Hub's 'All Probation Metrics' report.

- The Quality Assurance self-assessment scores will not be included in the data-driven assessment; however trusts or commissioners will be allowed to use these scores in the moderation process for 2013/14 where appropriate.

2.2 Governance and Regulation of the Performance Ratings

PTRS 13/14 is an annual performance assessment.

An interim snapshot will be produced for Q2 for the purposes of operational management; this is not subject to moderation.

The moderation process for 2013/14 is yet to be confirmed.

Through the NOMS commitment to transparency the following will be made publicly available at the end of 2013/14:

- The annual performance rating for each Trust,
- The underlying data used to produce the performance ratings,

2.3 A Significant Reduction in the Indicators Used

In 11/12 the number of metrics included in PTRS 11/12 was significantly reduced from 33 to 13. This marked a shift to a lighter touch performance framework, and a move away from input and process measurement in line with the government's commitments and changes to the Probation Trust Contracts.

To reflect the increasing focus on outcomes PAG reviewed the performance metrics with the driver tree (Annex A) to emphasise the links with the following outcomes:

- Protecting the public,
- Reducing re-offending,
- Effective delivery of the sentence of the court.

These outcomes form the domains within PTRS 13/14.

Given the reduction in the number of metrics it was decided that weightings should be introduced into the system. Further information can be found below.

HMI Probation is moving towards a more thematic approach to inspections, and has decided not to continue the OMI2 programme. Therefore, from 2013 onwards, OMI2 scores will no longer be produced. The five OMI metrics will be removed from the 2013/14 PTRS model, leaving eight metrics remaining.

3 PTRS 2013/14 Content

3.1 Structure

Background

For 2011/12 the PTRS reporting structure was aligned with the driver tree (Annex A) to focus the assessment on outcomes. This structure was reviewed and agreed by stakeholders at a workshop run by PAG.

We can see the overall outcome on the far right of the diagram. Feeding into this overall result we have the main outcomes (excluding 'Delivering value for money'):

- Protecting the public,
- Reducing re-offending,
- Effective delivery of the sentence of the court in a visible and responsive way.

A number of drivers link to each of the outcomes, for example the driver 'Victim needs, risks and rights' links to the outcome of 'Public Protection'.

Two drivers in *Figure 1* below link equally to two different outcomes. 'Quality and timeliness of offender assessments and reviews' feeds both the 'Protecting the public' and 'Reducing Reoffending' outcomes. Likewise the driver 'Effective offender engagement' links to both 'Reducing Reoffending' and 'Effective delivery of the sentence of the court in a visible and responsive way'.

This can be accounted for within the driver tree by drawing lines from each driver to both the relevant outcomes.

2013/14

Figure 2 shows the PTRS 13/14 report. The outcomes mentioned above can be considered to be the domains within PTRS and the underlying drivers in *Figure 1* are linked to a metric e.g. the driver 'Effective Offender Engagement' driver maps onto the performance metric 'Offender Feedback'.

As described above, this has resulted in performance measures contributing to two different domain ratings. This can be accounted for within the performance framework (to prevent double-counting) through the weighting assigned to each indicator. This was one of the factors considered when defining the metric weightings.

Figure 1: PTRS 2013/14 Drivers

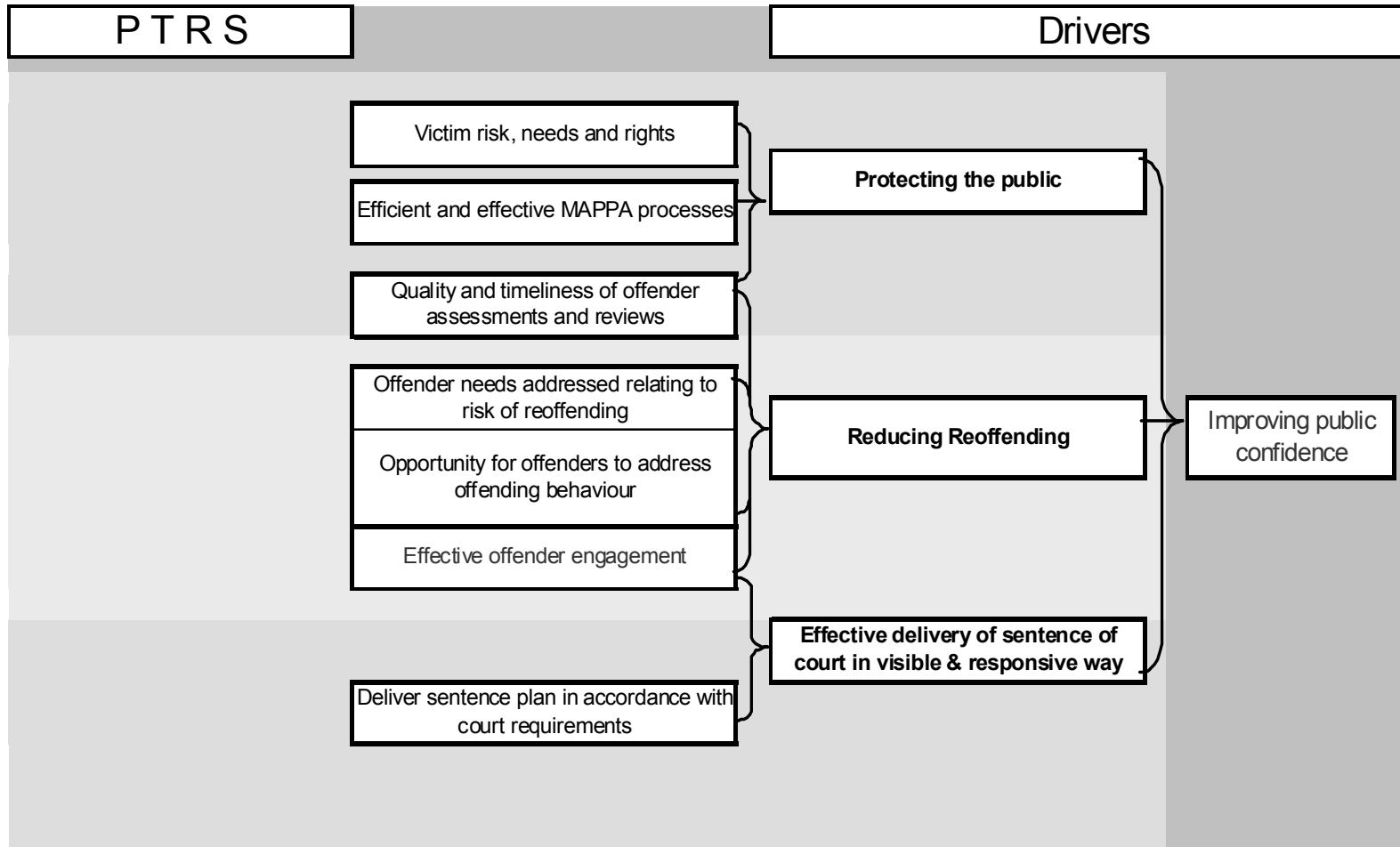
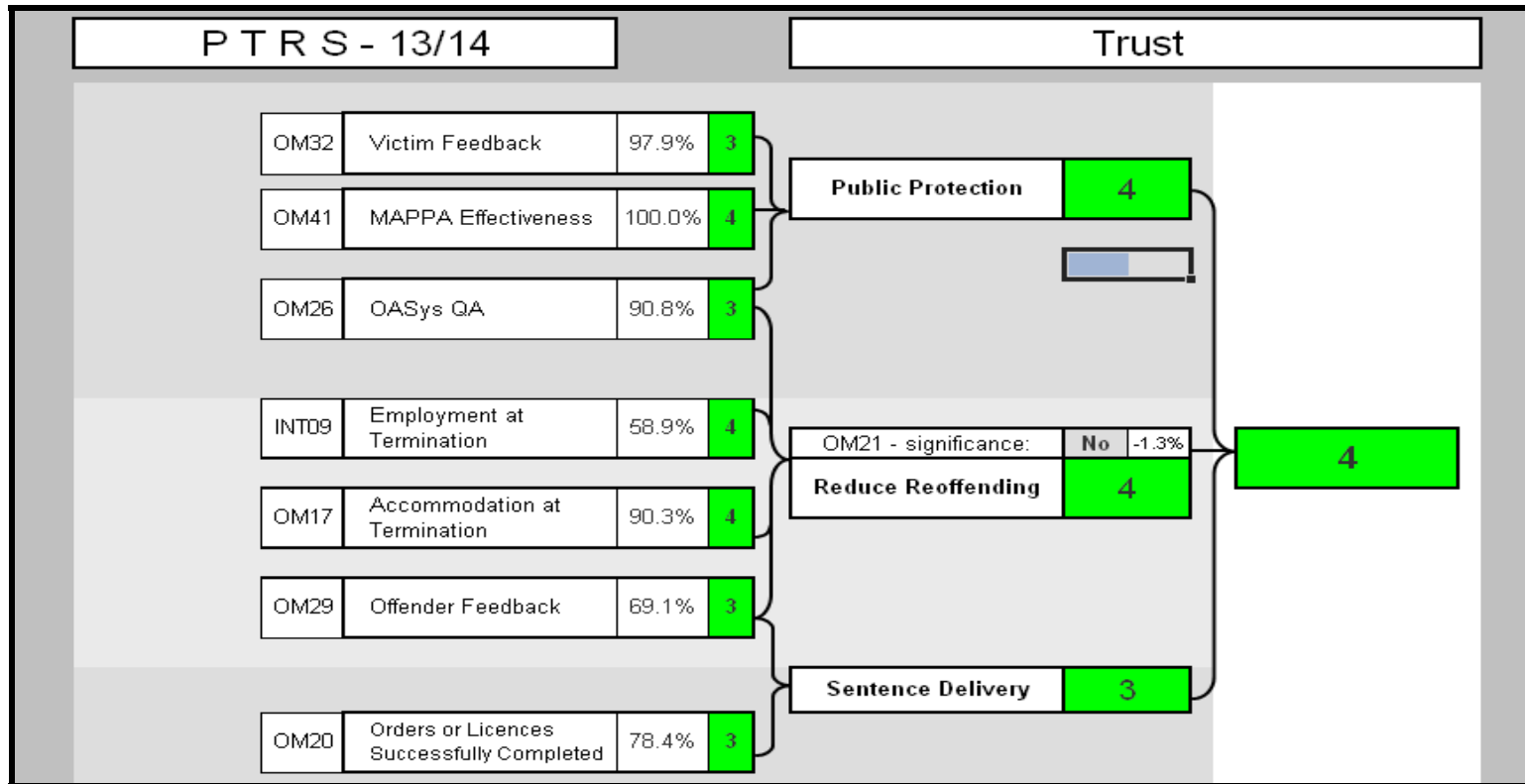


Figure 2: An example of the PTRS 13/14 report.



3.2 Ratings Methodology

For 2013/14 PTRS will retain the following four level ratings scale which will be applied to indicators, domains and overall:

| Key | |
|-------------------------|---|
| Exceptional Performance | 4 |
| Good Performance | 3 |
| Requiring Development | 2 |
| Serious Concerns | 1 |

Reports will continue to be colour-coded, with the RAG status having the same definitions as for PTRS 2012/13. PTRS will maintain the principle that hitting the target results in a minimum of a Level 3, good performance rating.

3.2.1 Indicator Ratings

The individual performance measures will be RAG rated on the scale above using performance against target and national average.

This assessment method is described in more detail below.

(i) Performance against target and national average

This methodology remains identical to that used in PTRS 12/13 and involves comparison with (1) local/national targets and (2) the national average. These two components are then combined to form the overall metric rating based on the rating grid below.

- (1) Performance against local/national targets is assessed according to:

| Band | Performance against target |
|------|----------------------------|
| 4 | Target met or exceeded |
| 3 | ≤ 2.5% below target |
| 2 | ≤ 5.0% below target |
| 1 | > 5.0% below target |

- (2) Performance against the national average is determined by ranking Trusts by their absolute performance as shown in the table below:

| Band | Rank |
|------|-------|
| 4 | 1-5 |
| 3 | 6-18 |
| 2 | 19-30 |
| 1 | 31-35 |

Under these conditions the value of the 5th ranked trust will be used as the boundary, so any trust ranked 5th or above will receive an exceptional rating if they are also achieving their target. The choice of the 5th ranked trust as the boundary will be reviewed over the course of the year.

This method aims to identify a maximum of 5 Trusts as exceptional. If any Trusts in positions 1-5 are equally ranked so that more than 5 Trusts are given a Level 4, the maximum number under 5 will be given the exceptional ratings. For example, if three Trusts were equal 5th, then only the top four would be given a Level 4 for this component.

The same general approach will be applied to the boundary between requiring development and serious concerns. The bottom 5 of the 35 Trusts will be awarded a 1 relative to national the average and the 30th ranked trust would be used as the boundary. This selection will also be reviewed over the course of the year. However, this approach is slightly different to the top 5 as all Trusts ranked equal 30th will be given a 2 for this component.

Ratings for these two components are combined using the grading grid below to produce the overall rating for this metric:

| Rating Grid | | (2) Against National Average | | | |
|--------------------|---|------------------------------|---|---|---|
| | | 4 | 3 | 2 | 1 |
| (1) Against Target | 4 | 4 | 3 | 3 | 3 |
| | 3 | 3 | 3 | 2 | 2 |
| | 2 | 3 | 2 | 2 | 1 |
| | 1 | 2 | 1 | 1 | 1 |

Trusts with a 4 against target (those achieving or exceeding their target) are guaranteed at least a Level 3. Trusts who are the best performers relative to the national average will always achieve a grading of 2 or above (not red).

Sample Based Indicators

The Victim Feedback and Offender Feedback measures will use the 'performance against target and national average' assessment method. However, these metrics are both based on samples of the true victim/offender population so confidence intervals will be used in the performance against target aspect of this assessment method. The confidence intervals will be calculated using the following standard 95% confidence interval formula:

$$Confidence\ Interval = 1.96 * \sqrt{\frac{p(1-p)}{N}},$$

Where p is the percentage figure and N is number of returns received. This formula assumes that the sample is randomly taken. The performance figure compared with the target = $p + \text{Confidence Interval}$.

The actual performance figure will be used for the comparison with the national average.

3.2.2 Indicator Weighting

Due to the significant reduction in the number of metrics in PTRS in 2011/12 the decision was made to introduce weightings into the system.

The weightings will be applied when calculating the domain ratings from the underlying metric ratings. Each domain will also be weighted, as described below.

The metric weightings take into account:

- The relative importance,
- How one metric may feed into more than one outcome/domain,
- Any limitations inherent in the design of each metric.

A workshop with key stakeholders (including representatives from Trusts, policy leads and HMI Probation) was held by PAG to determine the metric weightings. A pair wise comparison of the metrics in each domain was completed independently by the attendees and the resulting weights were discussed during the workshop. The weights/structure of PTRS were revised following the workshop and circulated for comment.

The tables on the following page show the resulting weights that each metric had within each domain for 2012/13 and the weights they have for 2013/14 in light of the removal of the metrics corresponding to the OMI2 scores.

Public Protection:

| Public Protection | | |
|----------------------|--------------|--------------|
| | PTRS 2012/13 | PTRS 2013/14 |
| OMI2 Risk of Harm | 45% | 0% |
| Victim Feedback | 25% | 45.5% |
| MAPP A Effectiveness | 20% | 36.4% |
| OASys QA | 10% | 18.2% |

Reducing Reoffending:

| Reducing Reoffending | | |
|--------------------------------|--------------|--------------|
| | PTRS 2012/13 | PTRS 2013/14 |
| OMI2 Likelihood of Reoffending | 30% | 0% |
| Employment at Termination | 25% | 41.7% |
| Accommodation at Termination | 15% | 25% |
| OASys QA | 10% | 16.7% |
| Offender Feedback | 10% | 16.7% |
| OMI2 Interventions | 10% | 0% |

Sentence Delivery:

| Sentence Delivery | | |
|--|--------------|--------------|
| | PTRS 2012/13 | PTRS 2013/14 |
| Order or Licenses Successfully Completed | 40% | 72.7% |
| Offender Feedback | 15% | 27.3% |

3.2.3 Domain Weighting

The NOMS Agency Board decided to also weight the three domains for similar reasons to those given above. The following weightings will be applied when calculating the overall rating:

| Overall | |
|----------------------|-----|
| Public Protection | 30% |
| Reducing Reoffending | 40% |
| Sentence Delivery | 30% |

The following table shows the contribution of each metric to the overall rating:

| Overall | | |
|---------|---|-------|
| 1 | Orders or Licenses Successfully Completed | 21.8% |
| 2 | Employment at Termination | 16.7% |
| 3 | Victim Feedback | 13.7% |
| 4 | Offender Feedback | 14.9% |
| 5 | OASys QA | 12.2% |
| 6 | MAPP A Effectiveness | 10.9% |
| 7 | Accommodation at Termination | 10.0% |

3.2.5 Domain Ratings

PTRS 13/14 consists of three domains, reflecting the outcomes given in section 3.1 above. Each domain will be rated on a scale of 1-4, calculated from the ratings of the underlying metrics using the same general method as in PTRS 12/13.

To produce an aggregated domain rating, each indicator within that domain is allocated a numerical value according to the rating for that indicator, as in the following table:

| Indicator Rating | Numerical Value |
|------------------|-----------------|
| 4 | 3 |
| 3 | 2 |
| 2 | 1 |
| 1 | -2 |

Using these numerical values and the weights detailed in section 3.2.2 a weighted average is calculated for each domain using the following formula (for a domain containing N measures):

$$\text{Domain Average} = \sum_n^N \text{Weight}_n * \text{Numerical Value}_n$$

(An example of this calculation is given on the following page).

This domain average is compared to the following boundaries and the corresponding rating is assigned to each domain:

| Boundary | Domain Rating |
|----------|---------------|
| > 2.25 | 4 |
| ≥ 1.5 | 3 |
| ≥ 0.75 | 2 |
| < 0.75 | 1 |

This approach means that extreme under or over performance in one indicator will not have a dominant effect on the domain or overall score.

Example Domain Rating Calculation

The example below shows how the weighted average is calculated for a fictional 'Reducing Reoffending' domain made up of the following metrics/ratings:

| Reducing Reoffending | | | |
|------------------------------|--------|--------|-------|
| Metric | Weight | Rating | Value |
| OaSys QA | 16.7% | 3 | 2 |
| Employment at Termination | 41.7% | 4 | 3 |
| Accommodation at Termination | 25.0% | 2 | 1 |
| Offender Feedback | 16.7% | 1 | -1 |

The first column contains the metric name, the second shows the weighting applied to each metric, the third column contains the fictional rating and the final column displays the corresponding numeric value as described on the previous page.

For this domain the weighted average would be calculated as follows:

$$\begin{aligned} \text{Domain Average} &= \sum_n^N \text{Weight}_n * \text{Numerical Value}_n \\ &= (16.7\% * 2) + (41.7\% * 3) + (25\% * 1) + (16.7\% * (-1)) \\ &= 1.67 \end{aligned}$$

When compared with the boundaries on the previous page this average corresponds to a Level 3, good performance rating for this domain.

Reducing Re-offending Rate

Reducing Re-offending is one of the key outcomes for probation and forms one of the domains in PTRS 13/14. PTRS 2012/13 used an automatic rule as part of the data-driven assessment and, as the local re-offending rate data will continue to be produced quarterly, the same rule will be adopted in 13/14.

The local re-offending rate produced by Justice Statistics & Analytical Services (JSAS) is published quarterly alongside a predicted re-offending rate. The actual re-offending rate is tested for statistical significance with the predicted rate and a performance rating is given; amber (Level 2) if the actual rate is statistically significantly higher than predicted and green (Level 3) if the rate is significantly lower than predicted, otherwise it is shown as grey.

Where the most recent and previous re-offending data both show a Trust to have a statistically worse re-offending rate than predicted, a Trust will only be able to score a maximum of a Level 3 rating for the Reducing Reoffending domain – a Trust will not be able to achieve a Level 4 for this domain.

A Trust may still provide evidence to the moderation panel in Q4 to show that this automatic rule is not representative of performance for the period in question.

Proven Re-offending statistics

In the recent 'Proven Re-offending' statistical bulletin one of the measures this bulletin aimed to replace was the Local Adult Re-offending rate used in PTRS.

An Outcome Measures Working Group has been established by PAG, working to replace the Local Adult Re-offending rate in PTRS with a measure derived from the Proven Re-offending bulletin at some point in the future.

3.2.6 Overall Rating

The overall ratings will be calculated from the domain ratings using the same method as in PTRS 13/14. The domains will be weighted as described in section 3.2.4 above.

Levels 1-3

The Level 1-3 overall ratings will be determined in a very similar way to the method for calculating the domain ratings as described above.

Each domain will be assigned a numerical value according to:

| Indicator Rating | Numerical Value |
|------------------|-----------------|
| 4 | 3 |
| 3 | 2 |
| 2 | 1 |
| 1 | -2 |

The numerical values (one for each domain) are averaged and compared to the following boundaries:

| Boundary | Domain Rating |
|-------------|---------------|
| ≥ 1.5 | 3 |
| ≥ 0.75 | 2 |
| < 0.75 | 1 |

An example of a Level 3 overall rating is given below.

Example Level 1-3 Overall Rating

The example below shows how a Level 1-3 overall rating is calculated for the following fictional spread of domain ratings:

| Domain | Weighting | Rating | Numeric Value |
|----------------------|-----------|--------|---------------|
| Public Protection | 30% | 2 | 1 |
| Reducing Reoffending | 40% | 3 | 2 |
| Sentence Delivery | 30% | 2 | 1 |

Taking an average of these numeric values gives the following calculation:

$$\begin{aligned} \text{Overall Average} &= \sum_n^N \text{Domain Weight}_n * \text{Numerical Value}_n \\ &= (30\% * 1) + (40\% * 2) + (30\% * 1) \\ &= \underline{1.4} \end{aligned}$$

When compared with the boundaries on the previous page this average corresponds to a Level 2, 'requiring development' rating for this Trust.

Level 4 Overall Rating

The overall Level 4 rating will be determined by direct comparison, rather than setting a bar for performance (though all Trusts considered must be at least operating at Level 3 using the method above). This will primarily be done by selecting the trusts with the highest 'overall average' calculated as given in the formula above.

PTRS will aim to identify provisionally **3 - 7** (10-20%) trusts as exceptional – the moderation process may, however, lead to the decision to award a number of Exceptional (4) grades outside this range. We also aim to identify the smallest number of trusts as exceptional e.g. if there are 3 trusts with the highest overall score and 2 trusts with the second highest score then only the top 3 will be rated provisionally as Level 4, even though the combined number is still within the target range.

Three scenarios occur:

1. The number of trusts achieving the highest overall score (n_t) is in this pre-determined range.
 - ⇒ All awarded exceptional, no further action needed.
2. n_t is more than the range.
 - ⇒ Provisionally all will be identified but the moderation panel will be provided with the differentiation data as discussed below
3. n_t is less than the lower end of the target range.
 - ⇒ Look at the number of trusts scoring the top two overall scores:
 - If the number of trusts scoring the top two highest overall scores ($n_t + n_{2t}$) is within the target range all are provisionally awarded exceptional and no further action is needed.
 - If this number ($n_t + n_{2t}$) is more than the upper end of the target range then the same procedure as for step 2 is used, with differentiation data supplied to the moderation panel.
 - If this number ($n_t + n_{2t}$) is less than the lower range of the target range then look at the trusts with the top three highest overall scores and apply the same three scenarios etc.

Under some circumstances (for example where more than the target maximum number of Trusts (20% = 7) have equal overall grading scores) an additional method to provide differentiation between those trusts with the same overall domain average will be used to provide the moderation panel with additional ranking data. This differentiation will be achieved by ranking all Trusts with equal domain averages in order of the sum of their overall domain scores. Using this data (and/or any other factors that the moderation panel elect to include) the panel will decide what level of overall performance equates to the Exceptional rating and which Trusts have achieved this.

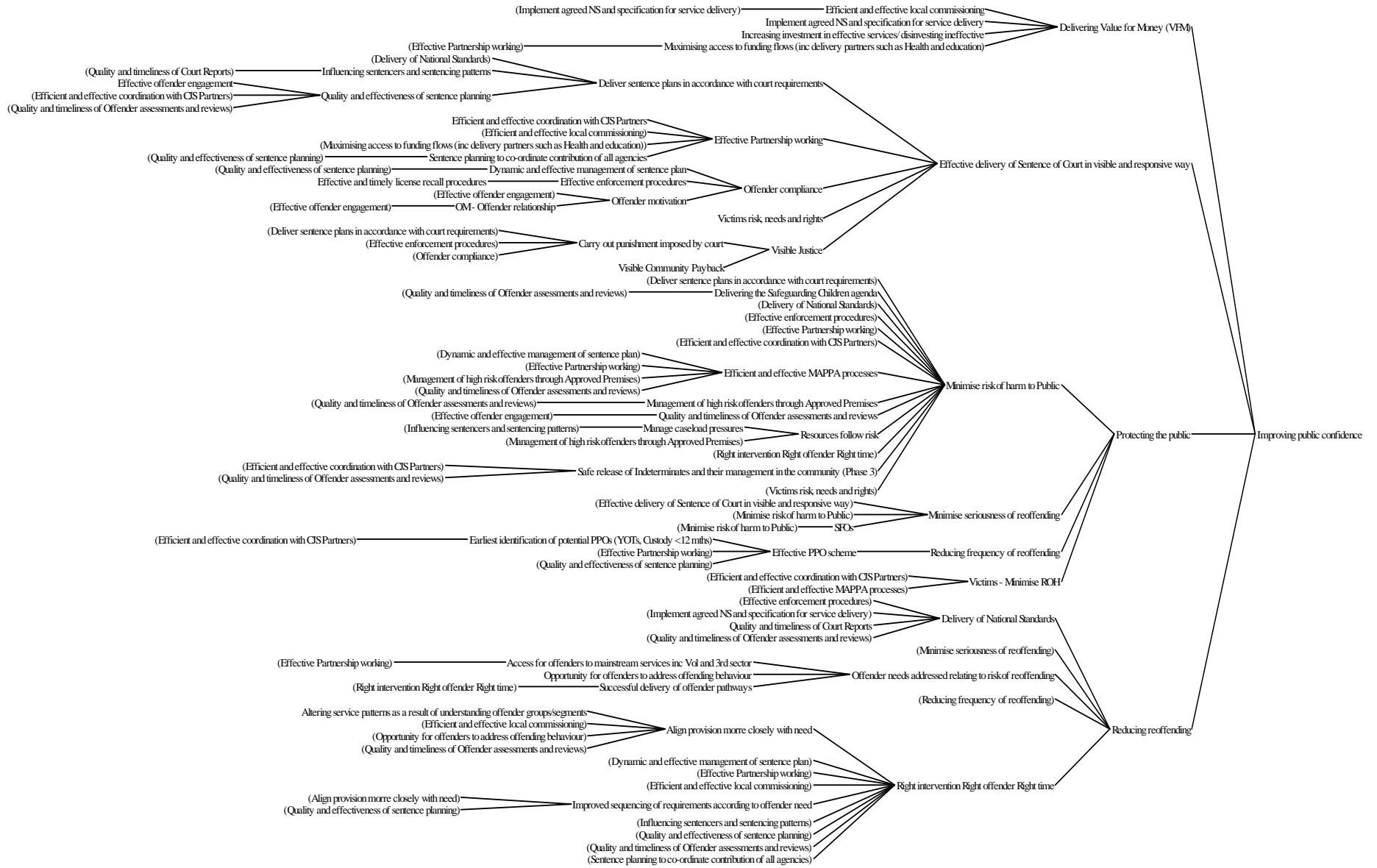
3.3 *Regional and National Aggregation*

Regional and national ratings will be calculated by taking a simple average over the ratings of all the trusts in a region at both domain and overall level.

To avoid regions with small numbers being 'bumped up' a rating, we will take each cut off point to be $> \#.5$, rather than $\geq \#.5$, so an average of 2.5 will correspond to a Level 2 rating rather than a Level 3.

The same process – a simple average for each domain and the overall ratings will be used to derive the national level domain and overall ratings.

Annex A – Driver Tree



Annex B – 13/14 Indicator Detail

| GTN Ref | Short Name | Measure | Outcome | Data Type | Target Type & Level | PTRS Rating Methodology |
|---------|---|--|--|---------------------------|---------------------|---|
| OM21 | Reduce Reoffending | Reduce the rate of proven re-offending whilst under the supervision of probation | Reducing Reoffending | Rolling Annual | Predicted rate | Significant difference from predicted rate |
| OM20 | Orders or Licences Successfully Completed | x% of orders or licences successfully completed | Effective Delivery of the Sentence | YTD | Locally set | Performance against target & national average |
| OM29 | Offender Feedback | X% of offenders surveyed have an overall positive perception of engagement | Effective Delivery of the Sentence; Reducing Reoffending | Annual (available for Q3) | National (67%) | Performance against target & national average (incl. sampling adjustment) |
| OM26 | OASys Quality | X% of OASys assessments are assessed as either "Sufficient" or "Good" on the OASys QA | Protecting the Public; Reducing Reoffending | Bi-annual (Q2 & Q4) | National (90%) | Performance against target & national average |
| OM41 | MAPPA Effectiveness | <p>This measure will comprise three elements:</p> <ul style="list-style-type: none"> • % of Category 2 MAPPA offenders managed at Levels 2 and 3 that appear on VISOR • % attendance by each probation SMB member at the SMB quarterly meetings • % attendance by the appropriate grade from the Probation Service at each level 2 and 3 MAPP meeting | Protecting the Public | Quarterly | Locally set | Performance against target & national performance |

| GTN Ref | Short Name | Measure | Outcome | Data Type | Target Type & Level | PTRS Rating Methodology |
|---------|------------------------------|---|-----------------------|-----------|---------------------|---|
| INT09 | Employment at Termination | The percentage of offenders in employment at termination of their sentence, order or licence to be at least x% (PSA 16) | Reducing Reoffending | YTD | Locally set | Performance against target & national average |
| OM17 | Accommodation at Termination | The percentage of offenders in settled and suitable accommodation at the end of their sentence, order or licence to be at least x% (PSA 16) | Reducing Reoffending | YTD | Locally set | Performance against target & national average |
| OM32 | Victim Feedback | X% of victims responding to survey are satisfied or very satisfied with service received | Protecting the Public | YTD | National (90%) | Performance against target & national average (incl. sampling adjustment) |

